The Sedona Conference Journal®

Volume 11 ❖ Fall 2010

Articles

The Sedona Conference®

Antitrust Law, Complex Litigation and Intellectual Property Rights

The Sedona Conference Journal®

Volume 11

Fall 2010

The Sedona Conference®

The Sedona Conference Journal® (ISSN 1530-4981) is published on an annual basis, containing selections from the preceding year's Conferences and Working Groups. The *Journal* is available on a complementary basis to courthouses and public law libraries and by subscription to others ($95; $45 for Conference participants and Working Group members). Send us an email (rgb@sedonaconference.org) or call (1-866-860-6600) to order or for further information. Complete Conference Notebooks for each of our Conferences are separately available. Check our website for further information about our Conferences, Working Groups, and publications: www.thesedonaconference.org.

Comments (strongly encouraged) and requests to reproduce all or portions of this issue should be directed to:
Executive Director, The Sedona Conference®,
180 Broken Arrow Way South, Sedona, AZ 86351-8998
(toll free) 1-866-860-6600 or (tel.) 928-284-2698;
fax 928-284-4240; email rgb@sedonaconference.org.

The Sedona Conference Journal® designed by
Margo Braman of Studio D: mbraman@sedona.net.

Cite items in this volume to "11 Sedona Conf. J. ____ (2010)."

Publisher's Note

Welcome to Volume 11 of *The Sedona Conference Journal*®
(ISSN 1530-4981), an annual collection of articles originally presented at our conferences and papers prepared by our Working Groups over the past year. The Sedona Conference® was founded in 1997 to provide a forum for advanced dialogue by the nation's leading attorneys, academics and jurists of cutting-edge issues of law and policy in the areas of antitrust, intellectual property rights, and complex litigation. We host Regular Season Conferences and several Working Group meetings each year, providing unique and rewarding opportunities to seriously explore the boundaries of various areas of the law with those who are creating them. This volume of the *Journal* contains articles selected from our most recent Conferences on antitrust law and litigation (Fall 2009), patent litigation (Fall 2009), complex litigation (Spring 2010), our Working Group on Electronic Document Retention and Production (WG1), and an article contributed in connection with an upcoming complex litigation conference.

We hope that you will find that the papers in this Journal reflect the same mix of theory and experience found at our Conferences and Working Groups, including the creativity and constructive irreverence required to challenge traditional thinking. The views expressed herein are those of the authors, and we encourage the submission of counterpoint pieces. Submissions can be sent to us electronically at rgb@sedonaconference.org, or by mail to The Sedona Conference®, 180 Broken Arrow Way South, Sedona, AZ 86351-8998, USA. If you are interested in participating in one of our Regular Season Conferences (limited to 45 participants in addition to the 15-person faculty, to encourage the dialogue that is our hallmark), or in joining one of our Working Groups, please visit our website for further information (www.thesedonaconference.org).

Richard G. Braman
Executive Director
The Sedona Conference®
September 2010

The Sedona Conference® gratefully acknowledges the substantial contributions of its Conference faculties, Working Group Series℠ Sustaining and Annual Sponsors, participants, members and observers, and our Advisory Board members, whose volunteer efforts and contributions make The Sedona Conference® a "thought-provoking and inspiring" experience providing content of immediate benefit to the Bench and Bar.

The Sedona Conference® Advisory Board

Table of Contents

REEXAMINATION PRACTICE WITH CONCURRENT DISTRICT COURT LITIGATION OR SECTION 337 USITC INVESTIGATIONS

Robert Greene Sterne, Jon E. Wright, Lori A. Gordon & Byron L. Pickard[1]
Sterne, Kessler, Goldstein & Fox, P.L.L.C.
Washington, DC

AUTHORS' NOTE

Patent reexamination was first selected as a topic for presentation at The Sedona Conference® on Patent Litigation in 2006. Version 1 of this paper was first published as part of that conference. The Sedona Conference's® on Patent Litigation in 2007, 2008 and 2009 each addressed reexamination and concurrent patent litigation, and subsequent versions of this paper accompanied those Sedona dialogues. Other versions accompanied presentations made at ACPC, IPO and PLI Conferences. Now in Version XI, it will accompany the Sedona dialogue on this topic that will take place on October 21, 2010, at the Sedona Patent Litigation Conference XI (2010).[2] In all versions, the authors address current procedure, process, and cutting-edge topics in reexamination practice and concurrent litigation. This paper subscribes to a neutral Swiss approach of presenting all sides of an issue and does not advocate for any particular view so that discussion may ensue. Many have provided comments and information for this article, including judges, senior officials from the PTO, Congressional staffers, patent owners, patent litigators, patent prosecutors, academics, bloggers and interested members of the public. Moreover, the authors devote substantial portions of their practices to reexaminations on behalf of patent owners and third party requesters and are on the editorial board of the foremost Internet site on reexamination, The Reexamination Center (www.reexamcenter.com). However, the views expressed herein are for purposes of dialogue and do not necessarily reflect the individual views of the authors.

1 Version 11. The authors thank Pauline Pelletier of their firm for the in-depth reexamination data research that is included in this version of the article.
2 http://www.thesedonaconference.org/conferences/20101021.

TABLE OF CONTENTS

I. INTRODUCTION

This paper addresses the interplay between patent litigation before the federal courts or the United States International Trade Commission ("ITC") (collectively, "the courts") and co-pending reexamination proceedings involving the patent-in-suit before the United States Patent and Trademark Office ("PTO"). As independent arbiters of patent validity and patentability[3], each forum poses a distinct set of challenges and risks for those challenging or defending patent validity. These so-called parallel universes use different rules, standards, procedures, time lines, and results in cases involving the same patent. High-profile cases involving reexaminations and co-pending litigation include *NTP, Inc. v. Research in Motion, Ltd.*[4] (patents found to cover the BlackBerry), *TiVo v. Echostar*,[5] (TiVo's DVR patents), *i4i v. Microsoft*, (patent covering XML functionality), *Uniloc v. Microsoft*, (patent covering anti-piracy protection), *Cordis v. Abbott,* (drug eluting stents). In another well-known case, Amazon's patent covering its "one-click" Internet shopping method was recently confirmed in reexamination.[6] These high-profile cases, some involving highly profitable products or large damage awards, highlight the critical interplay between the parallel universes of the courts and the PTO.

Two primary factors have contributed to the increased use of reexamination as an alternate or additional venue to challenge patent validity where district court litigation has been initiated. First, in 2005, the PTO streamlined reexamination by creating the Central Reexamination Unit ("CRU"), making it a more viable venue for post-grant validity challenges. The CRU's sole responsibility is handling reexaminations. The CRU's organization and initiatives are described more fully below. But, if the continued growth in the number of reexamination filings is an indication, its formation has put teeth into a process previously perceived as pro-patent owner and plagued by delay and uncertainty.

3 In reexamination, the PTO reviews an issued patent unpatentability. The courts decide the issue of patent validity. This distinction is important. For convenience, the authorities refer to these distinct issues collectively as a questions of validity.
4 *NTP, Inc. v. Research in Motion, Ltd.*, 418 F.3d. 1282, 1326 (Fed. Cir. 2005).
5 *TiVo, Inc. v. Echostar, et al.*, 446 F. Supp. 2d (E.D. Tex. 2006).
6 Reexamination No. 90/007,946 for U.S. Patent No. 5,960,411.

Second, the legal landscape surrounding patent validity has been in great flux. The Supreme Court's 2007 decision in *KSR Int'l Co. v. Teleflex, Inc.*[7] dramatically changed the applicable standard governing determinations of a claimed invention's obviousness, articulating a more flexible standard than the prior teaching-suggestion-motivation standard promulgated by the Federal Circuit. That decision alone appears to have cast serious doubt on the validity of many issued patents. Further, Federal Circuit decisions including *In re Translogic Tech.*[8] and *In re Swanson*[9] dramatically impacted and illuminated the legal landscape and brought the interplay between the courts and the PTO into sharper focus.

Patent infringement defendants and patent owners now recognize that defending patent validity before the CRU is a serious challenge. For this reason, reexamination filings have continued their rapid growth with no sign of slowing in 2010.[10] The impact of a potential reexamination is now commonly considered by both parties in nearly *every* patent litigation and ITC investigation. For that reason, patent owners contemplating a lawsuit must have a strategy in place in the event the accused infringer puts the asserted patent into reexamination. Similarly, every patent infringement defendant should consider the benefits of reexamination as an additional, perhaps more favorable, and less expensive venue in which to challenge patent validity. Finally, district court judges and ITC administrative law judges should be aware of how reexamination can impact their respective proceedings.

This paper is organized differently from a typical journal article. **Section I** presents hot topics and cutting-edge legal developments in reexamination law and practice. This first section presumes the reader is familiar with the use of reexamination as a viable post-grant venue for challenging patent validity. More in-depth treatment of certain of these hot topics is found in other parts of this paper. In **Section II**, we describe in more detail the parallel universes of the courts and the PTO. These two tribunals are substantively different in both scope and standard of review. Understanding the differences is vital to any strategic reexamination decision making. In **Section III,** we present advanced reexamination strategy considerations where federal court or ITC litigation is threatened or pending. Again, this section presumes basic knowledge of reexamination practice. We discuss timing of reexamination requests, the risk-versus-reward calculus, general reexamination tenets, and other, often overlooked, considerations. Finally, **Section IV** is a concise discussion of important aspects of basic reexamination practice. There, we discuss what is required to institute a reexamination, common pitfalls, the differences between *ex parte* and *inter partes* reexaminations, and other concerns. Readers not familiar with basic reexamination practice should review this section first. **Section V** provides a brief overview of the PTO's CRU, the current issues it faces, and recent initiatives to improve its core processes. **Section VI**, analyzes the most recent reexamination statistics, from both the CRU, and our own independent data collection and analysis.

---

7 *KSR Int'l Co. v. Teleflex, Inc.*, 127 S.Ct. 1727 (2007).

8 *In re Translogic Tech., Inc.*, 504 F.3d 1249 (Fed. Cir. 2008) (The Federal Circuit was presented with simultaneous appeal from BPAI and District Court on same patent. The Court decided the BPAI appeal first, upheld BPAI's ruling that the patent was invalid, and vacated Translogic's $85 million damages award from the district court.).

9 *In re Swanson*, 540 F.3d. 1368 (Fed. Cir. 2008) (The Federal Circuit clarified when it is appropriate to base a SNQ on art previously considered by the Office.).

10 One metric for growth is the increase in reexamination filings from one year to the next as demonstrated by average annual increases of 46.4% for *inter partes* and 6.5% for *ex parte* reexaminations from PTO fiscal year 2005 through 2009. Overall, the number of *inter partes* filings in 2009 was 437% of the number filed in 2005 and *ex parte* filings 126%. Straight lining the filings for fiscal year 2010 Q1-Q3 indicates a further increase in both proceedings.

II. HOT TOPICS AND NEW DEVELOPMENTS

The marked increase in the use of reexamination has naturally caused more frequent and closer evaluation of its unique procedures. In short, reexamination and its satellite issues (such as litigation stays, protective orders, nonobviousness evidence, duty of disclosure) remain among the faster developing areas of intellectual property law. Indeed, in the words of top PTO officials, it is still a "work in progress." We have identified a number of hot topics that are currently confronting parties finding themselves in a parallel universe. These hot topics were selected and are discussed with special attention to the perspective of the judges who manage the interface between the parallel universes. Our hot topics include:

(A) Reexamination Pendency
(B) Litigation Stays
(C) Protective Orders
(D) Substantial New Questions of Patentability
(E) Post-Grant Review Proposals
(F) Multiple PTO Proceedings Involving the Same Patent
(G) Appeals of *Inter Partes* Reexaminations to BPAI and Federal Circuit
(H) Impact of Reexamination on Remedies
(I) Impact of Reexamination and Court Decisions on Stock Price
(J) Impact of Settlement Agreements on Inter Partes Reexamination
(K) Impact of Reexaminations on Willfulness and Inequitable Conduct, and
(L) Retained Reexamination Experts

A. Reexamination Pendency

Reexamination pendency is a perennial hot topic for a number of reasons. First, for reexamination to be an effective post-grant venue for challenging patent validity, it must be concluded in a timely, efficient manner—with the "special dispatch" required by statute. To avoid prejudice to patent owners, however, the PTO must act in a deliberate, fair manner by giving the patent owner a full opportunity to defend the patentability of the claims. Second, reexamination pendency is an important factor weighed by the courts when deciding whether to stay concurrent litigation. Third, predictable reexamination timelines are enormously helpful for third party requesters and patent owners when they work reexamination strategies into parallel district-court-litigation and ITC-investigation timetables. In short, uncertainty in pendency of a reexamination proceeding diminishes the fairness and effectiveness of reexamination and litigation for all of the involved parties.

PTO statistics on reexamination pendency are available, and the authors have done their own statistical analysis on reexamination pendency. The results are presented in Section IV.A. However, general statistics are useful only up to a certain point. To perhaps remove some of the uncertainty and provide a more nuanced analysis of reexamination pendency, we share below some general observations on pendency and illustrate how different reexamination milestones can result in widely varying timelines from as little as three months to periods well in excess of six years.

There are several early reexamination milestones from which important clues on pendency may be ascertained: (1) the request, (2) the reexamination order, and (3) the first Office action.

The Request

The request itself can provide a number of early clues to potential reexamination pendency. For example, for how many claims has reexamination been requested—all claims, or only a limited subset of claims such as only those asserted in litigation? How many separate substantial new questions of patentability ("SNQs") SNQs are alleged? What is the nature of the proposed rejections—do they primarily allege anticipation, or is obviousness in play? How many different references are used for the proposed rejections? Is the request *ex parte* or *inter partes*? Not surprisingly, the CRU is generally slower to act on complicated reexamination proceedings than it is to uncomplicated ones. The requester has complete control of these decisions and can therefore assert some control over likely pendency from the outset.

The Reexamination Order

The reexamination order, which is the first opportunity to see how the CRU received the reexamination request, is the second big pendency milestone. By its own rules, the CRU should rule on whether to initiate a reexamination within three months of the request. The CRU has complete discretion in this regard and a number of scenarios are possible, all of which have a direct impact on reexamination pendency.[11] For example, the reexamination request can be denied, granted in full, or anything in between. Clearly, an outright denial—a finding of no SNQ—does not bode well for the requester and can lead to a very short pendency. On the other hand, full or partial adoption of the alleged SNQ's indicates the CRU was at least persuaded that the reexamination should move forward.

The First Office Action

The first Office action will also have many clues to potential reexamination pendency. For *ex parte* reexamination, the first Office action cannot arrive until after the period for patent owner response has expired.[12] For *inter partes* reexamination, PTO procedures dictate that the reexamination order be accompanied by the first Office action on the merits.[13] Just because the CRU has adopted all or part of the alleged SNQs does not necessarily imply that a full-blown reexamination is in the works. For instance, the CRU is well within its authority order a reexamination, and then in the first Office action indicate that one or more of the claims under reexamination are allowable. This bodes well for short pendency. On the other hand, full adoption by the CRU of the proposed grounds of rejection indicates longer pendency. As with the reexamination order itself, the first Office action will provide the parties with a decent idea of how receptive the CRU is to the reexamination request. Most district court judges will not consider staying a concurrent litigation at least until the CRU has ruled on whether to initiate the reexamination. And no stay decision should be made without at least considering the details of the reexamination order and first Office action.

For *inter partes* reexamination, these first three milestones should all occur no later than three months after the initial request for reexamination. For *ex parte* reexaminations, these milestones should occur five to six months after the reexamination request. These first three milestones thus occur relatively early in the reexamination process. And they go a long

11 There are only limited mechanisms for challenging the PTO's decision to order (or not order) a reexamination.
12 37 C.F.R. § 1.530(b).
13 37 C.F.R. § 1.935.

way towards determining overall scope, level of involvement and complexity of the reexamination. It therefore behooves all parties (including the court) to wait for these early milestones before investing heavily in any previously predicted reexamination pendency.

Other Reexamination Milestones

If it appears that the reexamination will proceed on the merits, then there are a number of other milestones that will impact overall pendency. The involved parties, including the court, should recognize these milestones and be nimble enough to take them into full consideration. These milestones include the PTO's response to the patent owner reply (and requester's comments thereto in *inter partes* reexamination). In *ex parte* reexamination, this response takes the form of either a final Office action or a Notice of Intent to Issue a Reexamination Certificate, which is commonly referred to as a "NIRC." For *inter partes* reexamination, this response takes the form of an Action Closing Prosecution ("ACP"). The CRU has a stated goal to get to this point in the reexamination proceeding inside of two years from the filing date of the request. As with the first Office action above, this milestone provides the next set of clues as to reexamination pendency.

As a rule, *ex parte* reexaminations proceed more quickly from this point simply due to the *ex parte* nature of the proceeding. If the decision is favorable to the patent owner—such as a NIRC or indication that some claims are allowable—reexamination pendency will likely be shortened. However, if the patent owner elects to appeal all or part of the decision, pendency will be lengthened by at least another one and a half to two years.

For *inter partes* reexamination, on the other hand, the ACP and subsequent Right of Appeal Notice ("RAN") typically signal an appeal. This is due to the *inter partes* nature of the proceeding—the patent owner may appeal any final rejection of any claim, while the requester can appeal any decision favorable to patentability, including the CRU's refusal to adopt a presented SNQ, or refusal to maintain any proposed ground of rejection. Thus, unless one of the parties surrenders or is precluded from participation due to settlement terms, *inter partes* reexaminations routinely involve at least another one and a half to two years pending appeal to the BPAI. For both *ex parte* and *inter partes* reexamination, appeals to the Federal Circuit are an option.

As described in Section II.G. below, the time period for appeals of a reexamination proceeding from the CRU to the BPAI and then to the Federal Circuit can vary, but at a minimum will take three years with current backlogs and processing requirements. With *ex parte* reexaminations, only the patent owner can appeal a rejection and the third party requester has no standing to appeal a favorable decision concerning the claims in reexamination. Since the NIRC and the reexamination certificate cannot issue for the patent owner in an *ex parte* reexamination unless the patent owner is satisfied with the CRU decisions concerning all of the claims in the reexamination, the patent owner is forced to appeal the rejection of a single claim even though all of the other claims in reexamination are confirmed or allowed. This all or nothing aspect of the reexamination process can force appeals and is unlike original prosecution where allowed claims can be issued in a patent and rejected claims can continue to be prosecuted in a pending application. All in all, however, the *ex parte* reexamination proceeding is faster on appeal than *inter partes* reexamination because only the patent owner has standing.

The appeal process for *inter partes* reexamination proceeding is also described in Section II.G. below. This appeal process can easily take more than four years (through the Federal Circuit) even though the PTO is trying to address the various delays.

Courts need to assess carefully what is likely in the appeal process in coming to an studied estimate of the time of the reexamination proceeding. Blanket pendency statements often made in stay motions and in arguments before judges need to be challenged under the given circumstances and status of the actual reexamination proceeding in question.

Summary

In summary, reexamination pendency must be considered with an informed eye. Mere review of published PTO statistics does not provide a complete picture of reexamination pendency. Moreover, the past is not necessarily indicative of the future. Stay decisions are not (and should not be) routinely entertained in advance the first three milestones. And thereafter, any stay (or denial of stay) should be flexible enough to respond to subsequent events and milestones at the PTO. Finally, strategy decisions should take into consideration best- and worst-case scenarios, and be nimble enough adapt when the unexpected occurs.

B. Litigation Stays

Requesting a stay of a litigation following the filing of a reexamination request is a common strategy employed by accused infringers seeking to delay and perhaps avoid litigation. The grant of a stay is at the sole discretion of the presiding judge. Due in part to substantial uncertainty and confusion in reexamination pendency, the recent trend in many leading patent litigation jurisdictions is against the grant of stays.

Litigation stays are usually contested, with accused infringers typically arguing for the stay of the litigation and patent owners arguing against the stay. Accused infringers often point to the reexamination as a procedure that has the potential to avoid litigation completely, such as where the reexamination cancels all asserted claims, or has the potential to streamline validity issues. Patent owners, on the other hand, often argue that the pendency of the reexamination proceeding, feasibly lasting for more than six years, makes any stay highly prejudicial to patent enforcement. They argue the famous legal maxim: "justice delayed is justice denied."[14]

District court judges have broad discretion to decide a motion to stay. [15] The decisions are highly fact specific and vary greatly by judge and jurisdiction. Recent district court stay decisions highlight the varied factors that drive the results. For example, the United States District Court for the District of Minnesota denied a defendant's motion for a stay, finding the case would not be streamlined "because defendants never alleged invalidity as a defense in their answer nor did they disclose any such prior art by [the applicable deadline], as required by the pretrial scheduling order."[16] In a case in the Central District of California, the court granted a motion for stay citing as a relevant factor that the patent owner was not exploiting the patent-in-suit.[17] We discuss some of the more nuanced stay decisions in Section IV.C.

---

14 Attributed to William E. Gladstone—Laurence J. Peter, PETER'S QUOTATIONS, at 276 (1977).
15 *Allergan Inc. v. Cayman Chem. Co.*, No. 07-01316 (C.D. Cal. Apr. 9, 2009).
16 *Blazek Sklo Poderady v. Burton International,* No. 08-2342 (D. Minn. June 11, 2010). In this case, the plaintiff, Blazek, asserted two patents directed to glass nail files against defendants. Instead of preparing a defense on the merits and engaging in discovery in the district court, the defendants pinned their hopes entirely on an *ex parte* request for reexamination that cited prior art used to invalidate a Czech-issued patent related to the patents-in-suit.
17 *Pitney Bowes Inc. v. Zumbox Inc.*, No. 09-07373 (C.D. Cal. May 20, 2010).

For many judges, the single most important factor in determining whether to stay a litigation is the pendency of reexamination proceedings. Many judges will not grant a stay because of the length of time required to conclude a reexamination proceeding through all appeals. Creating confidence in a reasonable reexamination timeline can simplify the stay calculus for district court judges. For instance, if reexamination pendency is completely uncertain, or if the time to final decision extends years past an expected trial date, then stays are less likely as they may prejudice the patent owner. Conversely, if the reexamination is likely to conclude shortly, substantial resources may be saved by issuing a stay. Some courts are not dissuaded by the prospect of long reexamination proceedings, or reexamination proceedings of uncertain length, if the reexamination proceeding is ordered early in or prior to the litigation.[18]

Many practitioners believe that a sufficient number of factors exist to support either granting or denying at least a temporary stay in any given case. Because stay decisions are reviewed under an abuse-of-discretion standard, district court judges have almost unfettered authority. Nonetheless, more nuanced stay decisions are becoming the norm as ever more authority is developed on this issue. Practitioners and judges alike should be aware of all factors that are in play including the efforts by the PTO to address concerns over reexamination pendency. Reexamination stays are discussed more fully below in Section IV.C.

C. Protective Orders

A protective order dictates how confidential documents produced during a litigation are handled by the parties. How should parties craft a protective order in a concurrent district court litigation or ITC investigation to prepare for a possible reexamination proceeding at the PTO? Is it possible for a patent owner to satisfy its duty of disclosure while adhering to the guidelines of a protective order? What limitations does a protective order place on the resources available to a patent owner to prosecute the reexamination proceeding and the suit? What mechanisms are available to provide information of non-obviousness covered by the protective order in the reexamination proceeding? These issues and more are discussed in detail below in Section IV.D. In this Hot Topics section, we alert the reader to some pressing issues of which the authors are aware.

General Prosecution Bars May Be Insufficient

When crafting a protective order ("PO"), it is imperative for both parties to consider the possibility of a concurrent reexamination. The court's standard PO may include a general prosecution bar, or the parties may be relying on previous templates that include a general prosecution bar. Reexamination was not a significant issue in patent litigation in the recent past. Unfortunately, such prosecution bars typically refer only to "patent applications." But a reexamination proceeding is not a patent application and it unclear whether such a prosecution bar would apply to reexamination proceedings. Whether either of the parties are contemplating reexamination or not, the PO should explicitly deal with reexamination proceedings.

18 *Sorensen v. Spectrum Brands, Inc. et al.,* No. 09-00058 (S.D. Cal. Dec. 23, 2009). One newer tactic on the defense side is to file the reexamination before suit is filed in lieu of filing a DJ. Then the third party requester uses the pre-suit filing of the reexamination as a basis to seek a stay.

The PO Must Balance Competing Concerns

The parties will have competing concerns. The accused infringer will be concerned that its confidential information may be used to amend existing claims or craft new claims to better cover accused products. This concern is ameliorated to some degree by the fact that claims may not be broadened during reexamination beyond the claims in reexamination, so the concern is not as great as it is with a patent application with no such restrictions. The patent owner, on the other hand, must be able to freely communicate its trial strategy, including invalidity and claim construction positions, to its reexamination counsel lest the two teams inadvertently take inconsistent positions. This is especially a concern given the different claim construction regimes used by the PTO and the court. The accused infringer is not so restricted because it has no right to amend the claims in reexamination. Both parties may also need to rely on material produced by the other side over the course of discovery in support invalidity (or patentability). One example is evidence related to secondary considerations (objective indicia) for nonobviousness such as commercial success, copying, failures, and long felt need (or lack thereof). The PO must balance these legitimate competing concerns.

In short, the authors have seen a wide spectrum of POs. Some are overly restrictive and effectively wall off trial counsel from reexamination counsel on the patent owner side, putting the patent owner at severe risk of taking inconsistent positions between the two proceedings. Some are not restrictive enough, putting no restraints on trial counsel's participation. Some altogether fail to deal explicitly with reexamination, thereby leaving the issue open to further dispute, resolution of which may waste valuable court resources. As noted, we discuss POs in more depth below in Section IV.D.

D. Substantial New Questions of Patentability

Fundamental to every reexamination request is the substantial new question of patentability or the "SNQ." The SNQ is the cornerstone of patent reexamination and every reexamination request—both *inter partes* and *ex parte*—must include at least one SNQ. The purpose of the SNQ requirement is to create a threshold for reexamination to prevent serial reexamination proceedings on the same references, and to prevent harassment of the patent owner.[19] The SNQ requirement was included in the original 1980 *ex parte* reexamination statute as "a balance between curing allegedly defective patents [via reexamination] and preventing harassment of patentees."[20] While it may sound relatively simple, presentation of a SNQ is more subtle than many practitioners realize. This section remains a Hot Topic because it is an often misunderstood area of reexamination practice that is ripe for litigation—where the parties challenge the existence of a SNQ in a reexamination request through the Administrative Procedures Act or through appeal to the Federal Circuit.[21]

Review of SNQ Determination

On June 25, 2010, the PTO published a notice in the Federal Register "clarifying the procedure for seeking review of a determination that a substantial new question (SNQ) has been raised in an *ex parte* reexamination proceeding."[22] The notice announced that

19 H.R. Rep. No. 96-1307 (1980).
20 H.R. Rep. No. 107-120, at 1; *see also In re Recreative Technologies,* 83 F.3d 1394, 1397 (Fed. Cir. 1996).
21 The ability of a party to challenge the adoption or failure to adopt a SNQ is limited to *ex parte* reexamination proceedings. By statute, in *inter partes* reexaminations, a determination by the USPTO that no SNQ has been raised or that a reference raised a SNQ is final and non-appealable. 35 U.S.C. § 312(c).
22 75 Fed. Reg. 36357 (June 25, 2010).

Director Kappos delegated to the Chief Judge of the BPAI the authority to review issues relating to the grant of a request for reexamination, specifically a SNQ finding. The Chief Judge can further delegate the SNQ review to a panel of Administrative Patent Judges who are deciding the appeal in the *ex parte* reexamination proceeding. As is currently the case, the notice identifies that separate from the BPAI's consideration of the SNQ issue, a patent owner may file a petition under 37 CFR 1.181(a)(3) to vacate an *ex parte* reexamination as "ultra vires."

This right of appeal is not interlocutory. Instead, for an *ex parte* reexamination ordered before June 25, 2010, the patent owner may include a challenge to the finding of a SNQ as a separate ground in an appeal to the BPAI, even if the patent owner did not request that the CRU reconsider the finding of the SNQ during prosecution. For *ex parte* reexaminations proceedings ordered on or after June 25, 2010, the patent owner must request that the CRU reconsider the grant of the SNQ as part of a full response to the Office action in order to preserve the issue for appeal. Notably, the procedures do not apply to *inter partes* reexamination proceedings. A determination by the PTO in an *inter partes* reexamination either that no SNQ has been raised or that a reference raises a SNQ is final and non-appealable.[23]

A recent BPAI decision issued June 30, 2010, cited the clarification claiming jurisdiction to review a CRU examiner's SNQ finding.[24] Here the appellant argued that a particular reference was previously considered during original prosecution and thus not qualified to form the basis of a SNQ. One issue addressed by the BPAI was whether it had jurisdiction to review a CRU examiner's SNQ finding (and order granting a request for reexamination). Citing the June 25, 2010, clarification as the source of its authority, the BPAI sided with the examiner that during original prosecution the reference was not reviewed with "any reasonably detailed analysis," nor was it applied as a principal prior art reference, nor was the issue forming the basis of rejection during reexamination the same as that examined during original prosecution. Thus the reference was not "old" art and raised a valid SNQ. The panel further noted that the appellant did not timely file a 1.181 petition to review the original determination and instead brought the issue up on appeal after the claims had been finally rejected on a variety of grounds including lack written description, indefiniteness, anticipation, and obviousness. The panel affirmed the examiner's determination on all remaining issues.

Burden to Establish a SNQ

Congress envisioned the SNQ as a primary gate-keeping function to prevent harassment of a patent owner.[25] Many practitioners argue that this protection is illusory due to the implementation of the SNQ review by the PTO. Specifically, practitioners point to the lack of evidentiary burden placed on the requester to establish a SNQ in the request.

To establish a SNQ, the requester must demonstrate that a patent or printed publication presents a new, non-cumulative technological teaching that was not previously considered and discussed on the record during the prosecution of the application that resulted in the patent for which reexamination is requested, and during the prosecution of

---

23 *See* 35 U.S.C. 312(c).
24 *See* BPAI Appeal No. 2008-005992, Reexamination Control No. 90/006,572 (June 30, 2010).
25 "As part of the original 1980 reexamination statute, Congress struck a balance between curing allegedly defective patents and preventing the harassment of patentees. It adopted a standard requiring a request for reexamination to raise a 'substantial new question of patentability.'" *See* H.R. Rep. No. 107-120, at 1; *See also In re Recreative Technologies*, 83 F.3d 1394, 1397 (Fed. Cir. 1996).

any other prior proceeding involving the patent for which reexamination is requested.[26] Additionally, the reexamination request "***must*** point out how any questions of patentability raised are ***substantially different*** from those raised in the previous examination of the patent before the Office."[27]

Thus, the burden is placed on the third party requester to demonstrate that the questions of patentability raised are ***substantially*** different than those raised in previous examinations. However, many practitioners argue that the PTO accepts statements that a patent or printed publication presents a new, non-cumulative teaching without requiring any evidence or discussion of the prior prosecution record to support the statement. This practice effectively shifts the burden to the patent owner to disprove the existence of a SNQ adopted by the PTO in an *ex parte* reexamination. In an *inter partes* reexamination, the patent owner is left with no recourse for challenging that determination.

E. Post-Grant Review Proposals

Original patent claims in both *ex parte* and *inter partes* reexamination proceedings are only examined on the basis of patents and printed publications.[28] The full suite of defenses available in patent litigation, including statutory subject matter (35 U.S.C. § 101), written description and enablement (35 U.S.C. § 112), public use, prior sale, and inequitable conduct, are not available in reexamination proceedings. This disparity between invalidity challenges available in a district court and the patentability challenges in reexamination has spawned efforts to expand the available bases for requesting reexamination to create more equivalence between the two forums. The PTO takes the position that the reexamination statute does not permit such an expansion.[29] The CRU has therefore strictly enforced the narrow basis for reexamination of patent claims to rejections based on patents and printed publication. Under this regime, collateral prior art defenses—commercialization activities surrounding a printed publication—cannot be considered in reexamination even though such defenses can be considered in any parallel court litigation under prior art categories such as "on sale" or "public use."

The efforts to achieve identity between the invalidity challenges available in district court and those available in reexamination have centered primarily on changes to the underlying statute and the introduction of a new post grant review process. This post-grant review ("PGR") process has been proposed in several prior versions of Patent Reform legislation; and as of the publication of this paper, Senate Bill S515 contains language for a first-window post-grant review. Under the current version of the bill, the first-window concept refers to the time period from when the patent issues to when the period for filing a PGR would close—12 months in the current draft. The PGR would supplement, not replace reexamination review. As part of the legislation, *inter partes* review would be made available for all unexpired patents.[30] The PGR would be an adversarial two-party process in the PTO with expanded grounds for challenging validity of a patent beyond the limited patents and printed publication available in reexamination. It is also contemplated that the proposed *inter partes* review would be an adversarial two-party process in the PTO. Limited discovery would be available and oral testimony before administrative law judges ("ALJs") would be allowed. Under the current view, the ALJ would rule in a summary-judgment

26 MPEP § 2216.
27 MPEP § 2216. (emphasis added).
28 35 U.S.C. §§ 302, 311; 37 C.F.R. §§ 1.510(a), 1.915(b).
29 *See* 37 C.F.R. §§ 1.552, 1.906; M.P.E.P. §§ 2258, 2658.
30 Currently, *inter partes* reexamination can only be filed for a patent issued from an original application filed on or after November 29, 1999. *See* 37 C.F.R. § 1.913; *See also* M.P.E.P. § 2611.

style. The goal of the PGR and the *inter partes* review would be a faster, better, and cheaper alternative to the courts for decision on the validity of issued patent claims. The PGR concept as currently envisioned has generated considerable controversy from many stakeholder groups who question whether the professed goals are attainable and whether the PGR would actually improve the patent system and protect legitimate patent rights. Finally, the existing *ex parte* reexamination procedure would still exist, and current *inter partes* reexamination proceedings would continue to be conducted under the present rules.

Other efforts focus on expanding the examination performed by the CRU, after a request is granted. As discussed above, the PTO does not address rejections of original patent claims beyond those based on patents or printed publications. The premise for not addressing these grounds for invalidity, particularly statutory subject matter or 35 U.S.C. § 112 written description and enablement challenges, is that the original patent claims were found patentable under these statutory provisions during prosecution and therefore the claim status under these provisions has not changed. However, some practitioners argue that decisions by the Federal Circuit and Supreme Court occurring after allowance of the original patent claims undermine this reasoning. As an example, the Supreme Court, in the recent *Bilski v. Kappos* decision, set forth a more flexible standard for determining patentable subject matter, and the PTO is in the process of implementing guidelines for this more flexible standard.[31]

As a result of these cases, the authors predict increased pressure on the PTO and Congress to introduce statutory-subject-matter challenges, at a minimum, into reexamination proceedings. Although statutorily barred from being included in the reexamination request, some practitioners argue that the PTO could, on its own, raise these issues during reexamination, in the same manner that is done during original prosecution. Although potentially allowable under the current statute, this approach would require a change to the PTO rules and procedures. Under current PTO procedures, amendments to the original patent claims open the door to these additional invalidity challenges—but only against the amendments.

F. Multiple PTO Proceedings Involving the Same Patent - Merger and its Impact on Reexamination Proceedings

If the parallel universes of having the same patent subject to review in reexamination and court litigation seems complicated enough, the picture can become even more multi-dimensional if the patent is also subject to additional reexamination, reissue or interference proceedings. There are several scenarios possible under current practice. Another *ex parte* or *inter partes* reexamination can be filed on the same patent.[32] A reissue application, whether broadening or narrowing, can also be filed. Additionally, the patent in reexamination can be involved in an interference proceeding with another patent application. Under present practice, it is conceivable that a single patent could be involved in all of these proceedings concurrently.[33]

Multiple proceedings can have a significant impact on pendency, cost, and complexity of the reexamination.[34] Such multiple PTO proceedings involving the same

31 *Bilski v. Kappos*, 130 S.Ct. 3218 (2010).
32 If an *inter partes* reexamination has already been filed, the estoppel provisions of 35 U.S.C. § 317 prevent the third party requester or its privies from filing a subsequent reexamination request.
33 If proposed S515 is enacted, this patent could also be subject to PGR and *inter partes* review as well. Some critics argue such a regime will effectively make enforcement of a patent prohibitive or impossible.
34 Current legislation contained in S515 in the U.S. Senate could expand this complexity with the creation of a "first window" post grant review process and creation of an *inter partes* review process that will be available for all unexpired patents.

patent are not rare. In fact, multiple reexaminations of the same patent or a reexamination with a parallel reissue application are seen frequently, especially if the patent is perceived as being very valuable or is part of a hard fought litigation. Since 2000, only 2,560 unique patents have been involved in 5,680 reexamination proceedings at the CRU as of January 1, 2010.[35] Multiple reexamination proceedings on the same patent explain why the overall ratio of reexaminations to patents is over two to one. Specifically, of these 2,560 unique patents, 286 or 11% have been reexamined more than once, one as many as six times.[36] Three patents have been reexamined five times, 14 reexamined four times, 34 reexamined three times, and 234 reexamined twice. In Section V.F. below, the practice issues involving these parallel proceedings are examined in more detail.

Should a patent owner or third party requester initiate an additional concurrent proceeding? There are several important political and strategic perspectives to consider before initiating another concurrent proceeding. For instance, certain types of proceedings may be merged by the PTO. For example, multiple reexaminations of the same patent are typically merged into one proceeding. Because of the nearly certain merger, some practitioners have adopted the practice of first filing an *ex parte* reexamination and then following soon after with an *inter partes* reexamination. The *ex parte* reexamination typically involves fewer claims and fewer SNQs than the later filed *inter partes* reexamination. Some argue that this strategy allows the requester to obtain a quick filing date (to enhance the possibility of a stay from the court) and prevents a stay of the *inter partes* reexamination after merger due to the presence of the *ex parte* issues. Additionally, *ex parte* interview rights are typically extinguished after merger of the proceedings. But the Office of Patent Legal Administration ("OPLA") may be revisiting its merger rationales for this requester-initiated multiple reexamination strategy.

Other practitioners file multiple *ex parte* reexaminations (alone or in combination with an *inter partes* reexamination) on the same patent. This "rolling reexams" approach is not prohibited by statute or by the rules. Because the multiple reexaminations are typically merged, patent owners argue that rolling reexaminations delay conclusion of the reexamination proceeding, undermining their statutory right to a reexamination proceeding being handled with "special dispatch," waste valuable patent term, can effectively turn an *ex parte* proceeding into an *inter partes* proceeding, and could be perceived as harassment of the patent owner. There are instances where the PTO has denied second *ex parte* reexamination requests as not raising a SNQ.

Patent owners faced with multiple reexaminations on the same patent that result in mergers are not without recourse. There is sufficient ambiguity in the merger rules and sufficient statutory bases to allow patent owners to argue against merger in some situations. Patent owners can challenge merger decisions through petition practice and bring any perceived unfair application of the merger rules, or perceived harassing behavior, to the attention of OPLA, which handles many petitions in reexamination matters on behalf of the Director. Because the PTO has complete discretion in merger decisions, it can (and has) dissolved previously merged cases.[37] In the authors' experience, OPLA is receptive to well-crafted arguments and is willing to reconsider previously ordered mergers.

Reissue applications may also be merged with a co-pending reexamination proceeding. The merged reissue/reexamination proceedings are often transferred to the

35 Requests for Reexamination Noticed in Official Gazettes 1999-2010 (Jan. 1, 2010).
36 *See* Broadcom Corporation Patent No. 5,425,051 (three *ex parte* proceedings were requested by Qualcomm in 2006 and three more *ex parte* proceedings in 2008).
37 *See, e.g.*, Reexamination Control Nos. 95/001,205 and 90/009,370.

technology center for handling, rather than the CRU. Some practitioners argue that this strategy creates an unfair advantage by allowing the patent owner to prosecute the reexamination in a technology center that may be friendlier to the patent owner and that operates under different performance metrics for examiners. Practitioners also argue that this technique is used to introduce delay into the PTO proceedings. The authors have been told that the PTO is constantly reviewing its merger procedures for reissues and reexamination to remove the potential for "gaming the system."

A patent being reexamined may also be involved in an interference proceeding with at least one application.[38] The general policy of the PTO is that the reexamination proceeding will not be delayed, or stayed, because of an interference or the possibility of an interference.[39] In a concurrent reexamination and interference, the PTO follows the practice of making the required and necessary decisions in the reexamination proceeding and, at the same time going forward with the interference to the extent desirable.[40] Any party to the interference may make a miscellaneous motion under 37 C.F.R. § 41.121(a)(3) to suspend an interference pending the outcome of a reexamination proceeding.[41] Similarly, a party to the reexamination proceeding may file a petition to stay the reexamination proceeding because of the interference.[42]

G. Appeals of **Inter Partes** *Reexaminations to BPAI and Federal Circuit*

Not a single *inter partes* reexamination appeal from the BPAI to the Federal Circuit has been decided to date. The first *inter partes* reexaminations appealed to the Federal Circuit were voluntarily dismissed by the appellant prior to briefing and oral argument.[43] While a small percentage of *inter partes* reexamination certificates have issued from the CRU,[44] a search of BPAI decisions from July 1, 1997, to July 1, 2010, reveals that the BPAI has decided only 29 *inter partes* reexamination appeals to date.[45] Significantly, seven of these decisions resulted in a non-final, non-appealable decision, and were therefore remanded to the examiner for further prosecution.[46]

One reason for the BPAI's inability to issue final decisions seems to be a lack of clarity in both the rules and statutes for the role of the BPAI in reviewing CRU decisions. More specifically, it is not clear whether the BPAI should act as a judge between patent owner and the third party requester, or whether it should sit as a "super examiner," the role it often takes in appeals.[47]

---

38 *See* M.P.E.P. §§ 2284, 2686.02.
39 *See* M.P.E.P. §§ 2284, 2686.02.
40 *See* M.P.E.P. §§ 2284, 2686.02.
41 *See* M.P.E.P. §§ 2284.II, 2686.02II.
42 *See* M.P.E.P. §§ 2284.V, 2686.02IV.
43 *Cooper Cameron v. SAFOCO Inc.*, Reexamination Nos. 95/000,015 and 95/000,017, Nos. 2009-1435, -1459 (Federal Circuit).
44 A Reexamination Certificate is issued at the completion of reexamination canceling any claim of the patent finally determined to be unpatentable, confirming any claim determined to be patentable, and incorporating any amended or new claim determined to be patentable.
45 FOIA Reading Room and Database of Final Decisions Issued by the Board of Patent Appeals and Interferences http://des.uspto.gov/Foia/BPAIReadingRoom.jsp (Last Accessed July 1, 2010).
46 *See* Reexamination Control Nos. 95/000,006 (USPN 6,357,595); 95/000,009 (USPN 6,399,670); 95/000,030 (USPN 6,508,393); 95/000,038 (USPN 6,527,941), 95/000,064 (USPN 6,767,487); 95/000,066 (USPN 6,789,673); and 95/000,069 (USPN 6,789,673).
47 It has been posed to the authors that the reason no *inter partes* case make it from the BPAI to the Federal Circuit is due (1) to the very small number of *inter partes* reexaminations filed in the early years of the statute; and (2) to the practical reality that the *inter partes* reexamination process is a "work in progress" and like all such endeavors, it takes time to work out the kinks and "get it right." Both of these observations appear correct but belie the practical consequence that the goals of *inter partes* reexamination of being faster, better, and cheaper that the district courts in assessing patentability based on patents and printed publications have failed so far in practice.

The Chart entitled "Procedure Following Board Decision for Reexaminations Commenced On or After November 2, 2002" in MPEP § 2601.01 graphically shows at least one remand problem that currently exists. In this example, the remand from the BPAI to the CRU occurs where a proposed ground of rejection in the reexamination request is not adopted by the CRU examiners in the first Office action. On appeal to the BPAI the third party requester raises this as a "new ground of rejection" that should have been made by CRU. The BPAI agrees but does not have the factual and argument record to decide the rejection and thus remands the reexamination back to the CRU to address with the parties. This remand may occur for other reasons or may occur a second time. It is the specter of multiple remands that has created the concern that *inter partes* reexamination may be fundamentally flawed.[48]

Another possible reason for the lack of cases making it to the Federal Circuit, is that, for a certain small subset of *inter partes* reexaminations, the process is effective without the need to appeal to either the BPAI or the Federal Circuit. According to recent PTO statistics, 167 *inter partes* reexamination certificates have issued out of the 923 total requests filed as of June 30, 2010.[49] While this represents only 18% of total *inter partes* reexaminations, the outcome data indicates that third party requesters succeeded in having all claims cancelled or disclaimed in 49% (82) of the completed *inter partes* proceedings in which a reexamination certificate issued. In 43% (71) claims were changed in some way and in only 8% (14) did all claims survive reexamination unamended.[50] The high success rate is skewed by the significant non-response rate by the patent owner either after a first Office action or after actions later in prosecution. Because a large number of *inter partes* reexaminations are also involved in co-pending litigation, this data might simply reflect the cases in which the parties have settled or in which the district court has reached a determination regarding validity.

The practical effect of this uncertainty appears to be that the BPAI remands at least some reexaminations, especially *inter partes*, back to the CRU.[51] Based on the authors' interviews with senior BPAI and CRU personnel, we know the PTO is aware of this issue and is working to assess the issue and perhaps to devise changes that take into consideration the policy goals behind *inter partes* reexamination. Specifically, the PTO is seeking an approach that recognizes that *inter partes* reexamination should not be used as a tool to harass patent owners, but rather should function as a post-grant validity check on issued patents that is faster, cheaper, and better than federal court challenges. Those objectives formed Congress's basis for creating the *inter partes* reexamination process.

H. Impact of Reexamination on Remedies

Concurrent reexaminations can have a dramatic impact on remedies available to patent owners in infringement actions. Specifically, any substantive amendments during

48 At Sedona PL08 this possibility of multiple remands (dubbed the "infinite do loop" from computer programming parlance) was raised and seriously discussed. Senior PTO officials and others were concerned about its existence. Since then the authors have conferred with these senior officials and others to assess whether this problem is global or whether it is limited to two possible remands or to certain situations (e.g., the unadopted proposed ground of rejection example discussed). It seems that it is too early in the deployment of the current *inter partes* reexamination process to know whether the problem is global. Moreover, it may be that the BPAI will go more in the direction of acting as a judge who makes a final decision and sets the case for appeal to Federal Circuit, and less as a "super examiner" who remands for another round of prosecution in its review of *inter partes* reexaminations.

49 USPTO *Inter Partes* Reexamination Filing Data (June 30, 2010).

50 *Id.*

51 The question has been raised by several people as to why this multiple remand ("infinite do loop") problem does not exist with *ex parte* reexamination. One answer may be the structural difference of the third party requester not having standing once the *ex parte* request is instituted. A patent owner has no incentive to raise the "new ground of rejection" argument at the BPAI that the CRU should have adopted a proposed ground of rejection not used from the reexamination request.

reexamination to asserted patent claims can invoke the doctrine of intervening rights, which can reduce the damages available to a patent owner. Reexamination can also impact any injunctive relief that a patent owner has obtained. Both of these issues are discussed next.

Intervening Rights

During reexamination, the patent owner is permitted to make amendments to the claims undergoing reexamination. However, a claim amendment may have a significant impact on the ability of the patent owner to obtain damages for infringement of that claim prior to the issuance of the reexamination certificate. An amendment to a claim creates two important time periods for damages considerations – (1) the period between the issue date of the original patent and the issue date of the reexamination certificate and (2) the period after the issue date of the reexamination certificate until the patent is no longer in force.

The impact on damages is dependent on the scope of the amendment made in the reexamination proceeding. An amendment to a claim may extinguish past damages if the amended claim is ***not*** "substantially identical" to an original patent claim. In that case, the patent owner can only recover damages during the second period (after the issue date of the reexamination certificate). However, amended claims that ***are*** substantially identical to an original patent claim can be enforced (and damages recovered) from the issue date of the original claims.

The effect of amended reexamination claims during the second time period (after the issuance of the reexamination certificate) is governed by the doctrine of intervening rights.[52] Intervening rights is a defense to infringing activity occurring ***after*** issuance of the reexamination certificate.[53] Section 252[54] provides for two separate and distinct intervening rights defenses – absolute intervening rights and equitable intervening rights.

Absolute intervening rights are created by the first sentence of the second paragraph of 35 U.S.C. § 252:

> A reissued patent shall not abridge or affect the right of any person or that person's successors in business who, prior to the grant of a reissue, made, purchased, offered to sell or used within the United States, or imported into the United States, anything patented by the reissued patent, to continue the use of, to offer to sell, or to sell, or to sell to others to be used, offered for sale, or sold, the specific thing so made, purchased, offered for sale, used or imported unless the making, using, offering for sale, or selling of such thing infringes a valid claim of the reissued patent which was in the original patent.

Absolute intervening rights extend only to the "specific things" actually made, purchased, offered for sale, or used before the issuance of the reexamination certificate. Other continued activity after a reexamination certificate issues is covered by equitable intervening rights defense. Absolute intervening rights apply only to tangible articles existing prior to the date of the reexamination certificate.[55] The applicability of absolute

52 *See Fortel Corp. v. Phone-Mate, Inc.*, 825 F.2d 1577, 1580 (Fed. Cir. 1987).
53 *Seattle Box Co. v. Industrial Crating and Packing Inc.*, 756 F.2d 1574, 1579 (Fed. Cir. 1985).
54 Any proposed amended or new claim determined to be patentable and incorporated into a patent following either an *ex parte* or *inter partes* reexamination has "the same effect as that specified in section 252 of this title for reissued patents." 35 U.S.C. §§ 307(b), 316(b).
55 *Shockley v. Arcan, Inc.*, 248 F.3d 1349, 1353 (Fed. Cir. 2001).

intervening rights to a *process* practiced prior to the date of the reexamination certificate has not yet been addressed by the Federal Circuit.

Equitable intervening rights are created by the second sentence of the second paragraph of 35 U.S.C. § 252:

> The court before which such matter is in question may provide for the continued manufacture, use, offer for sale, or sale of the thing made, purchased, offered for sale, used, or imported as specified, or for the manufacture, use, offer for sale, or sale in the United States of which substantial preparation was made before the grant of the reissue, and the court may also provide for the continued practice of any process patented by the reissue that is practiced, or for the practice of which substantial preparation was made, before the grant of the reissue, to the extent and under such terms as the court deems equitable for the protection of investments made or business commenced before the grant of the reissue.

The equitable intervening rights provision gives a court the power to allow the continued manufacture, use, offer for sale, or sale of a "thing" in situations where a defendant has made substantial preparation to make, purchase, offer for sale, or use the "thing" or in the case of process "practice of the process."[56]

A court may apply the equitable intervening rights provision "to the extent and under such terms as the court deems equitable for the protection of investments made or business commenced before the grant of the reissue."[57] The statute provides no further guidance on how to make this determination. Generally, the equitable intervening rights defense protects a third party who relied to its detriment on an aspect of the original claims that was changed by reexamination.[58]

Injunctive Relief

Patent reexamination can also have significant impact on equitable remedies. In *Fexiteek Americas, Inc. v. Plasteak, Inc.*, the Southern District of Florida lifted a permanent injunction that was on appeal to the Federal Circuit, where the PTO had issue an Advisory Action in a concurrent *ex parte* reexamination of the patent-in-suit.[59] The Advisory Action, which follows the patent owner's response after a final rejection, stated that Flexiteek had failed to overcome all of the rejections of the patent-in-suit. After the Advisory Action issued, the defendant moved the district court to lift the permanent injunction, invoking language therein stating that "upon any decision by a court of the [PTO] that renders the '881 Patent invalid or unenforceable, [the defendants] may petition this Court to terminate this Permanent Injunction." The plaintiffs argued that the Advisory Action was not a final PTO decision and that they had "additional challenges," including a timely noted appeal to the BPAI.[60] The court sided with the defendants and terminated the permanent injunction: "the Court finds that the PTO's Advisory Action is not only the most recent decision regarding the '881 Patent's validity, it is a decision made after a thorough examination of the matter conducted by a body which holds particular expertise in such issues. The court

56 35 U.S.C. § 252.
57 *Id.*
58 *Slimfold Mf. Co. v. Kinkead Industries, Inc.*, 810 F.2d 1113, 1117 (Fed. Cir. 1987).
59 *Fexiteek Americas, Inc. v. Plasteak, Inc.*, Civ. Case No. 08-60996 (July 20, 2010 Order).
60 *Id.*

finds that the PTO's determination should control and the Court will terminate the Permanent Injunction."[61]The plaintiff also pointed out its ability to amend its patent claims based on a reissue application of the same patent, and that even if the current claims were canceled defendants will infringe the reissued claims. The Court rejected that argument as speculative.[62]

Some critics have argued that this decision is decided wrongly. The final Office action was on appeal to the BPAI and the claims of the patent had not been finally cancelled. The BPAI may reverse the CRU or the patent owner may seek additional appeals of the reexamination to the Federal Circuit. Simply put, these critics contend that the CRU's determination is only the first step in determining patent validity in a reexamination. Indeed, it is not uncommon for the BPAI to reverse final CRU determinations. Statistics show that a sizeable number of appealed *ex parte* reexaminations result in reversal and issuance of reexamination certificates with one or more original claims confirmed. If this decision is widely adopted, and permanent injunctions can be routinely lifted based merely on a final Office action, tremendous uncertainty could result for patent owners. Critic's views aside, the case underscores the significant impact that reexamination—even an ongoing one—can have on the remedies available in concurrent litigation. And, it highlights the need for parties to educate the court on the procedural nuances of reexamination so that well informed orders can be entered.

I. Impact of Reexamination and Court Decisions on Stock Price

The CRU, the courts, and the parties should all be aware of how their decisions involving reexaminations can impact the world outside the tribunal. In certain high profile situations in the past several years, the stock price (and company valuation) of a publicly traded patent owner has dropped precipitously or increased substantially due to a significant decision in a patent suit[63] or in a concurrent reexamination of the patent-in-suit such as the issuance of a non-final Office action rejecting some or all of the claims of the patent in reexamination.[64] A precipitous drop often is in large part due to short selling the stock of the patent owner. Putting aside whether such practices are proper, it seems fair that the PTO should indicate on its web site that rejections in non-final Office actions do not necessarily reflect the final disposition of such claims in reexaminations. This information may stem the negative implications of a non-final Office action and thus may allow the capital markets to more accurately and appropriately react to these reexamination developments.

The reaction of the markets to events in the reexamination proceeding underscores the importance of the *ex parte* communications procedures in place at the PTO. Advance

---

61 *Id.*

62 *Id.*

63 In January 2009, shares of Rambus fell nearly 40 % when the U.S. District Court for the District of Delaware ruled in a patent infringement suit that the "company cannot enforce 12 of its semiconductor patents in a suit against Micron Technology Inc. because Rambus destroyed documents about them." *See* IPLaw360, Document Shredding Voided Rambus Patents (Jan. 9, 2009).

64 Tessera Technologies shares dropped nearly 40% following a non-final Office action in a patent reexamination. Tessera's stock recovered somewhat, but only after its general counsel contacted investors and assured them non-final actions were not unusual and that "[c]laims of a patent can not be invalidated in reexamination until the process if fully complete, including all appeals." (*See* http://www.forbes.com/markets/2008/03/04/tessera-chip-patent-market-equity-cx_md_0304markets37.html). In another reported case, 01 Communique's stock fell more than 61% following an announcement that its patent infringement case against Citrix had been stayed pending reexamination of the patents underlying the claims. (*See* http://network.nationalpost.com/np/blogs/tradingdesk/archive/2008/03/13/hot-stock-01-communique-down-60-on-patent-re-examination.aspx). Finally, Avistar Communications blamed Microsoft initiated reexaminations for its need to cut its U.S. and European workforce by 25%. (*See* http://www.siliconbeat.com/2008/03/26/local-firm-blames-25-job-cut-on-microsoft-action/).

knowledge of the issuance of an Office action, the issuance of a NIRC, or the issuance of the reexamination certificate is potentially valuable information to a trader. The PTO has guidelines in place that limit communications regarding the substance of a reexamination proceeding. It is imperative that all employees of the PTO strictly adhere to these policies regarding the discussion of substantive activities in reexamination proceedings in order to prevent possible improper insider trading or other violations of SEC rules.

Coupled with precipitous stock drops arising from non-final Office action rejections, litigation and reexamination counsel for the requester have issued press releases touting the significance of the development. While the First Amendment encourages full disclosure of public information, critics contend that these press releases are so slanted that they cross the line of what is proper conduct for attorneys when "litigating in the press."

The authors have interviewed corporate executives and securities litigators regarding what legal rights and responsibilities patent owners have with and against short sellers using non-final reexamination and other patent enforcement developments to make quick big profits. In addition, they have interviewed experts in media relations and received excellent feedback on how best to deal with court and PTO decisions in the parallel universe. In these days of instant corporate "news"—created by mass communications and the Internet, bloggers and message boards, mainstream press and self proclaimed anonymous pundits, investors and manipulators—perception is often more powerful than the "truth" and it behooves senior executives, their counsel and advisors to be current on best practices on how best to deal with this ever present challenge.

The authors commend the book *STOP THE PRESSES*[65] by Richard S. Levick, Esq. and Larry Smith as an excellent source of best practices for meeting this challenge. The following eight suggestions were provided by Melissa Arnoff of Levick Communications on how best to control and frame instant corporate news.

(1) Be positive: Don't repeat negative language or focus on negative verdicts. Instead, find a way to position the news in a positive, or at least neutral, light.

(2) Embrace the media: Instead of hiding from reporters, get to them early to tell your story and put decisions in context so they can tell balanced, accurate stories. If you avoid commenting, your opponent will control the entire story. This is the era of transparency; you cannot hide.

(3) Be an educator, not an enemy: Patent law is complicated. Help inform reporters and investors so that they better understand the re-exam process and don't over-react to small decisions that are part of a long process. At first, this practice may seem awkward, but it will pay great dividends in better and fairer media coverage.

(4) Stay in front of the news: Prepare statements and news releases for each possible court decision before the verdict so that you can deliver your position immediately and not be delayed by the approval process and wordsmithing.

(5) Tell your story: Know what story you want to tell beyond the litigation. What company image do you want to project? Use that image in all your interviews so that you have something to talk about beside the legal process. Every public utterance is ultimately a marketing opportunity.

65 RICHARD S. LEVICK & LARRY SMITH, STOP THE PRESSES: THE CRISIS AND LITIGATION PR DESK REFERENCE (2d ed. 2007).

(6) Repeat. Repeat. Repeat: Don't be afraid to tell your story to as many audiences as possible as many times as possible. You don't have to use the same words every time, but keep the message consistent.

(7) Use your friends: Enlist third-parties to help tell your story and validate the strength of your company, or the details of patent law. It adds credibility and gives you more ways to tell your story.

(8) Apply the "Power of Three": Why spend so much time preparing a speech or writing a paper only to use it once? Maximize your effort by finding at least three uses for each product you create. If you deliver a presentation before a group, re-package the talk as an article for a legal or IP publication, edit it for use as a blog post, and pitch it to the news media as the core of a feature story.

Compounding the issue of adverse impacts on stock is the delay by the PTO in posting documents on its publicly searchable database ("PAIR"). The PTO has improved the delay between filing and posting documents on PAIR. According to a PTO official, the goal is to have documents posted within 2 business days (48 hours). But this delay is still unacceptable in the context of the worldwide public markets.

J. Impact of Settlement Agreements on **Inter Partes** *Reexamination*

What is the effect on *inter partes* reexamination proceedings where a settlement agreement is reached in a concurrent litigation and one of the parties to the reexamination concedes or stipulates either to the validity or invalidity of the patent? It is well settled that, during litigation, patent owners may make admissions to which they may be bound during reexamination proceedings. For instance, a patent owner may make a binding admission as to whether a particular reference is prior art. If a third party requester concedes patent validity in a settlement agreement, is the PTO then obligated to decide the reexamination on such an admission? Would a settlement be considered a final decision such that the estoppel provisions now apply in the reexamination? Does a third party requester lose standing to participate in an *inter partes* reexamination automatically upon settlement or should removal from the reexamination proceeding be a condition of settlement? If a third party requester agrees to step out of the reexamination, would the *inter partes* reexamination continue or does the PTO convert the *inter partes* proceeding to a de facto *ex parte* reexamination (without interview rights and the right of substantive communications)?

A settlement agreement, by itself, will not operate to terminate a co-pending *inter partes* reexamination, even when an identity of claims and issues exists between the two proceedings. The estoppel provisions of 35 U.S.C. § 317(b) states that the estoppels arise "[o]nce a final decision has been entered against a party in a civil action … that the party has not sustained its burden of proving the invalidity of any patent claim in suit." A settlement agreement is not a final decision that a party has not sustained its burden of proving invalidity of a patent claim. The PTO takes this position and will not terminate an *inter partes* reexamination based solely on a settlement agreement between parties. However, depending on the facts of the case, a Consent Order filed in the district court may be sufficient to trigger the estoppel provisions.

The PTO also will not automatically remove standing of the third party requester when the co-pending litigation settles. Therefore, any settlement agreement must address the ability of the third party to continue participation in the *inter partes* reexamination request. A third party can waive its ability to participate further in the reexamination. When the third party steps out of the reexamination, the *inter partes* reexamination effectively turns *ex parte* in nature (i.e., only the patent owner remains). Other interested entities are not permitted by the PTO to step into the shoes of the settling requester.[66] But whether a reexamination can upset a prior settlement agreement is another story, as noted in *In re Swanson*.[67]

K. Impact of Reexamination on Willfulness and Inequitable Conduct.

Willfulness

Another Hot Topic in concurrent reexamination and litigation is the use of reexaminations as a defense against willful infringement. In *In re Seagate Technology, LLC*[68] the Federal Circuit overturned the then existing standard for willful infringement. Under the new standard, "to establish willful infringement, a patentee must show by clear and convincing evidence that the infringer acted despite an objectively high likelihood that its actions constituted infringement of a valid patent."[69] If this threshold objective standard is satisfied, the patentee must also demonstrate that this objectively-defined risk was either known or so obvious that it should have been known to the accused infringer.[70]

Under the new standard, some practitioners argue that the granting of a reexamination request by the PTO defeats a claim of willful infringement. To date, district courts have declined to establish a *per se* rule regarding the impact of reexamination on a claim of willfulness. Instead, the district courts have viewed the granting of a reexamination request as one factor, among a totality of the circumstances, to consider in examining whether a party can meet the requirements of *In re Seagate*.[71] At least one court found that "[i]t does appear that a reexamination order may be taken as dispositive with respect to post-filing conduct."[72]

Decisions on the impact of reexamination on the willfulness inquiry have focused on the status of the claims at the time of the willfulness determination. For example, if a reexamination certificate issued without amendments to the claims or claims in suit, a court may be hesitant to assign much weight to the reexamination request in the willfulness inquiry. However, before the reexamination certificate issues, the validity of a patent remains questionable and allegations of deliberate or reckless actions by a defendant may lack sufficient factual or legal grounds.[73]

Inequitable Conduct

Inequitable conduct is often based on an allegation that the patent owner attempted to deceive the PTO by failing to disclose a known piece of material prior art. A

---

66 In litigations involving multiple accused infringers, the accused infringers that are not requesters to an *inter partes* reexamination run the risk that the requester will settle and use this prospect as settlement leverage with the patent owner who benefits from having the reexamination effectively converted to *ex parte.* However, these same accused infringers benefit because they are not subject to the estoppels of the *inter partes* statute.

67 *In re Swanson*, 540 F.3d 1168, at n.5 (Fed. Cir. 2008) (noting that "an attempt to reopen a final federal court judgment of infringement on the basis of a reexamination finding of invalidity might raise constitutional problems").

68 *In re Seagate Technology, LLC*, 497 F.3d 1360 (Fed. Cir. 2006).

69 *Id.* at 1371.

70 *Id.*

71 *See Lucent Technologies v. Gateway, Inc.*, 2007 U.S. Dist. LEXIS 95934 (S.D. Cal. 2007); *See also St. Clair Intellectual Propery Consultants, Inc. v. Matsushita Electronic Industrial Co., Ltd.*, 2009 U.S. Dist. LEXIS 49882 (D. Del. 2009).

72 2007 U.S. Dist. LEXIS 95934 at *19 (S.D. Cal. Oct. 30, 2007).

73 *Ultratech Int'l, Inc. v. Swimways Corp.*, No. 05-134 (M.D. Fla. Mar. 3, 2009).

finding of inequitable conduct can render the patent unenforceable. Whether the alleged failure to disclose the prior art reference amounts to inequitable conduct rests on a sliding scale of materiality and intent. If the reference is deemed to be highly material, then the required intent showing is diminished. If the reference is deemed to be less material, then the required intent showing is heightened. Because it is difficult to prove intent, those attempting to prove inequitable conduct typically try to convince the court that the omitted reference is highly material.

Parties may attempt to use a reexamination proceeding to buttress the materiality prong of the inequitable conduct analysis. To do so, the omitted prior art reference is used as the basis for a SNQ in a reexamination request. If the examiner is persuaded that the omitted prior art reference forms a SNQ and then orders a reexamination, this will be taken as further evidence as to the materiality of the reference.

However, some reexamination practitioners have noted problems with this approach. First, there are no evidentiary standards associated with a reexamination request. There is no requirement that the attorney argument supporting the request be backed up by an expert declaration or other evidence. Second, discovery is not available in PTO proceedings. The patent owner is very limited in its ability to challenge an improvidently granted SNQ. Third, the use of a SNQ finding is a litigation tactic used by the defense bar to "create" an inequitable conduct charge. The PTO is becoming more aware of this issue as more reexaminations requests are being challenged as tools not for reexamination but for bolstering an inequitable conduct charge.

L. Retained Reexamination Experts

Expert witnesses play a critical role in patent reexaminations, district court litigation, and ITC investigations. Depositions of retained experts in district court litigation and ITC investigation are the norm; expert witness depositions are not permitted in patent reexaminations. Because retained experts are expected to be deposed and will ultimately testify at trial, parties must consider how their expert witness will demean himself while testifying, withstand cross-examination, and appeal to judges and juries. These considerations typically do not enter into a decision to retain an expert witness in reexamination proceedings, especially in *inter partes* reexamination where there is no opportunity to interview examiners.

Where there is concurrent litigation, however, some practitioners have expressed concern that a retained reexamination expert will be subject to deposition in the co-pending district court case or ITC investigation. District courts addressing this issue have uniformly declined to allow a party to a patent infringement suit to depose an expert retained solely for use in a reexamination of the same patent.[74] In disallowing deposition of retained reexamination expert witnesses, these courts have based their decision on the parties' decision not to designate that same witness as an expert in the concurrent litigation.

In *Roy-G-Biv*, the district court denied defendant's motion for reconsideration of a magistrate's order barring the deposition of an expert who submitted a declaration in the reexamination of the patent-in-suit.[75] There, the plaintiff successfully argued that the witnesses were not identified as testifying experts under Federal Civil Procedure Rule 26,

---

74 *Roy-G-Biv v. Fanuc*, 2009 WL 2046488 (E.D. Tex. July 9, 2009);*Goss International Americas, Inc. v. Man Roland, Inc.*, 2006 WL 1134930 (D.N.H. Apr. 28, 2006).
75 *Roy-G-Biv v. Fanuc*, 2009 WL 2046488 (E.D. Tex. July 9, 2009).

nor had the witnesses submitted expert reports in the litigation. The defendants countered that the witness should be deposed as a fact witness.[76] The district court found that the defendants' "attempt to re-characterize the witnesses as fact witnesses" was misplaced and declined to "displace the normal protections of Rule 26(b)(4)," disallowing the depositions of the reexamination expert unless the plaintiff identified that expert as a testifying expert in the litigation.[77]

The same result was reached in *Goss International Americas, Inc.*, where a district court denied defendant's motion to compel the production of documents related to plaintiff's expert's declarations submitted to the PTO during prosecution of the patents-in-suit. The court reasoned that the witness "was retained to provide declarations during the prosecution of the patents-in-suit, he is neither a testifying nor consulting expert in this litigation."[78] There, defendant relied on *Hewlett-Packard Co. v. Bausch & Lomb, Inc.*, 116 F.R.D. 533 (N.D. Cal. 1987), where the court ordered plaintiff to produce draft declarations submitted to the PTO. The Goss Court distinguished *Hewlett-Packard*, stating "[t]hat case was decided under Fed. R. Civ. P. 26(b)(4), because the expert in question, who had provided a declaration to the PTO, was also a testifying expert at trial. Here, by contrast, Rule 26(b)(4) does not apply, because [the expert] is not a testifying expert."

Finally, submission of expert testimony in reexamination is less rigorous than its admission in evidence in litigation. Trial courts act as gatekeepers, deciding whether the expert evidence is reliable before allowing into evidence. The CRU however does not have written criteria by which it decides whether declarations submitted from experts are reliable or even whether the declarant should be considered an expert. Given the lack of cross-examination and depositions, it may be easier to submit a disingenuous expert declaration to the CRU. In many cases, it may be difficult or impossible for a CRU examiner to discern that an expert declaration is not credible, because the examiner can only assess the declaration on the cold written record and there is not the benefit of any cross-examination to reveal shortcomings or fallacies in an expert opinion.

III. THE PARALLEL UNIVERSES EXAMINED

The term "parallel universe" has been used to describe the situation where patent validity is considered simultaneously by both a district court or the ITC and the PTO. It should be noted that the proceedings are "parallel" only in the temporal sense. To the extent that the term connotes any other identity of procedure, the term is a misnomer. In actuality, the proceedings are quite different in both scope and procedure. Understanding those differences is critical to any informed decision making on parallel reexamination proceedings. Before tackling reexamination strategy considerations in Section IV, we first explore some basic differences in the two proceedings.

A. Scope of Proceedings

The scope of available validity challenges is far broader in district court litigation than it is in patent reexamination proceedings. In district court litigation, patent validity may be challenged under any statutory provision, including provisions set forth at 35 U.S.C. §§ 101, 102, 103 and 112. Further, patent challengers also may argue that the asserted patent is unenforceable due to either inequitable conduct or laches. Patent

76 *Id.*
77 *Id*
78 *Goss International Americas, Inc. v. Man Roland, Inc.*, 2006 WL 1134930 (D.N.H. Apr. 28, 2006)

reexamination, on the other hand, is far more limited in scope. By statute, reexaminations may only be initiated when the PTO is presented with a "substantial new question of patentability" or ("SNQ.")[79] A SNQ only may be predicated on prior art printed publications and may not be cumulative to information already considered by the PTO in original prosecution or in prior reexaminations. A SNQ may not be predicated on any other statutory provisions, including whether the claims contain statutory subject matter under section 101 or whether there exists an "on-sale-bar" or "public use" under section 102(b). Any party considering a parallel reexamination should be aware of the limited scope available to challenge the patentability during a patent reexamination.

B. Standard of Review

The standard of review for patent validity is different in district court litigation than it is before the CRU in patent reexamination proceedings. In district court, patent claims enjoy a presumption of validity, which may be overcome only by clear and convincing evidence. In contrast, no such evidentiary presumption exists during reexamination before the PTO. The PTO and the CRU use a "preponderance of the evidence" standard for adjudicating patentability.[80] For this reason, challenging a patent's validity[81] should be easier before the PTO than in the district court. In addition, a patent owner faces many practical limitations in its ability to amend claims during reexamination.[82] One of the most important is that liability for past damages is put at serious risk if claims are amended substantively during reexamination.

C. Claim Construction

The standards for claim construction are very different in district court litigation compared to patent reexamination proceedings. During reexamination proceedings, claims are construed with their broadest reasonable interpretation, consistent with the specification.[83] For this reason, a *Markman* order in district court has no preclusive effect on the PTO and is not binding thereon.[84] A broad claim construction draws in more potential prior art.[85]

In U.S. district courts, on the other hand, claims are often construed so they remain valid in view of prior art. The court looks to get the "correct" claim construction after reviewing the parties' respective positions. Such a construction will typically be narrower than that used by the PTO and may thus limit the world of available prior art.

This is not a merely academic distinction—the difference in claim construction can have real world effects in the parallel universe. Consider a situation where a court issues a

---

79 35 U.S.C. §§ 303, 304, 312, 313.

80 MPEP § 706.I. ("The standard to be applied in all cases is the 'preponderance of the evidence' test. In other words, an examiner should reject a claim if, in view of the prior art and evidence of record, it is more likely than not that the claim is unpatentable.").

81 The district courts determine "validity," while the PTO reexamines "patentability." The authors use these terms interchangeably, but they are technically different.

82 No proposed amended or new claim enlarging the scope of claims of a patent is permitted in a reexamination proceeding. 35 U.S.C. § 305. The test for when an amended or new claim enlarges the scope of an original claim is the same as that under the 2-year limitation for reissue applications that add enlarging claims under 35 U.S.C. § 251, last paragraph. MPEP § 2250 (*citing In re Freeman*, 30 F.3d 1459, 1464 (Fed. Cir. 1994)). After expiration of a patent undergoing reexamination, no amendments may be proposed for entry. 37 CFR § 1.121(j). Further, any amendments and all claims added during the proceeding are withdrawn if a patent expires during pendency of a reexamination proceeding. MPEP § 2250.

83 *In re Yamamoto*, 740 F.2d 1596, 1571 (Fed. Cir. 1984).

84 *In re Trans Texas Holdings Corp.*, 498 F.3d 1290, 1297 (Fed. Cir. 2007).

85 When a patent owner loses their ability to amend the claims (e.g., when a patent term expires during the reexamination proceeding), the standard for claim construction moves from the broadest reasonable interpretation standard to a standard "pursuant to the principle set forth by the court in *Phillips v. AWH Corp.,* 415 F.3d 1303, 1316, 75 USPQ2d 1321, 1329 (Fed. Cir. 2005)(words of a claim 'are generally given their ordinary and customary meaning' as understood by a person of ordinary skill in the art in question at the time of the invention)." MPEP §2258.I.G.

claim construction order in a litigation and the patent owner is ultimately successful in defending its intellectual property right against an invalidity challenge based on this claim construction. The defendant (or another third party) may subsequently challenge the patentability of the same patent in the PTO. In this situation, the PTO, using a broader construction, creates a different scope for the claims, and arguably a different intellectual property right. In such cases, the patent owner is not permitted to adopt the claim construction of the court. Due to these different claim construction standards, a patent owner may be forced into the difficult circumstance of having to amend claims to incorporate the court's construction and potentially lose past damages, or continue to argue the issue in the reexamination proceeding and potentially extinguish all intellectual property rights in the patent. To be sure, more than a few patent owners have faced this exact situation.

It is technically true that a patent owner has the ability to amend claims during reexamination, provided the amendments do not enlarge the scope of the claims. In the parallel universe situations, however, this ability is severely circumscribed. First, substantive amendments to asserted claims could literally wipe out a district court or ITC *Markman* ruling. If the litigation or ITC investigation has progressed to trial, substantive amendments to asserted claims could result in an enormous waste of judicial and party resources—sometimes to the tune of tens of millions of dollars. Second, if a patent owner is forced to amend claims to preserve patentability, it risks the loss of any claim to past damages under the intervening rights doctrine, which is fully applicable to reexamination proceedings.[86] In many instances, this not an insignificant prospect. Third, in *inter partes* proceedings, the patent owner is not allowed to interview the examiner. If the patent owner decides to first argue around the prior art, and then is faced with a final Office action, there is little room for the type of negotiation necessary to arrive at claim amendments likely to be successful in overcoming pending final rejections. Thus, in reality, patent owners have an extremely limited ability to amend claims. This is one area of law that deserves careful attention.

D. Decision Makers

The ultimate arbiter of patent validity is different in the district court than it is in patent reexamination proceedings. This may be stating the obvious, but this fact has very real consequences. Patent validity challenges in district court are determined by a judge or jury that more often than not has absolutely no technical background in the relevant art. Before the CRU, on the other hand, patentability is determined by technically trained, experienced patent examiners. Moreover, the CRU assigns a three examiner team to each reexamination.

E. District Court v. Central Reexamination Unit

"Courts do not find patents 'valid,' only that the patent challenger did not carry the 'burden of establishing invalidity in the <u>particular case</u> before the court …..'"[87] A prior holding of validity by a district court is therefore not inconsistent with a subsequent holding of invalidity by the PTO.[88] While the PTO may accord deference to factual findings made by the court, the determination of whether a SNQ exists will be made independently of the court's decision on validity, since the decision is not controlling on the

86 35 U.S.C. §§ 307(b) and 316(b).
87 *In re Swanson*, 540 F.3d 1368, 1377 (Fed. Cir. 2008).
88 In an *inter partes* reexamination brought by a party to the litigation, a final holding of validity in the civil action triggers the estoppel provisions of 37 C.F.R. § 317(b) in the reexamination proceeding. All issues that were raised or could have been raised by the party in the civil action will not be maintained in the *inter partes* reexamination proceeding.

PTO.[89] A non-final holding of claim invalidity or unenforceability also will not be controlling on the question of whether a SNQ is present. Only a final holding of claim invalidity or unenforceability (after all appeals) is controlling on the PTO. In such cases, a SNQ would not be present as to the claims held invalid or unenforceable.[90] In other words, the PTO will not reexamine patent claims that previously were invalidated by a district court. In sum, only a final, non-appealable, ruling on invalidity is binding on the PTO.

F. Cumulative Effect

The cumulative effect of the "non-parallel" aspects of district court and reexamination proceedings is profound. It is the authors' perception that broader claim construction used by the PTO, combined with the lack of any presumption of validity and skilled decision makers, results in far easier prior art validity challenges. Moreover, the CRU often rejects all of the claims in the first Office action and puts the burden on the patent owner to prove the patentability of the claims even if the claims have been subject to extensive prior art attack in prior court actions.

With a high level understanding of the fundamental differences between validity challenges before the district court and the CRU, we now can explore various reexamination strategy considerations when district court litigation is threatened or pending.

IV. REEXAMINATION STRATEGY CONSIDERATIONS WHEN LITIGATION IS THREATENED OR PENDING

A comprehensive strategy should be in place before filing a reexamination request. The requester should have a clear objective and should be fully aware of the consequences of filing a reexamination request. Once filed, the requester cannot "unfile" the reexamination as part of a settlement with the patent owner.[91] After the reexamination is instituted, it will take on a life of its own. Many pitfalls await the unwary requester who has not fully thought out the consequences of filing a reexamination request. The following sections are presented roughly in chronological order.

A. Reexamination Pendency

Reexamination pendency has attracted high-level scrutiny. For example, then Chief Judge Paul R. Michel commented on the ability of the PTO to handle post-grant proceedings in an Address at the *Federal Trade Commission Hearing on the Evolving IP Marketplace*, held on December 5, 2008:

> To me, the proposed alternative for weeding out bad patents is convincing. Can we really get a faster, better, and cheaper review of challenged patents at the PTO than in the courts? Experience with the existing PTO reexamination procedures raises doubts.
>
> And the PTO is already overwhelmed by ex parte examination with average pendencies over three years, in some arts, far longer. Is it realistic

89 *Ethicon, Inc. v. Quigg*, 849 F.2d 1422, 1428-29 (Fed. Cir. 1988); *In re Swanson,* at 12-18 (citing *Ethicon*, 849 F.2d at n. 3 and *Stevenson v. Sears Roebuck & Co.,* 713 F.2d 705, 710 (Fed. Cir. 1983)) (emphasis in original).

90 *See* MPEP § 2686.04.

91 In a recent example, we are told that a third party requester attempted to halt a reexamination proceeding by stopping payment of a check for the reexamination fee. The PTO indicated that such actions could result in OED imposing a suspension of the attorney from PTO practice for taking the action.

> to expect the PTO to be able to conduct a new form of *inter partes* reexamination faster and cheaper than the courts? And more accurately? Unless its new procedures, competencies, and powers can be clearly defined, how will we know what consequences would follow? How will we know this is not a mirage in the desert that looks like an oasis, but has no water?

We provide below some insight into external and internal CRU and BPAI procedures, as well as the latest information regarding reexamination pendency.

1. *Pendency before the CRU*

All reexaminations are required by statute to be handled with "special dispatch."[92] Nonetheless, higher priority is afforded to reexaminations of patents involved in litigation. Even higher priority is afforded when trial proceedings have been stayed pending the outcome of reexamination. The highest priority is assigned to reexaminations that have been pending for at least two years. The rules require patent owners to notify the Office of prior or concurrent proceedings,[93] and the CRU has dedicated paralegals that search litigation databases for the status of pending litigation during the pendency of the reexamination proceeding. The CRU thus assigns priority based on its own statistics and research, and based on patent owner notifications. Therefore, it is critical for the patent owner to keep the PTO informed of the existence and status of related co-pending district court or ITC proceedings.

Upon filing, reexamination requests first undergo review by the CRU staff to ensure compliance with the rules.[94] Failure to comply with the provisions may result in a Notice of Incomplete Request or Failure to Comply, vacating the filing date until a response is filed within 30 days to remedy any defects.[95] For instance, the staff will ensure that each reference cited by requester is used to support at least one proposed rejection. Further, the staff recently began ensuring that the requester properly and affirmatively demonstrate that each SNQ is non-cumulative of the art previously considered during original prosecution or previous reexaminations. Of the requests received in fiscal year 2010 Q1-Q3, 10% of both *ex parte* and *inter partes* requests were terminated during preprocessing by CRU staff for failure to comply with the requirements of 37 C.F.R. §§ 1.510 and 1.915 before even reaching an examiner.[96]

If the reexamination request passes muster by the CRU staff, a notice of the request is made public in the Official Gazette. When a request is deemed to satisfy all the requirements of the *ex parte* or *inter partes* rules,[97] the filing date becomes the reexamination filing date. Just because the notice of request is published in the Official Gazette does not necessarily mean that the reexamination request was satisfactory. Roughly 10% of requests are later vacated by the examining panel for informalities. While this number has fallen from roughly 15% in previous years, the authors have noted a recent uptick in denial rates based on the requester's failure to adequately prove that the proposed SNQ is not cumulative to the art considered during original prosecution.

---

92 35 U.S.C. §§ 305 and 314.
93 37 C.F.R. §§ 1.565 and 1.985.
94 37 C.F.R. §§ 1.510 and 1.915.
95 Instructions Regarding Notice of Failure to Comply (PTOL 2077).
96 USPTO Reexamination Filing Data, Footnote 1 (June 30, 2010).
97 35 U.S.C §§ 510 (*ex parte* reexaminations) and 919 (*inter partes* reexaminations).

Once a satisfactory request has been made, the CRU has a three month deadline to issue a decision on the request based on whether a SNQ has been raised.[98] According to PTO operational statistics, the amount of time from filing the request to an order granting reexamination averaged 1.6 months for *ex parte* and 1.8 months for *inter partes* proceedings. The amount of time from filing the request to an order denying the request averaged 1.7 months for *ex parte* and 1.1 months for *inter partes* proceedings.[99] While an order denying the request can be petitioned under 1.181 for review of the examiner's determination, a substantive denial of this nature is often a strong indicator that the request will not move forward.

For *ex parte* requests, the grant starts a two month window in which the patent owner may respond to the request. In that response, the patent owner may amend claims or argue that the claims under reexamination are patentable. If the patent owner elects to file a patent owner's statement, then the third party requester may reply. This is the only opportunity a third party requester gets to participate in an *ex parte* reexamination proceeding. Most reexamination practitioners advise against filing a patent owner's response for this reason. For *inter partes* requests, CRU procedures suggest that a first Office action on the merits accompany the grant, but that is not required. Overall, the CRU has established an internal goal to issue a final Office action or an Action Closing Prosecution ("ACP") within two years of the filing date of a request for reexamination. As the CRU works through its backlog, this goal is becoming more and more realistic. According to operational statistics, the average amount of time from grant of a request to the first Office action is 5.8 months for *ex parte* and 1.8 for *inter partes* proceedings. This reflects the CRU's practice of seeking to mail the first Office action along with, or shortly behind, the order granting *inter partes* reexamination. Thus, this milestone for *inter partes* proceedings is better approximated as 3.6 months from the date of filing to the first Office action.[100]

Official statistics suggest that "overall pendency" is reflected by the average time from filing the request to issuance of a NIRC (the often cited duration of 27.4 months for *ex parte* and 34.8 months for *inter partes* proceedings), however, a range of scenarios confound this statistic. First, averages lack the granularity to indicate reasons for issuance of a NIRC, a milestone that serves to notify the public of the final disposition of the claims, including circumstances where all claims have been canceled or abandoned by the patent owner. Averaging together reexaminations prosecuted until finality and those defaulting due to lack of participation may also distort the result. Second, any pendency bookended by issuance of a NIRC must exclude all proceedings pending on appeal. Without understanding the amount of time expected from a first Office action to a final Office action or Right of Appeal Notice, the overall reexamination timeline remains ambiguous and situation specific.

Approximately 46% of reexamination requests involve the electrical/software/business method arts. In addition, the mechanical arts make up around 31% and this number is growing. While there may be a perception that reexamination is disfavored in the chemical/biological arts, over 23% of reexamination filings were in these arts.[101] The lower number of requests in the chemical/biological arts likely mirrors current trends in the technology centers and the fewer overall number of issued patents in the chemical/biological fields. Current pendency rates reflect the large number of reexaminations in the electrical arts.[102]

---

98 The substantial new question of patentability (SNQ) is discussed in more detail in subsequent sections of this paper.
99 USPTO Reexamination Operations Statistics Fiscal Year 2010, Quarter Ending June 30, 2010.
100 USPTO Reexamination Operations Statistics Fiscal Year 2010, Quarter Ending June 30, 2010.
101 USPTO Reexamination Filing Data (June 30, 2010).
102 *See* Section VII.

2. *Pendency before the BPAI*

Once prosecution closes, an appeal to the BPAI is available to the patent owner in *ex parte* reexamination and to both the patent owner and the third party requester in *inter partes* reexamination. After hearing an appeal, the BPAI has a stated goal of then rendering a reexamination decision in six months. In an April 2008 briefing on *inter partes* reexaminations, the Institute for Progress estimated the average pendency for an un-appealed *inter partes* reexamination as more than 3.5 years, and the expected pendency for appealed *inter partes* reexamination as at least 6.5 years.[103]

An independent survey of BPAI decisions rendered between January 1, 2007, and January 1, 2010, and the file histories associated with their reexamination control numbers revealed that (not inconsistent with official performance metrics for fiscal year 2009 provided by the BPAI)[104] of the 173 *ex parte* and 15 *inter partes* proceedings in this sample, the time from docketing at the BPAI to a final decision averaged 4.5 to 6.5 months over the past three years. Nevertheless, the median time from Notice of Appeal (or Right of Appeal) to docketing at the BPAI (representing the duration of briefing) averaged 18-20 months.[105]

At present, the apparent BPAI bottleneck threatens to dwarf any pendency driven by the CRU or even briefing before the BPAI. At least six or seven months can pass between the filing of a notice of appeal, the patent owner's brief, the third party requester response (if *inter partes*), the examiner's answer and subsequent replies. In the *inter partes* reexaminations this period is often longer due to the time between Respondent's Brief and the Examiner's Answer.

To better approximate the briefing period and average pendency of reexaminations after a Notice of Appeal (or Right of Appeal), the file histories of all requests filed between January 1, 2006, and January 1, 2009, as of June 11, 2010,[106] were analyzed. Of the 1,738 *ex parte* proceedings filed during this period, 375 had a Notice of Appeal with opening appeal briefs filed in 246 (66%) of proceedings. Of the 246 appeal briefs filed, the pendency from Notice of Appeal to acceptance of the brief averaged 96 days, with a median of 62 days, and a range of 21-370 days. Of those where acceptable briefs were filed, 149 had received an Examiner Answer in an average of 112 days, with a median of 99 days, and a range of 2-360 days. Of these, only 99 had been docketed at the BPAI. The amount of time from the Reply Brief (if filed) to a notice of docketing averaged 96 days, with a median of 68, and a range of 6-598. Notably, 122 of those reexaminations on appeal received a NIRC prior to any decision by the BPAI.

Similarly, of the 415 *inter partes* proceedings filed during this time period, 131 had a Notice of Appeal with opening appeal briefs filed in 95 (73%) of proceedings. Of the 95 briefs filed, the pendency from Notice of Appeal to acceptance of the last proper brief averaged 162 days, with a median of 133 days, and a range of 60-351 days. Of those where acceptable briefs were filed, 30 had received an Examiner Answer in an average of 108 days, with a median of 77 days, and a range of 24-443 days. Of these, only 17 had been docketed at the BPAI. The time from the last Reply Brief to a notice of docketing averaged 159 days,

103 *Reexamining Inter Partes Reexamination,* Institute for Progress (April 2008).
104 *See* Board of Patent Appeals and Interferences Performance Measures for Fiscal Year 2009 http://www.uspto.gov/ip/boards/bpai/stats/perform/FY_2009_Performance_Measures.jsp.
105 Sterne *et al.*, *Appeals from the Central Reexamination Unit*, 5th Annual Advanced Patent Law Institute USPTO-PV10, United States Patent and Trademark Office, Alexandria, VA (Jan. 21, 2010).
106 Public file histories were obtained from the Patent Application Information Retrieval (PAIR) system on June 11, 2010, and analyzed at their current stage of reexamination.

with a median of 96 days, and a range of 46-521. That this sample is limited reinforces the notion that *inter partes* appeals are less abundant than their *ex parte* counterparts, as described in Section II.G. Only 10 of those on appeal received a NIRC prior to a decision by the BPAI.

A rapidly growing backlog at the Board of Patent Appeals and Interferences ("BPAI" or "Board") is perhaps the most closely watched pendency issue. The backlog has continued in 2010.[107] However, because the BPAI also falls under the statutory mandate of special dispatch, handling appeals from the CRU in the same way as appeals from the initial examining corps would result in impermissible delays. Accordingly, the PTO has taken some recent steps to decrease pendency for appeals coming from the CRU.

For example, many appeals were bounced for non-substantive informalities such as failure to adhere to the PTO's briefing rules, which are not always consistently applied. Or even for failure of the examiner to initial every reference in an IDS. For example, of the 1,738 *ex parte* proceedings filed during the period surveyed, 375 had a Notice of Appeal with opening appeal briefs filed in 246 (66%) of those on appeal. Of the 246 appeal briefs filed, 75 (30%) received a Notice of a Defective Appeal Brief and only 99 have been docketed to date at the BPAI. Similarly, of the 415 *inter partes* proceedings filed during this time period, 131 had a Notice of Appeal with opening appeal briefs filed in 95 (73%) proceedings. Of the 95 briefs filed, 39 (41%) received a Notice of a Defective Appeal Brief and only 14 have been docketed at the BPAI to date. Critics argue that this appeal process in practice does not satisfy special dispatch.

To achieve consistency and streamline procedures for appeals in *ex parte* reexamination, the PTO announced in May 2010 that the BPAI's Chief Judge (or his representative) will now have the sole responsibility determining whether appeal briefs comply with the applicable regulations.[108] This BPAI review will be completed prior to forwarding appeal briefs to the examiner for consideration. Previously, this task fell to the examiners. The PTO expects to "achieve a reduction in *ex parte* reexamination proceeding appeal pendency as measured from the filing of a notice of appeal to docketing of the appeal by eliminating duplicate reviews by the examiner and the BPAI."[109] The PTO also expects a further reduction in pendency because the streamlined procedure will "increase consistency in the determination, and thereby reduce the number of notices of noncompliant appeal brief and non-substantive returns from the BPAI that require appellants to file corrected appeal briefs in *ex parte* reexamination proceeding appeals."[110] It is unclear why this review procedure was not implemented for *inter partes* reexamination.[111] But given the state of pendency last year at this time, any action to streamline appeals from the CRU is generally welcomed by reexamination practitioners.

3. Pendency conclusion

As noted above in the Hot Topics section, reexamination pendency must be considered with an informed eye. Mere review of published PTO statistics does not provide a complete picture of reexamination pendency. Moreover, the past is not necessarily indicative of the future. Stay decisions should be flexible enough to respond to events at the PTO. And strategy decisions should take into consideration best- and worst-case scenarios, and be nimble enough adapt when the unexpected occurs.

107 *See* http://www.uspto.gov/web/offices/dcom/bpai/docs/receipts/fy2009.htm.
108 http://edocket.access.gpo.gov/2010/pdf/2010-12534.pdf.
109 *Id.*
110 *Id.*
111 Speculation is that such changes are in the works.

B. Settlement

Reexamination could help force an early settlement. For example, some practitioners suggest presenting a reexamination request to the opposing party patent owner prior to submitting the request to the PTO (i.e., a "pocket reexamination request"). The idea is to encourage early settlement of pending litigation on favorable terms. The patent owner may need to be educated on the risks that reexamination poses to its patent-in-suit. Further, the patent owner will be put on notice of invalidity risks its patent-in-suit faces at trial. Even if a pocket reexamination does not immediately drive settlement, a grant of reexamination, or an Office action that is adverse to the patent claims, especially a final rejection, may improve the accused infringer's settlement negotiating position. Potential requesters should keep in mind, however, that, once launched, the reexamination cannot be recalled.

C. Litigation Stays

Because the issue of patent validity is running concurrently in two separate proceedings, judicial economy would seem to counsel a stay of one or the other proceeding in all instances. In reality, this is far from true because each venue is bound by different rules and standards for assessing patent validity. Further, each venue is bound by very different rules and standards for determining whether a stay is appropriate, and each stay decision is highly fact specific. Some general patterns may be discerned, however, and these are described more fully below.

District court judges have inherent and almost unfettered control over their own dockets. A decision to stay a case is reviewed for abuse of discretion—a very difficult standard to overcome on appeal. Further, district court judges have great flexibility in the types of stays they issue. For instance, in one Eastern District of Texas case, Judge Everingham granted a motion to stay the litigation based on the accused infringer's *ex parte* reexamination request. In the order granting the stay, the court crafted a stipulation that the accused infringer must agree not to challenge the validity at trial of the patents-in-suit based on prior art patents or printed publications that were considered in the reexamination proceedings. Ordinarily, these estoppels only apply to *inter partes* reexaminations, and only after the proceeding has concluded. Further, the accused infringer was barred from directly or indirectly instituting any further reexamination proceedings, despite being statutorily allowed to do so.

Despite this nearly unfettered discretion, however, courts will generally consider at least the following high level factors in making stay determinations: (1) whether a stay will unduly prejudice or present a clear tactical disadvantage to the nonmoving party; (2) whether a stay will simplify the issues at trial; and (3) whether discovery is complete and whether (or when) a trial date has been set. These broad factors are discussed more fully below.

First, a court looks at the extent to which the non-moving party would be prejudiced in delaying the litigation. Inherent in this factor is consideration of reexamination pendency. Further, as noted above in the Hot Topics, the extent to which the parties are competitors should be considered carefully. In certain instances, the district court will mitigate any potential prejudice to the patentee by requiring a stipulation that the accused infringer will not challenge the patent on grounds considered during reexamination. By doing so, the court reasons, the patentee "is afforded both the advantage of *ex parte* proceeding and an estoppel effect."

Second, courts take into account the possibility of simplifying issues with a stay. Under this factor, the status of the reexamination is often considered. The further along the reexamination, the more likely a stay will be granted. Typically, stays are rarely granted on the basis of the reexamination grant; at least a first Office action rejection is required. Of course, potential invalidation of the only patent-in-suit would simplify many issues, but cases are often more complicated. For example, an accused infringer may have a strong case for patent invalidity based on statutory subject matter or an on-sale bar. Reexaminations may not be instituted on this basis and the court may still, therefore, have to determine patent invalidity on these grounds if the patent survives reexamination on the prior art. As the PTO develops more information about the reexamination process, in particular the statistics of *inter partes* reexamination, courts will be better able to make an informed decision as to whether a stay will simplify a subsequent trial.

Finally, the court asks if discovery is complete and whether trial dates have been set. Judicial economy naturally favors requests made early in the litigation. Therefore, in view of (1) and (2) above, any patent challenger hoping to stay the more costly district court litigation should strive to get its reexamination filed as soon as possible and to request a stay of a co-pending litigation as soon as feasible. Denials due to premature requests are usually without prejudice and stay requests can be renewed based on developments during reexamination.

In a particularly nuanced stay decision, Judge Selna of the Central District of California articulated a number of factors supporting the grant of a stay in *Allergan Inc. v. Cayman Chem. Co.*:[112]

1. Prior art presented to the Court will have been considered by the PTO, with its particular expertise.

2. Discovery problems relating to prior art can be alleviated by the PTO examination.

3. It the PTO invalidates the patent-in-suit, the case will likely be dismissed.

4. The outcome of reexamination may encourage settlement.

5. The record of reexamination will likely be entered at trial, thereby reducing the complexity and length of the litigation.

6. Issues, defenses and evidence will be more easily limited in pre-trial conferences after a reexamination.

7. The cost will likely be reduced for both the parties and the court.

Judge Selna also noted the following factors that would support denial of a stay:[113]

1. Delay and changing market conditions over time may dramatically lower the value of injunctive relief.

---

112 *Id.*
113 *Id.*

2. Substantial expense and time invested may be wasted when litigating issues also under reexamination.

3. A delay may grant a tactical advantage to the moving party.

4. The reexamination outcome might not affect the civil litigation.

In *Allergan*, Judge Selna denied the stay primarily because the parties were competitors, and "the PTO might not conclude its reexamination for several years, and that this delay would prejudice Allergan's rights to exclusive use of its patented technology and would cause it irreparable harm."[114]

Of course, reexamination pendency remains a big factor. Consider this extreme example of a litigation stayed pending an *inter partes* reexamination: In July 2002, Harry Shannon filed an *inter partes* reexamination request (Control No. 95/000,005) challenging the validity of a patent asserted in a patent infringement litigation in the Middle District of Florida (*Enpat, Inc. v. Shannon, et al.*, 6:02-cv-00769). In September 2002, the district court stayed the litigation pending a final decision on the validity of the claims by the PTO. In the reexamination proceeding, the right of appeal notice was issued by the CRU in August 2005, following an action closing prosecution. The appeal has yet to be decided by the BPAI. In May 2004, the district court directed administrative closure of the case pending final decision by the PTO. In that order, the parties were required to provide periodic status reports on the reexamination proceeding to the district court. In February 2005, the plaintiff Enpat filed an unopposed motion to dismiss the case without prejudice due to the pending reexamination request.

Judges deciding a motion to stay, or presiding over a concurrent litigation stayed pending a reexamination, may consider contacting the CRU.[115] The authors have been informed that calls from judicial clerks to the PTO have occurred. Some commentators argue that this ability to contact the CRU when deciding a motion to stay is a valuable tool to a judge deciding whether to grant a stay motion. What restrictions, if any, should be placed on communications between federal judges and CRU officials? Should these communications be limited strictly to procedural details or as fellow government officials should judges be provided less restrictive communication? These questions remain open.

If a party is successful in obtaining a stay in the district court litigation pending resolution of a reexamination, that fact immediately should be brought to the attention of the CRU. The CRU has set forth procedures to increase the pace at which reexaminations involved in concurrent litigations are handled. For example, in situations where a stay is granted in a concurrent litigation, the PTO will take up a reexamination request within six weeks of filing and "all aspects of the proceeding will be expedited to the extent possible."[116] It is therefore critical for parties to keep the CRU informed of the status of the concurrent litigation.

Stays are also technically available in reexaminations. Unlike the district courts, however, the PTO does not have an unfettered ability to control its docket. The rules

---

114 *Id.*
115 The Office of Patent Legal Administration has a hotline for questions dealing with "Reexamination and Reissue legal and policy guidance." The number is (540) 772-7703. The authors have found the staff to very helpful and responsive to all manner of requests. However, OPLA is prohibited from discussing any substantive aspect of a specific *inter partes* proceeding.
116 MPEP § 2686.04.

provide the ability for the patent owner to request a stay.[117] Generally, the PTO has been unwilling to grant such stay requests due to the statutory mandate to handle reexaminations with special dispatch. However, in an *inter partes* reexamination, the PTO may be amenable to a stay where the reexamination proceeding is at its beginning stages, the litigation is near a final resolution, and estoppel would render all issues in the reexamination moot when the litigation becomes final. To avoid a possible stay of the reexamination proceeding, the requester should file a reexamination request as early as practical in a concurrent litigation.

Finally, the Federal Circuit will not likely issue a stay of any case before it. If an appeal arrives from the BPAI, the Federal Circuit will rule and any decision adversely affecting the validity of any patent claim would trump any district court decision to the contrary. If an appeal from the district court arrives first, the Federal Circuit will likewise rule on the district court case. The PTO would be bound by any ruling invalidating a claim, but the reverse is not true. An unsuccessful validity challenge in the district court is not binding on the PTO as it reviews patent validity under different standards. To the extent that the cases arrive simultaneously at the Federal Circuit, the court may review the BPAI decision first. For instance, the Federal Circuit was presented an appeal from a district court decision and from the BPAI on the same patent.[118] In the district court decision, a jury awarded over $85 million for Hitachi's infringement of Translogic's patent. In a parallel decision, the BPAI found the patent to be invalid as obvious. Both appealed to the Federal Circuit. The Federal Circuit first heard the appeal from the BPAI and affirmed the patent's invalidity in a precedential decision. The Court then vacated the district court's decision and remanded for dismissal.

In the end, the best source for how a particular district court judge will deal with a motion to stay is local counsel. Local counsel should have their finger on the pulse of the court and its judges at any moment in time. For the PTO, stays are highly unlikely given the statutorily imposed mandate to deal with reexaminations with special dispatch. Finally, the Federal Circuit likely will deal with appealed cases as they are presented to it, without issuing any stay. This is especially true where the BPAI decision arrives prior to, or simultaneously with, a district court decision.

D. Protective Orders

A protective order dictates how confidential documents produced during a litigation are handled by the parties. When crafting a protective order, it is imperative for both the plaintiff and defendant to consider the possibility of a concurrent reexamination. For example, the interplay between the duty of disclosure in a reexamination proceeding and a protective order in a concurrent litigation is a critical and difficult issue facing both the patent owner and the third party requester. Another important issue is the extent to which individuals involved in the litigation (and privy to confidential materials) should be involved in prosecution of a reexamination.

How should parties craft a protective order in a concurrent district court litigation or ITC investigation to prepare for a possible reexamination proceeding at the PTO? Is it possible for a patent owner to satisfy its duty of disclosure while adhering to the guidelines of a protective order? What limitations does a protective order place on the resources available to a patent owner to prosecute the reexamination proceeding? What mechanisms are available to provide information of nonobviousness covered by the protective order in the reexamination proceeding?

---

117 See 37 C.F.R. §§1.565(b) and 1.987.
118 *In re Translogic Tech., Inc.*, 504 F.3d 1249 (Fed. Cir. 2008).

1. Scope of the Duty of Disclosure in a Reexamination Proceeding

In a reexamination proceeding, each individual associated with the patent owner has a duty of candor and good faith in dealing with the PTO. The duty of candor includes a duty to disclose to the PTO all information known to that individual to be material to patentability in a reexamination proceeding.[119] Individuals who have a duty to disclose to the PTO all information known to be material to patentability in a reexamination proceeding include "the patent owner, each attorney or agent who represents the patent owner, and every other individual who is substantively involved in a reexamination proceeding".[120]

Some practitioners argue that the scope of the duty of disclosure in 37 C.F.R. § 1.555 is ambiguous. Are all employees of the patent owner as well as every attorney or agent that represents the patent owner subject to the duty of disclosure regardless of their involvement in prosecution activities? For example, are litigation counsel covered by the duty of disclosure? Are retained experts?

Several district courts have limited the scope of the duty of disclosure to only those attorneys or agents substantively involved in preparation or prosecution. In *Intelli-Check v. Tricom Card Techs., Inc.*[121] the District Court for the District of New Jersey subscribed a duty of candor on each named inventor, prosecuting attorney, and other individuals who are substantively involved in the preparation or prosecution of the patent application. Similarly, Chief Judge Spencer in the Eastern District of Virginia found that a "party is only bound by the duty of candor, and therefore can only be penalized for failure to disclose material information, *if they are substantively involved in the preparation and prosecution of the patent application.*"[122]

Therefore, a key question facing a patent owner is who can be called upon to prosecute or consult on strategy for the reexamination proceeding or the prosecution of other pending applications. To what extent may the patent owner's trial team participate in prosecution of its pending patent applications, reexaminations, reissues, or interferences using information derived from the litigation? To what extent may the accused infringer's trial team participate in prosecution of their own patent applications in the same subject matter as the patents in suit, in reexaminations of the patent owner's patents, or in interferences involving the patent owner? Many litigation attorneys take a conservative approach to this issue and strictly avoid any involvement or discussions related to the prosecution of pending applications or reexaminations. Other litigation attorneys take a less conservative approach and participate in consultation on specific reexamination issues.

Additionally, a patent owner also must consider to what extent their reexamination expert may be considered to be "substantively involved in the preparation or prosecution of the patent application." This consideration may influence whether a patent owner utilizes that same expert for the litigation as in the reexamination proceeding.

---

119 37 C.F.R. §§ 1.555(a) and 1.933(a).
120 37 C.F.R. §§ 1.555(a) and 1.933(b).
121 2005 U.S. Dist. LEXIS 38794 at *14 (D.N.J. Nov. 10, 2005.)
122 *TeleCommunication Sys., Inc. v. Mobile 365, Inc.*, No. 06-485 at 8 (E.D. Va. Mar. 31, 2009) (emphasis added).

2. Considerations for Crafting a Protective Order

When crafting a protective order for district court litigation or ITC investigations, the parties must consider the possibility of a reexamination being filed. For example, the parties may establish procedures to allow the filing of materials obtained during discovery that are subject to a protective order in the PTO reexamination proceeding. Absent sufficient procedures, a patent owner may be barred by the protective order from presenting materials to support patentability (e.g., evidence of secondary considerations of nonobviousness) to the PTO. Similarly, the third party requester may be barred from submitting evidence supporting its obviousness position to the PTO.

Another important consideration is how to craft the prosecution bar provision of the protective order. Prosecution bar provisions set forth the parameters for the involvement of individuals associated with a litigation in a PTO proceeding involving the patent-in-suit or a family member of the patent-in-suit. Patent owners typically will want the attorneys handling their reexamination to be able to discuss strategy with litigation counsel so consistent positions can be advanced in the multiple forums. Furthermore, it is more cost effective for the patent owner to utilize the same counsel in both proceedings. When negotiating the prosecution bar provisions, the patent owner typically argues for a very narrow prosecution bar or even no prosecution bar at all. Defendants typically want broader, more restrictive, prosecution bars.

As a general matter, it is often argued that no party having access to another party's highly confidential technical information under a protective order should be allowed to amend or supervise the amendment of pending claims in applications or claims under reexamination in the same technical space, nor should they be allowed to draft new claims.[123] In-depth knowledge of a competitor's highly confidential technical information, combined with the ability to amend or draft claims, may convey an unfair advantage to the claim drafter. This applies equally to patent infringement plaintiffs and defendants, and applies equally whether the highly confidential information is received from an adversary or a party with temporarily aligned interests such as a co-defendant.

Unresolved questions remain regarding the interplay between non-prosecution clauses in a protective order and a patent owner's duty of disclosure under 37 C.F.R. §§ 1.555 and 1.933. We expect significant developments in this area as the Office of Patent and Legal Administration ("OPLA") and the courts wrestle with this issue.

3. Handling Conflicting Duties

In concurrent litigation, a patent owner may be faced with the circumstance in which it has the duty to disclose materials to the PTO under the duty of disclosure but has a conflicting duty to maintain the confidentiality of these materials under the protective order. How does a patent owner handle these conflicting duties?

Several practitioners have argued that the duty of disclosure to the PTO takes precedence over a protective order of a district court. Chief Judge Spencer of the Eastern District of Virginia addressed the issue of whether the PTO's duty of disclosure overrides the protective order of a district court such that the party does not have to obtain authorization

123 The protective order should explicitly and separately address applications and reexaminations. A reexamination is not an "application" and if the order simply addresses "applications," then reexaminations would not expressly be covered.

to file materials subject to the protective order in the PTO.[124] Judge Spencer found that there was "no evidence – legal or logical – to support" the patent owner's (plaintiff's) contention that the protective order is overridden in its entirety by the PTO's disclosure requirements. A patent owner must therefore carefully consider the provisions of the district court's protective order before submitting materials covered by the protective order to the PTO.

In sum, both parties must consider the possibility of concurrent reexamination and must pay close attention to the protective order. Patent owners must understand their disclosure obligations under the reexamination rules. Patent owners should also consider carefully the duties and restrictions imposed upon them by receipt of such confidential or classified information. Accused infringers likewise need to be concerned about the impact of disclosing highly confidential technical information to other parties capable of drafting and amending patent claims.

The PTO has struggled with the issue of submission of third party confidential materials covered by a protective order. In recent decisions, the PTO has stated that information covered by a court's protective order is not to be submitted without evidence that permission has been granted for the public disclosure of the information. If such evidence is not provided, the PTO expunges the entire filing containing the information.

4. Submission of Evidence Supporting Patentability

Another important consideration in drafting a protective order is the ability to submit to the PTO evidence favorable to patentability. Often in patent litigation, the best evidence of patentability originates from the accused infringer or another third party. This is especially true where the defendant is advancing obviousness-based invalidity challenges. If evidence of secondary considerations of non-obviousness exists, the patent owner will want to present this evidence to the PTO in a reexamination. However, if this evidence is marked confidential pursuant to a protective order, the patent owner will not be able to submit the evidence without permission from the owner of the information or the Court.

The issue of submission of third party confidential materials to the PTO in support of patentability is particularly difficult. Currently, no PTO procedures exist to file evidence in support of patentability under seal. Once filed, the materials are placed in the public record. In contrast, documents material to unpatentability may be filed under seal. Some practitioners argue that this provides an unfair advantage to the third party because a third party can allow the filing of documents counter to patentability in the PTO but block the filing of documents in support of patentability.

When negotiating the protective order, the patent owner must consider the potential need to file confidential evidence in support of patentability to the PTO and put provisions in place. One avenue to consider is a proactive procedure to challenge the designation of materials as confidential. Often protective orders require challenges to the confidentiality designation of a document to be made promptly, often long before the prospect of reexamination has particularized in the minds of the parties. Therefore, the parties should address exceptions to allow for later challenges to the designation of documents in the event of a reexamination proceeding. Finally, there is a need for the PTO to be mindful of this issue and implement procedures to address it.

---

124 Memorandum Order, March 31, 2009, *TeleCommunication Systems, Inc. v. Mobile 365, Inc.*, C.A. No. 3:06CV485, p. 9 (E.D. Va. Mar. 31, 2009).

E. Impact on Trial[125]

Can the existence of or developments in the reexamination proceeding, such as a final rejection of the claim(s) in suit, be brought to the attention of the jury, or is this inadmissible because it is too prejudicial? One patent litigator shared with the authors that her mock jury research indicated that the mere knowledge by the jury of the existence of the reexamination reduced the likelihood that the jury would find invalidity because it assumed that the PTO, the administrative agency expert in deciding validity, now had assumed responsibility for this issue. In effect, the jury would "punt" on the invalidity issue. Assuming this jury research is reproducible, it further argues that the judge should be concerned about the prejudicial impact on the jury of any information about the existence of or developments in the reexamination. Indeed, the possibility that unexpected developments in the reexamination could be admissible at trial has been the basis for some trial counsel choosing not to seek reexamination even when there is strong prior art.[126]

However, at least one district court decision deemed non-final reexamination proceedings to be too prejudicial to present to a jury.[127] In that case, the court ruled that "without any [final] conclusions of the PTO to rely upon, evidence that the PTO is currently reexamining the patent may work to unduly alleviate Defendants' 'clear and convincing' burden for both invalidity and willfulness in front of the jury." Similarly, the Federal Circuit recently recognized that "a requester's burden to show that a reexamination order should issue from the PTO is unrelated to a defendant's burden to prove invalidity by clear and convincing evidence at trial."[128]If the trend of these recent decisions continue, it may be difficult to get any evidence at all related to reexamination proceedings before a jury.

Finally, some trial counsel take the view that, if the prior art is not successful before the CRU, it will be even less successful before the court. This is based on the use of *KSR* at the PTO, the broader claim construction rules at the PTO, and the lower burden of proof of invalidity at the CRU. To put it in sound bite terms – "If you can't win it at the CRU, then you have even less chance to win it in court." Such counsel take this view particularly in jurisdictions having a reputation for upholding the validity of patents.[129]

F. Damages

Official PTO statistics indicate that 65% of *ex parte* and 43% of *inter partes* reexaminations result in some change to the claims. In 12% of *ex parte* and 49% of *inter*

---

125 This is a Hot Topic based on discussions with many top trial counsel. While it is still anecdotal, it appears that the more "trial experienced" the counsel is, the less she is concerned that the existence of or developments in a concurrent reexamination will be allowed by the judge to be brought to the attention of the jury. We solicit feedback on this topic.

126 This concern seems to be intertwined with several factors. The first is that many judges have little practical knowledge about reexaminations and what really happens at the CRU or the BPAI. They could be "easily swayed" by initial developments from the CRU, whether it is rejection of all of the claims or allowance of some of the claims in suit in the first Office action. The second is that some judges would like to defer to the PTO the validity issue. The third is the possibility of a litigation stay, which if granted initially and then lifted later could present a dilemma in explaining to the jury why there has been a time delay in the suit.

127 *Microsoft Corporation et al v. Commonwealth Scientific & Industrial Research Organisation*, No. 06-00549 (E.D. Tex. Apr. 9, 2009).

128 *Procter & Gamble Co. v. Kraft Foods Global, Inc.*, 549 F.3d 842, 848 (Fed. Cir. 2008).

129 The court of particular interest is the U.S. Dist. Ct. (E.D. Tex.) The perception is that most of the judges in this "pro patent" court resent the intrusion of the reexamination process into their judicial proceeding. Thus, the concern is that if there is a "favorable" development in the CRU for the patent owner, the judge would be more inclined to let the jury know of this development than if there has been an "unfavorable" development. Trial counsel who have faced this issue with these judges observe that there is a low likelihood that the judge will allow any developments in the reexamination to be brought to the attention of the jury because they need to be balanced on such admissibility determinations. As one trial counsel put it in opining that all developments in reexaminations will be excluded from the jury, the balanced approach of exclusion will be based on the adage of "what is good for the goose [patent owner], is good for the gander [third party requester]".

partes reexaminations all claims are canceled outright.[130] Where claims are substantively amended, the accused infringers may not be liable for past damages under intervening rights law. This can be crucial where the patent term is short or the accused infringers have clear and inexpensive design-around options.

More specifically, substantive amendments made during reexamination may defeat damages for past infringement under the statutory doctrine of intervening rights.[131] A patent owner cannot seek damages for claims that are not substantially identical to the original claims.[132] There is no *per se* rule for determining whether a claim is not "substantially identical."[133] The analysis includes examining "the claims of the original and the reexamined patents in light of the particular facts, including prior art, the prosecution history, other claims, and any other pertinent information."[134] The determination is a legal one, and a claim is changed if its scope is changed.[135]

If a claim is not substantially identical, then a patentee may not seek damages for product sales prior to issuance of the reexamination certificate. If damages have already been awarded, a defendant may seek to have damages vacated since the claims were void *ab initio*. Because of the potential impact on damages, litigation counsel should consider filing reexamination requests on all patents-in-suit, if possible, and on all the asserted claims.

Accused infringers should consider the following CRU statistics as of June 30, 2010:

- 75% of *ex parte* reexaminations initiated by a third party, reaching the issuance of Reexamination Certificate, resulted in some or all of the claims being canceled or amended. Only 25% survived with all claims being confirmed.[136]

- 49% of *inter partes* reexaminations completed resulted in all claims being canceled. Combined with those amended, 92% resulted in some change to the claims. Only 8% survived with all claims being confirmed.[137]

A full set of current CRU statistics is included at the end of this paper.[138]

G. Potential risks for accused infringers

The clearest risk for an accused infringer is that at least one asserted patent claim survives the reexamination process unamended and without any adverse prosecution history estoppels. The reexamination may allow the patent owner to have the CRU consider all of the prior art in the litigation and to present arguments and declarations that support patentability of the claims over this prior art. This could include, for instance, possible secondary (objective) considerations of non-obviousness that were not present when the claims were originally prosecuted.

130 USPTO Reexamination Filing Data (June 30, 2010). Some critics assert that there is insufficient granularity in these PTO statistics to reflect situations where some but not all asserted claims are either found to be invalid or are substantially amended.
131 35 U.S.C. §§ 252, 307(b), 316(b).
132 *Bloom Eng'g Co. v. N. Am. Mfg. Co.*, 129 F.3d 1247, 1250 (Fed. Cir. 1997).
133 *Laitram Corp. v. NEC Corp.*, 952 F.2d 1357, 1358 (Fed. Cir. 1991).
134 *Id.* at 1362-63.
135 *Tennant Co. v. Hako Minuteman, Inc.*, 878 F.2d 1413, 1417 (Fed. Cir. 1989).
136 As of June 30, 2010, 7,586 *ex parte* reexamination certificates have been issued by the PTO. This represents approximately 80% of all *ex parte* reexaminations granted. Recall that *ex parte* reexaminations may also be initiated by request of the patent owner or the PTO Director.
137 As of June 30, 2010, only 167 *inter partes* reexamination certificates have been issued. This represents approximately 21% of all reexamination requests granted by the PTO.
138 The PTO updates these statistics quarterly. *See* http://www.uspto.gov/web/patents/cru.html.

An *ex parte* reexamination may also give the patent owner a significant advantage in dealing with the prior art because, once begun, the third party requester is excluded from the process, while the patent owner can interview the examiner. Although the court can find a patent invalid even if it survives reexamination, most judges likely will defer to the presumed administrative expertise of the PTO, CRU, and BPAI.

Further, putting an asserted patent into reexamination could allow the patent owner to correct other defects in the patent, such as potentially ambiguous claim language, antecedent basis problems, or other perceived issues with the claims. This is especially true with newly issued patents where the potential for damages lies in the future, rather than with past damages. In the same vein, patent owners can also add claims during reexamination, provided that the added claims are not broader in scope than the original claims in reexamination. The added claims could strategically cover aspects of the accused infringing products not included in the issued claim set, although prohibitions against broadening amendments may bar such activity.

Finally, a reexamination request filed early on in the litigation could impact trial. This is especially true where the PTO decisions are favorable to patentability. However, as noted above, there are questions and concerns as to the admissibility of any non-final PTO action at trial.

H. Timing of Reexamination Requests – When to File?

Once a decision is made to proceed with a reexamination strategy, one of the most important considerations is deciding when to file. The timing of a reexamination request ultimately will be determined by the requester's overall goals.

Early filing should be considered where the goal is to stay a more costly district court litigation until the validity of the asserted patent is adjudicated by the PTO. Most courts will not consider staying the litigation until at least a first Office action rejection is received. Statistics indicate that the issuance of the first Office action could be one year or more after the request is granted and a filing date is accorded in an *ex parte* reexamination request. In an *inter partes* reexamination, the rules state that "[t]he order for *inter partes* reexamination will usually be accompanied by the initial Office action on the merits of the reexamination."[139] In the authors' experience, however, this is not always the case, particularly with the increased popularity of *inter partes* reexaminations in the past several years. But, based on anecdotal evidence, it appears the CRU is trending towards issuance of the initial Office action with the reexamination order.

Early filing is also highly recommended where the reexamination is launched as an insurance policy against an adverse district court decision. Overall, at least two years are typically necessary for a final decision from the CRU[140]—waiting too long to file a request could reduce the effectiveness of such a strategy. For such strategy, keeping tabs on the trial date is a must.

Sometimes, seeking a litigation stay is not feasible. Further, there is always a chance that an adverse decision by the CRU could have a negative impact on trial. In such cases, it may be desirable to delay reexamination filing to a point somewhat less than one

---

139 37 C.F.R. § 1.935.

140 The CRU has indicated an internal goal of 24 months from instituting the *inter partes* reexamination request to an Action Closing Prosecution (ACP).

year prior to trial. This mitigates the chance of an adverse CRU decision impacting trial, but may still be early enough for the reexamination to have a positive impact.

In some cases, parties have waited until after an adverse trial decision to file a reexamination request. Appeals to the Federal Circuit are notoriously uncertain, and a remand on an issue of claim construction or damages, for example, could result in a new trial on those issues. In this case, a reexamination may have time to run its course prior to a subsequent final decision or appeal.

Finally, reexaminations should also be considered as a settlement tool. Early preparation of a "pocket reexamination" to show to the patent owner could help drive negotiations in favor of an accused infringer. Even if the reexamination is not immediately filed, the efforts in preparing the pocket reexamination are directly applicable to an accused infringer's invalidity case and would likely not be wasted.

Given the above timing considerations, the authors recommend considering the following factors:

1. What overall goals should a reexamination strategy accomplish?
2. When is the trial scheduled and how firm is the trial date?
3. How has the court reacted to reexaminations in the past?
4. How strong is the prior art and are there one or more SNQs to support one or more proposed grounds of rejection?
5. Are *all* of the asserted claims subject to a SNQ?
6. How complicated is the invalidity case and what are the realistic chances of success before a jury or judge?
7. Are there pertinent dates on the discovery docket that might counsel delay in filing—e.g., after close of discovery to ensure all discovered prior art is included or after inventor depositions?

One additional consideration was raised in a recent Federal Circuit decision where a party attempted to obtain relief from a final judgment under Fed. R. Civ. P. 60(b)(6) based on alleged disclaimers made during post-trial reexamination.[141] In this case, the accused infringer waited until the district court's entry of judgment to file its reexamination request. The reexamination had progressed to a point where the patent owner had to respond to an Office action rejection. In that response, the patent owner allegedly made "representations to the [PTO]" the "limited the scope" of one of the accused claims.[142] The district court denied the Rule 60(b)(6) motion stating that the accused infringer "waited until after judgment in this case to file its Petition for Reexamination, while simultaneously failing to appeal the jury's finding of validity."[143] Because Rule 60 motions are reviewed under an abuse of discretion standard and typically require "extraordinary circumstances," the Federal Circuit affirmed the decision.

141 *Amado v. Microsoft Corp.*, 517 F.3d 1353 (Fed. Cir. 2008).
142 *Id.* at 1363.
143 *Id.*

Nonetheless, the authors can envision circumstances where a motion under Rule 60 might be granted. For instance, if the reexamination is timely filed and the patent owner voluntarily amends the asserted claims, or the asserted claims are finally declared invalid, after a final judgment is reached in the district court on the original claims, it seems relief from such a final judgment would be warranted. We are not aware of such a case, but relief under Rule 60 opens another potential window of time where a positive reexamination result could be useful.

I. Multiple **Ex Parte** *Reexamination Requests*

Where a party has a choice in filing an *ex parte* or *inter partes* reexamination request, what considerations go into the choice? One factor to consider is that there is no legal limit on the number of *ex parte* reexamination requests that can be filed. However, the bar for establishing a valid SNQ may become higher with each reexamination request. Because the Examiner makes a determination whether a reference raising a SNQ is cumulative to earlier considered art, the more references that have already been considered raises the bar for references in subsequent reexamination requests.

In what situations should multiple *ex parte* reexamination requests be considered? The ability to file more than one request can be a valuable tool where the patent owner is seen to mischaracterize the prior art, to make inconsistent statements between the reexamination and the parallel court proceeding, or where there is newly discovered prior art that surfaces after the previous reexamination request has been filed.

Another factor to consider is the impact that multiple or "rolling" reexamination requests may have on a district court judge. Could multiple reexamination requests impact an already granted litigation stay? At least one district court precluded an accused infringer from filing further *ex parte* reexamination requests as a condition of granting its motion to stay a litigation on the eve of trial.[144] The ability to file subsequent reexamination requests may be an important tool in an overall *ex parte* reexamination strategy. Relinquishing that ability should be carefully considered.

On March 1, 2005, the PTO issued the "Notice of Changes in Requirement for a Substantial New Question of Patentability for Second or Subsequent Request for Reexamination While an Earlier Filed Reexamination is Pending."[145] In the Notice, the PTO set forth a new policy:

> Under the new policy, the second or subsequent request for reexamination will be ordered only if that old prior art raises a substantial new question of patentability that is different than that raised in the pending reexamination proceeding. If the old prior art cited (in the second or subsequent request) raises only the same issues that were raised to initiate the pending reexamination proceeding, the second or subsequent request will be denied.
>
> … Further, 35 U.S.C. 303(a) states "[w]ithin three months following the filing of this title, the Director will determine whether a substantial new question of patentability affecting any claim of the patent concerned is

144 *See Visto Corp. v. Research in Motion Ltd.*, No. 06-181 (E.D. Tex. complaint filed Apr. 28, 2006).
145 Official Gazette Notice (Mar. 1, 2005).

> raised by the request." It is reasonable to interpret this provision as requiring each request for reexamination to raise its own substantial new question of patentability as compared not only to the original prosecution (in the application for the patent) and any earlier, concluded reexamination proceedings, but to pending reexamination proceedings as well.

Id.

When faced with the specter of multiple reexamination requests, a patent owner may consider filing a petition with the PTO arguing that subsequent reexamination requests are being filed for the purposes of harassment or to delay prosecution of a pending request.[146] If the prior art provided in the subsequent request unquestionably presents a new SNQ, the petition may not have a high likelihood of success. However, when coupled with an argument that the SNQs in the subsequent reexamination requests are cumulative, these petitions may be successful. This should also be a consideration when determining whether to file multiple reexamination requests.

J. Additional Strategic Questions to Consider

1. Withholding of prior art

Should an accused infringer withhold prior art from a reexamination request? If the reexamination request was an *ex parte* request, such art could become the basis for subsequent reexamination requests, if necessary and non-cumulative. If the reexamination request was an *inter partes* request, the withheld art may still be available for use at trial. Estoppel does not attach to the withheld or applied art until a final decision is reached in the *inter partes* reexamination proceeding.[147] Therefore, in litigation with concurrent *inter partes* reexamination proceedings, the withheld or applied art could be "ripped" from the litigation if the reexamination finishes before the litigation, and *vice versa*.

The authors are aware of some district court judges who have crafted stays to preclude the third party requester from using art that could or should have been brought during the reexamination proceeding. Also note that there are express limitations regarding subsequent submissions of prior art in *inter partes* reexaminations. Specifically, after the reexamination order, the third party requester may only cite additional prior art that is (1) necessary to rebut a finding of fact by the examiner, (2) necessary to rebut a response of the patent owner, or (3) which for the first time became known to the requester after filing the request provided certain conditions are met.[148]

Further, the trial team may be subject to the PTO's duty of disclosure requirements, as discussed above.[149] If the withheld art was disclosed during the litigation, for example as part of the accused infringer's invalidity contentions, then the patent owner or the patent owner's reexamination team may have the ability, or indeed the duty, to submit that withheld art and have it considered during the reexamination. One question to consider is whether the litigation team's knowledge of material prior art could be imputed to a patent owner. In short, many traps exist for the unwary, and patent owners must take care to avoid conduct that could result in inequitable conduct charges.

---

146 *See* MPEP § 2240.
147 35 U.S.C. § 317(b).
148 *See* 37 C.F.R. § 1.948.
149 *See* 37 C.F.R. §§ 1.555, 1.933; *see also* § II.A.3., *supra*.

2. Experts' independence

Should technical or legal experts have access to reexamination requests not yet filed prior to preparing reports? Also, do experts have a role in preparing a reexamination request? Since the *KSR* decision, it is important to create a full record in the reexamination proceeding of evidence supporting or negating nonobviousness and expert testimony. Affidavits or declarations may be the preferred mechanism for creating this record in many situations. Care must be taken by both parties to a litigation that expert testimony in the form of affidavits in the reexamination before the PTO is consistent with any expert testimony or reports to be used at trial, and vice versa. A potential impeachment or inequitable conduct minefield awaits the uninformed.

Another question to consider is whether an expert is an "individual associated with the patent owner" in the context of the Rule 56 duty of disclosure. Put differently, does a patent owner have the duty to inquire of their experts if they know of any references that should be cited to the PTO as part of the reexamination proceeding?

3. Privilege issues

Is there a waiver of privilege when a PTO submission, prosecution event, or meeting uses litigation work product or reveals trial strategy? The patent owner has an obligation to inform the PTO of any concurrent district court litigation in *ex parte* and *inter partes* reexaminations. While the CRU monitors the concurrent litigation, it is prudent for the patent owner to view this as an ongoing duty. For *inter partes* reexamination, any person can file a paper notifying the PTO of a concurrent proceeding.

If a party thinks information must be disclosed, the protective order should specify procedure for the parties to "meet and confer" to resolve any disclosure issues. If parties cannot agree, then the issue may need to be raised to the presiding judge. One way to protect information that a party feels must be disclosed, but is designated under the protective order, is to file the information under seal at the PTO with a petition to expunge at the conclusion of the proceeding.[150] The PTO specifies detailed procedures for filing information under seal in pending applications. The information to be protected is submitted in a labeled, sealed envelope. A petition to expunge the information accompanies the sealed documents. If the examiner does not believe the sealed information is material to patentability, the petition is granted and the information is expunged from the file. If the information is deemed material to patentability, the petition is denied, the information will become part of the application record, and the information will be available to the public.

It is important to note, however, that the information filed under seal may be made public at the conclusion of the reexamination process, if the information is deemed material to patentability. Therefore, the court and the party owning the confidential information should be involved in the decision to disclose and how the disclosure is made to the PTO.

4. Fast courts versus slow courts

How might the perceived speed of a court affect a decision to file a request for reexamination? Since the average pendency of a reexamination through the CRU, the BPAI,

150 *See* MPEP §§ 724.04 - 724.06.

and the Federal Circuit is 48 to 96 months, is the vehicle of reexamination more suitable for a slow court? As noted herein, the timing of a reexamination request depends on the requester's overall goals. If the purpose of a threatened reexamination is to drive settlement, then the speed of the court is irrelevant. If the goal of the reexamination request is to stay the litigation, then the speed of the request vis-à-vis the speed of the court is paramount.

ITC patent actions are notoriously fast and bear some special attention. As most readers are aware, ITC investigations proceed more quickly than most district court actions. Some reports indicate that an ITC investigation is generally completed within 15 months, whereas the average patent litigation in district court takes approximately 22 months.[151] In the past, the speed with which the ITC had to proceed was strictly mandated by statute. Since the statute was amended in 1994, the ITC now must "conclude any such investigation and make its determination under this section at the *earliest practicable time* after the date of publication of notice of such investigation."[152]

Due to this statutory mandate for a swift investigation, it should not be surprising that the ITC is hesitant to stay its investigations. One recent ITC case confirms the ITC's reluctance. In this case, the presiding administrative law judge ("ALJ") granted a stay pending reexamination of the patents-in-suit.[153] On appeal to the Commission, the stay was rejected, and proceedings were reinstated. The defendant petitioned the Federal Circuit for a *writ of mandamus* that the stay be reinstated. The Federal Circuit refused to reinstate the stay, finding that the Commission had justified its action and that hardship, inconvenience, and avoidance of a particularly complex trial are not sufficient reasons to grant *mandamus*.

At the same time, however, the ITC has not adopted any *per se* rule regarding staying an investigation in light of a concurrent reexamination at the PTO. Rather, the ALJ will weigh several factors, including: (1) the stage of discovery and the hearing date; (2) the issues in question and trial of the case; (3) the undue prejudice or clear tactical disadvantage of any party; (4) the stage of the reexamination at the PTO; (5) the efficient use of ITC resources; and (6) the availability of alternative remedies in federal court.[154]

Because of the speed with which the ITC conducts its investigations, litigants should expect an ITC investigation to proceed at its normally rapid pace with a low likelihood that the ALJ will stay the investigation.

5. Cases with multiple defendants

In cases with multiple accused infringers, how should a patent owner deal with reexamination threats by a single defendant, or a plurality of defendants? For the multiple defendants contemplating a reexamination strategy, what if the defendants are not of a single mind when it comes to reexamination strategy? Should reexamination be explicitly dealt with in any joint defense agreement? Again, any reexamination request can be timed so that it will not likely affect any trial proceedings. A requester could further allay fears by

151 *See* Vivek Koppikar, *Evaluating the International Trade Commission's Section 337 Investigation*, 86 J. PAT. & TRADEMARK OFF. SOC'Y 432, 433 (2004).

152 19 U.S.C. § 1337(b)(1) (emphasis added); see also 19 C.F.R. § 210.2 ("It is the policy of the Commission that, to the extent practicable and consistent with requirements of law, all investigations and related proceedings under this part shall be conducted expeditiously. The parties, their attorneys or other representatives, and the presiding administrative law judge shall make every effort at each stage of the investigation or related proceeding to avoid delay.").

153 *In re Freescale Semiconductor, Inc.*, 2008 WL 2951399 (Fed. Cir. June 25, 2008) (non-precedential) (order denying "petition for a writ of mandamus to direct the [ITC] to vacate its opinion denying petitioner's motion for stay pending reexamination of the patents at issue" in Investigation No. 337-TA-605).

154 *In re Certain Personal Computer/Consumer Electronic Convergent Devices, Components Thereof and Products Containing Same*, Inv. No. 337-YA-558, ALJ Order No. 6, 2006 ITC Lexis 52, at *12-*22 (U.S.I.T.C. Feb. 7, 2006) (order granting temporary stay).

committing to the other non-participating defendants not to request a litigation stay should the claims be rejected by the PTO. Of course, a consensus strategy is most desirable, but nothing is likely to bar a single defendant from launching a reexamination request if it believes its interests are best served by doing so.

One final consideration is how the "real party in interest" rule is addressed in multiple-defendant cases, as discussed above. Do the requester's non-participating co-defendants fall under the estoppel provisions? If not, could the same art be "litigated" at the PTO and at the district court by these non-participating defendants? We are not aware of this tactic having been tested, but joint defense groups may want to consider such a strategy.

6. The judge's perception of reexamination requests

Might a judge view a reexamination request as usurping the judge's authority? Does it help if the reexamination request is submitted by counsel not associated with trial counsel? Keep in mind that *ex parte* reexamination requests may be filed anonymously. Could it be in the requester's interest, where there is co-pending litigation, to anonymously file the reexamination request? Patent owners should consider interrogatories and/or production requests directed to whether the accused infringers have filed a reexamination request or caused a reexamination request to be filed.

Historically, many district court judges viewed reexaminations, particularly *ex parte* reexaminations, with disbelief and have been reluctant to grant stays especially if their court operates on a "fast track." More recently at Sedona discussions, some judges have expressed the view that they may rethink their approach in the future now that the CRU has been created and the PTO statistics seem to indicate prompt processing of reexaminations and a high probability of the reexamination resulting in some or all of the claims being found unpatentable. However, other judges are troubled by the time delay of reexaminations.[155] Litigants thus should pay special attention in the court filings relating to stays to explain the current reexamination environment.

7. Impact on laches

Do reexamination proceedings or the issuance of a reexamination certificate have any impact on the six-year statutory laches provisions of 35 U.S.C. § 286. A recent decision by the Federal Circuit seems to indicate that a reexamination certificate will not restart or impact the six-year laches presumption for enforcing an issued patent.[156] This appears consistent with the notion that a patent is fully enforceable on its original claims, even when it is involved in reexamination proceedings.

8. Duty of Disclosure

It is clear that the patent owner remains under a duty of disclosure while the patent is in reexamination proceedings under 37 C.F.R. §§ 1.555 and 1.933. Further, a third party requester participating in an *inter partes* reexamination owes a duty of candor and good faith to the PTO under 37 C.F.R. § 11.18. The ongoing duty of disclosure for the patent owner raises some interesting strategic questions that we consider below.

---

155 *See* Order Denying Sun Microsystem, Inc.'s Renewed Motion For Partial Stay in *Network Appliance Inc. v. Sun Microsystems, Inc.*, No. 07-06053 (N.D. Cal. Nov. 4, 2008) (citing Sedona PL08 version of this paper).

156 *See Serdarevic v. Advanced Med. Optics, Inc.*, 87 U.S.P.Q.2d 1481, 1484-85 (Fed. Cir. 2008) (holding that the presumption of laches applied against inventorship claim despite intervening reexamination proceeding).

For instance, who should handle the reexamination? Options include using patent attorneys from the trial team, or prosecution attorneys from the litigating law firm. Alternatively, outside patent attorneys or patent attorneys inside the patent owner's company could be used. Critical issues to consider include efficiencies, maintenance of privilege, and perhaps most importantly, compliance with protective orders. Best practices may warrant use of outside patent attorneys for the prosecution of the reexamination who are walled off from the litigation team, especially where the protective order includes a non-prosecution clause.

Another strategic question is determining what to cite to the CRU. The MPEP requires citation of "patents or printed publications which (A) are material to patentability in a reexamination proceeding, and (B) which have not previously been made of record in the patent file."[157] Validity decisions in reexamination may not be made on the basis of fraud on the Office, or on the basis of prior use or sale. Therefore, the patent owner arguably has fewer categories of potentially relevant material to cite to the PTO. However, it appears as if the obligation remains to disclose the same broad scope of prior art printed publications as would be the case in the original prosecution.[158]

However, because the prevalence of inequitable conduct charges in patent litigation remains unabated,[159] many reexamination practitioners reasonably err on the side of caution in preparing information disclosure statements. For instance, in the recent case of *Larson Mgf. Co. v. Aluminart Prods. LTD.*,[160] the district court found a reexamined patent to be unenforceable for failure to cite information from a related co-pending application. Specifically, the patent owner failed to cite a number of references applied in the co-pending application, as well as two Office actions where the application examiner had considered the same art as the CRU examiner. The Federal Circuit overturned the inequitable conduct charge finding that the uncited references were cumulative. But in doing so, the Court also determined that the Office actions *were* material. The case was remanded to consider whether the requisite intent was present as to the uncited Office actions. There was no mention by the Court of the fact that the MPEP does not appear to require citation of material "previously … made of record in the patent file."

Arguably the patent owner is required in the reexamination to disclose the same broad scope of prior art as would be the case in the original prosecution. The result of cases such as *Larson Mfg.* is that the filing of a voluminous IDS citing all types of prior art uncovered in the course of a concurrent patent litigation are regularly filed in reexamination proceedings—filings that include re-citation of all the art previously filed in any parent or otherwise related applications during regular prosecution. Where a patent owner feels compelled to cite material that may not be strictly required by the MPEP, it may simplify matters for the CRU. There are many examples of IDS filings in reexaminations where hundreds of possible prior art items disclosed during a concurrent litigation are dumped on the CRU by using separate IDS pleadings for different categories of information. For instance, one pleading may clearly cite to art previously cited during original prosecution of related applications, while a second pleading may cite to art newly discovery during the concurrent litigation or ITC investigation. The CRU examiner can then more easily discern what she may consider relevant to a particular case.

157 MPEP § 2280.
158 *See* MPEP § 2280 and 2684.
159 *Larson Mgf. Co. v. Aluminart Prods. LTD.*, 559 F.3d 1317 (Fed. Cir. 2009) ("The ease with which inequitable conduct can be pled, but not dismissed, is a problem of our own making.") (Concurring Opinion) (J. Linn).
160 *Id.*

Yet another strategic question revolves around what information to cite from a co-pending district court litigation or ITC investigation. The MPEP states that the duty of disclosure under Rule 555 "is consistent with the duty placed on patent applicants" by Rule 56. In that regard, Rule 56 requires citation of "[i]nformation from related litigation."[161] Such information includes "pleadings, admissions, discovery including interrogatories, depositions, and other documents and testimony."[162] Such categories of information may include assertions that are contradictory to assertions made to the examiner. Where there are concurrent litigation or ITC investigation proceedings, these classes of information must be continually monitored and cited to the office where relevant. Again, because of the prevalence of inequitable conduct charges, many practitioners err on the side of caution and tend to be over-inclusive in this regard. If information is cited that is beyond the scope of the reexamination proceedings—e.g., information related to allegations of prior sale or use—the CRU will simply note the issue as unresolved.[163]

The reexamination proceeding is strictly limited to considering prior art in the categories of patents, printed publications and admissions, and cannot address other forms of prior art such as public use, offer for sale, public knowledge, etc. Considerable CRU resources are wasted by Examiners considering art that cannot form the basis of a rejection of the original issued claims in a reexamination.[164] However, only submitting patents and printed publications in an IDS does not relieve the patent owner and its reexamination attorneys of the risks of violation of the duty of disclosure and possibly committing inequitable conduct.

Further 37 C.F.R. § 11.18 requires the patent owner's reexamination attorneys to conduct a reasonable inquiry of submissions to the PTO, which means that some level of review of each prior art submission should be done prior to filing. Currently it seems that due to the risk of inequitable conduct, patent owners are erring of the side of voluminous IDS filings and are not culling out possible prior art that does not fit into the categories of patents, printed publications and admissions.

Throughout the *inter partes* reexamination proceeding, the patent owner also has an obligation to "call the attention of the Office to any prior or concurrent proceedings in which the patent is or was involved, including but not limited to interferences, reissue, reexamination, or litigation and the results of such proceedings."[165] Additionally, any party may file a paper in an *inter partes* reexamination proceeding notifying the Office of the same. How much information is a party required to submit from a concurrent litigation? Is bare notice of a concurrent proceeding sufficient to meet the obligation imposed by Rule 985? Many patent owners are submitting voluminous court documents from concurrent litigation. Many of these documents would not qualify as prior art during original prosecution or reexamination prosecution.

One unresolved issue is the extent to which the duty of disclosure under Rules 555 and 933 would apply to a patent owner's trial team. What if the trial team is completely barred by the protective order from participating in the prosecution of a

161 MPEP § 2001.06(c).
162 *Id.*
163 MPEP § 2280.
164 If a patent owner chooses to add new claims or amend claims during a reexamination proceeding, the added language is subjected to a broader examination, similar to that of an examination of an original application. For example, the new claim or new limitation is considered for compliance under 35 U.S.C. §112. *See* MPEP §2258.II. Furthermore, some practitioners argue that non-patent, non-publication prior art (e.g., prior use, prior offer for sale) is material to new claims and/or new limitations added during reexaminations.
165 37 C.F.R. § 1.985; *see also* § 1.565 for similar *ex parte* rule.

concurrent reexamination? How high and impervious must the wall be between the trial team and the reexamination team to avoid implicating the duty of disclosure rules? What if one or more members of the trial team are registered patent attorneys? What disclosure obligations does a patent owner have, and can these obligations be avoided by remaining intentionally ignorant of prior art or other potentially relevant information that is confidentially disclosed over the course of the litigation?

These are very important questions as the trial team is likely to become aware of material prior art or other material information either through its own investigation or simply through the accused infringer's disclosure of its invalidity contentions. At least one recent district court decision suggests that the duty of disclosure does NOT trump a trial attorney's obligations under the protective order to maintain the confidentiality of protected information.[166] As a final matter, while possibly not under the same duty of disclosure as a patent owner, the requester nonetheless has the duty of candor during the PTO proceeding, as Rule 11.18[167] applies equally to reexaminations and all other proceedings before the PTO.[168]

V. BASIC REEXAMINATION PRACTICE

A. Generally

As noted above, reexamination can be *ex parte* or *inter partes*. In *ex parte* reexamination, a third party requester will receive copies of Office actions and patent owner replies, but cannot otherwise participate in the reexamination proceeding and cannot appeal PTO decisions. In *inter partes* reexamination, when the patent owner submits a reply to an Office action, the third party requester is entitled to file comments in response thereto. The third party's comments must, however, be limited to issues raised by the Office action or in the patent owner's response. The third party requester is entitled to certain appeals, but is also subject to certain estoppels. Both *ex parte* and *inter partes* reexaminations are discussed below.[169]

In order to assure timely delivery of papers, both the third party requester and patent owner should provide the PTO with current correspondence address information. In the past, both the patent owner and third party requester utilized the same form when changing correspondence address. The use of the same form resulted in situations where the patent owner correspondence address was used by the PTO as the third party requester correspondence address. The PTO has recently introduced a "Patent Owner Change of Correspondence Address" form and a separate "Third Party Requester Change of Correspondence Address" to address these concerns.

166 *See* Hot Topics Section II.C.
167 37 C.F.R. § 11.18.
168 The uncertainty between the required and also preferable bounds of the duty of disclosure in reexaminations and the ever present specter of a violation of the duty of disclosure, where even many at the PTO will say they do not know where the real lines are, is argued by some as just another example of why significant change in the law of inequitable conduct (or the duty of disclosure) needs to occur, whether by court decision or legislative change.
169 Other practitioners and commentators have provided summaries of the mechanics of reexamination practice. *See, e.g.*, J. Steven Baughman, *Reexamining Reexaminations: A Fresh Look at the* Ex Parte *and* Inter Partes *Mechanisms for Reviewing Issued Patents*, 89 J. PAT. & TRADEMARK OFF. SOC'Y 349, 360 (2007); Roger Shang & Yar Chaikovsky, Inter Partes *Reexamination of Patents: An Empirical Evaluation*, 15 *Tex. Intell. Prop. L.J.* 1 (2006); Sherry M. Knowles, *et al.*, Inter Partes *Patent Reexamination in the United States*, 86 J. PAT. & TRADEMARK OFF. SOC'Y 611 (2004).

B. The Request and the SNQ

Rule 1.510(b) sets forth the mandatory elements of an *ex parte* reexamination request, and Rule 1.915(b) sets forth the mandatory elements of an *inter partes* reexamination. Both *ex parte* and *inter partes* reexamination requests require (1) a statement pointing out each SNQ and (2) a detailed explanation of the pertinence and manner of applying cited patents and printed publications to every claim for which reexamination is requested. The SNQ must be based on prior patents and/or printed publications.[170] Other patentability issues, such as prior public use or insufficiency of the disclosure, will not be considered for instituting a reexamination. The PTO will only reexamine those claims for which a SNQ is alleged and found.

1. The substantial new question ("SNQ") generally

Both the *ex parte* and *inter partes* statutes require that a request for reexamination include at least one SNQ.[171] The legislative history of the *ex parte* reexamination statute describes the SNQ as new, non-cumulative information about preexisting technology that may have escaped review at the time of the original examination of the patent application and in subsequent reexaminations of the patent, if there have been any. The SNQ could therefore be more aptly named a "substantial new technical teaching." The establishment of a SNQ has tripped up many practitioners. It is not enough for a reference to be "new," the reference must also be non-cumulative to the technological teachings previously considered by the PTO during prosecution. Therefore, even a newly discovered reference may not raise a SNQ if the reference merely is cumulative to similar prior art already fully considered by the PTO in a previous examination or reexamination.[172] This is an important point when determining whether to file a reexamination request and which references to use.

The CRU rejects many reexamination requests on first filing for failure to clearly point out a SNQ. Specifically, the authors note a perceived increase in the number of reexamination requests that are receiving a Notice of Incomplete Reexamination Request or requests that are being denied because the requester has failed to establish the reference used as a basis for a SNQ is "non-cumulative."

Further, there is a common but fundamental misunderstanding by many as to the difference between a SNQ and a proposed ground of rejection. A reexamination request must include both at least one SNQ in addition to a detailed explanation of the pertinency and manner of applying a proposed SNQ to every claim for which reexamination is requested—the proposed ground of rejection.[173] A SNQ is therefore a separate and distinct concept from the proposed ground of rejection (a *prima facie* case of unpatentability).

Not all previously considered references ("old art") are ineligible to support a SNQ. Old art previously considered in original or prior prosecution may be used to support a SNQ if shown in a new light. Previously unconsidered art may not provide de facto support for a SNQ if it is merely cumulative to art already considered by the Office. We discuss new light for old art in the following section.

Finally, the PTO will consider an undated document in reexamination if it is accompanied by reliable evidence, such as an affidavit or deposition transcripts, supporting

170 37 C.F.R. §§ 1.510(b)(1), 1.915(b)(3).
171 35 U.S.C. §§ 303 and 312.
172 MPEP § 2242.
173 37 C.F.R. §§ 1.510(b) and 1.915(b).

an asserted publication date. Where a reference originated with the patent owner, the CRU might consider issuing a request to the patent owner for additional information under 37 C.F.R. § 1.105. For reexaminations ordered on or after November 2, 2002, a finding of a SNQ, and claim rejections, can be based solely on previously cited/considered "old" prior art, or in combination with other prior art.[174]

2. In re Swanson *and the SNQ*

The Federal Circuit in *In re Swanson*[175] clarified what it takes to support a SNQ where a reference was previously used to reject the claims in original prosecution. The *ex parte* reexamination statute sets forth the universe of references that can be used to raise a SNQ.[176] In addition to a newly discovered reference, a previously applied reference can raise a SNQ if the previously applied reference is presented in a "new light." Section 303(a) makes this explicit—"[t]he existence of a substantial new question of patentability is not precluded by the fact that a patent or printed publication was previously cited by or to the PTO or considered by the PTO."[177] *In re Swanson* provides some guidance on what constitutes a "new light" for old art. For example, a SNQ based on previously applied art could arise because the examiner in the original examination misunderstood the actual technical teaching, because the examiner failed to consider a portion of the reference that contained the now cited teaching, or if the examiner applied the reference to a different limitation or claim than that to which the reference is currently being applied.

But a reference does not raise a SNQ if the examiner in the original examination understood the actual technical teaching but got it "wrong" in the rejection. This is a subtle but critical distinction.

3. KSR *and the SNQ*

A further unsettled issue for many practitioners is the impact of *KSR* on reexamination practice. Did *KSR* open the door to reexamination challenges based on prior art overcome during original prosecution by arguing lack of teaching, suggestion, or motivation to combine?[178] The PTO pondered this critical issue for over sixteen months and

174 MPEP §§ 2242(II)(A), 2258.01(A).

175 *In re Swanson*, 540 F.3d. 1368 (Fed. Cir. 2008).

176 35 U.S.C. § 303(a) (patents and printed publications).

177 This sentence was added in the 2002 amendment to 35 U.S.C. § 303 to specifically address *In re Portola Packaging, Inc.*, 110 F.3d 786 (Fed. Cir. 1997). In *Portola*, the Federal Circuit "interpreted the statutory intent [of the *ex parte* reexamination statutes] as precluding reexamination based on 'prior art previously considered by the PTO in relation to the same or broader claims.'" *In re Swanson*, at 11 (citing *Portola*, 110 F.3d at 791). Congress explained that the amendment to 35 U.S.C. § 303(a) "overturns the holding of *In re Portola Packaging, Inc.,* a 1997 Federal court decision imposing an overly-strict limit that reaches beyond the text of the Patent Act." H.R. Rep. No. 107-120, at 2.

178 When *KSR* was decided at the end of April 2007, the PTO feared an avalanche of reexaminations based solely on an argument that the obviousness standard applied in the original prosecution had been relaxed. This argument was advanced by the third party requester in *Ex Parte* Reexamination Control No. 90/008,949. In this request, the third party requester argued that the Supreme Court's decision in *KSR* provided a "new light" in which to view the references under the doctrine of obviousness. *See* Order Denying Request for *Ex Parte* Reexamination in Reexamination Control No. 90/008,949, at 4. The *Ex Parte* Reexamination request was denied by the PTO. *Id.*, at 5. In the denial, the PTO clarified the standard for determination of whether a SNQ exists based on "old art":

> Reexamination is limited to review of new information about preexisting technology, which may have escaped review at the time of initial examination of the patent application. It was not designed for harassment of a patent owner by review of old information about preexisting technology, even if a third party feels the Office's conclusion based on that old information was erroneous. The reexamination legislative history nowhere provides for review of such old information, each time a court clarifies or re-interprets a standard or point of law that effects the patentability determination. If it did, the reexamination process would be unwieldy, because case law is constantly evolving.
>
> … The KSR decision does not *per se* create new information about preexisting technology that may have escaped review at the time of the initial examination of the patent application. And, in this instance, the KSR decision does not present or view the "old art" in a different way, or in a "new light," as compared to what was already considered in the '7628 reexamination proceeding."

Id., pp. 6-9. (emphasis in original). The feared avalanche did not materialize but there is no doubt that *KSR* spawned more reexaminations than would have occurred otherwise.

then addressed this question explicitly in Revision 7 of the M.P.E.P. (Manual of Patent Examining Procedure or MPEP), which became publicly available in August 2008. The MPEP now states:

> The clarification of the legal standard for determining obviousness under 35 U.S.C. § 103 in *KSR International Co. v. Teleflex Inc.* (KSR), 550 U.S. 550, 82 USPQ2d 1385 (2007) does not alter the legal standard for determining whether a substantial new question of patentability exists. The requirement for a substantial new question of patentability remains in place even if it is clear from the record of a patent for which reexamination is requested that the patent was granted because the Office did not show "motivation" to combine, or otherwise satisfy the teaching, suggestion, or motivation (TSM) test. Thus, a reexamination request relying on previously applied prior art that asks the Office to look at the art again based solely on the Supreme Court's clarification of the legal standard for determining obviousness under 35 U.S.C. § 103 in *KSR*, without presenting the art in new light or different way, will not raise a substantial new question of patentability as to the patent claims, and reexamination will not be ordered.[179]

Following that amendment to the MPEP, the Federal Circuit issued its decision in *In re Swanson* on September 4, 2008. The *In re Swanson* decision did not address the impact of *KSR* on the determination of whether references raise a SNQ. The issue therefore remains unsettled.

KSR states that patent examiners, as well as the courts, can review the factual predicates underlying the obviousness calculus and reach the ultimate legal conclusion whether the subject matter is obvious.[180] Thus, it makes logical sense that it would be of great interest to the reexamination examiners to know what a person of ordinary skill in the art ("POSITA") would have known at the time of filing of the original application for which reexamination is requested. The third party requester is advised to consider providing a description of what the POSITA would have known preferably in the reexamination request (or less preferably in a later response to an Office action). This POSITA technical description can be presented in a separate section of the reexamination request, but regardless of how it is provided, it is necessary that the SNQ basis be set forth for each technical reference referred to in this technical description and that each of these SNQs be used in a least one proposed rejection. What is believed to be the first reexamination request that employed such a POSITA technical description is found in *Inter Partes* Reexamination Control No. 95/000,353 ("the '353 request").

The '353 request provided an extensive discussion of the various factors, articulated by the Federal Circuit and discussed in the Examination Guidelines, which may be considered in determining the level of ordinary skill in the art.[181] The specific factors addressed in the '353 request included the types of problems encountered in the art, prior art solutions to those problems, rapidity with which innovations are made, the sophistication of the technology, and the educational level of active workers in the field.

179 MPEP § 2216, Rev. 7, July 2008.

180 *KSR*, 540 U.S. at 413 ("While the sequence of [the Graham factors] might be reordered in any particular case, the factors continue to define the inquiry that controls. If a court, or patent examiner, conducts this analysis and concludes the claimed subject matter was obvious, the claim is invalid under § 103.").

181 *See e.g.*, *In re GPAC*, 57 F.3d 1573, 1579 (Fed. Cir. 1995); *Customer Accessories, Inc. v. Jeffrey-Allan Indus., Inc.,* 807 F.2d 955, 962 (Fed. Cir. 1986); *Envtl. Designs, Ltd. v. Union Oil Co.,* 713 F.2d 693, 696 (Fed. Cir. 1983).

It is not uncommon for reexamination requests, particularly *inter partes* reexamination requests, to be hundreds of pages long. Some commentators note that such lengthy requests are unduly long and amount to an abuse of the reexamination process. However, other commentators note that, particularly in *inter partes* reexamination where the requester is faced with "use it or lose it" estoppel provision for known references, a third party requester is forced into lengthy requests in order to fully develop all SNQs available when the reexamination request is filed.

C. Impact of **KSR** *on Reexamination Practice*

The Supreme Court's decision in *KSR v. Teleflex* altered the obviousness calculus in a fundamental way by making the obviousness determination more subjective. The full scope of *KSR*'s impact on patent reexamination remains to be seen, but initial reports indicate that the results may be dire for patent owners, particularly in the predictable arts. Extensive research about the impact of *KSR* has been done by many groups. That research demonstrates the impact is significant in original prosecution. For instance, perhaps the most extensive publicly available sample and analysis was done by Microsoft's Corporate Vice President for IP Policy and Strategy, Marshal Phelps and his team. His research was presented at Sedona Patent Litigation 2008.[182] As his analysis shows, the most effective way to challenge an obviousness rejection in predictable arts is to persuasively argue, with factual support, that a claim feature is not taught by the references. If the references in fact show each element, either explicitly or inherently, then it seems to be very difficult to overcome an obviousness rejection.[183]

It is clear from recent Federal Circuit and BPAI decisions that mere attorney argument is not sufficient in many cases to prove non-obviousness.[184] The attorney is typically not an expert in the technology of the claimed invention and is not a person of ordinary skill in the art.[185] To prove non-obviousness it seems the best approach is to tell the story of the invention in its full glory so that the factual predicates are found in the record to support the desired legal conclusion of non-obviousness. While *KSR* makes many statements about what is or is not obvious, it is clear from Supreme Court law that what the decision-maker requires is all of the relevant facts about the invention and its predecessor technology. Thus it behooves the patent owner to put all of the necessary factual predicates into the reexamination record to support the desired legal conclusion of non-obviousness. Failure to do so could result in the CRU finding the claims not patentable and the BPAI and Federal Circuit on appeal being limited to a record that will not permit a reversal.

KSR is seen by some judges as providing examples of what might constitute good factual predicates to support non-obviousness, but not as a definitive guide on how best to set forth the full story of the invention. These factual predicates include the so-called "secondary considerations" or "objective evidence" of non-obviousness, such as unexpected results, long felt need, failure by others and commercial success. But this list is not definitive and counsel for the patent owner should be vigilant and creative in ferreting out and presenting all factual evidence that supports patentability.

182 *See* Microsoft Obviousness Data Research Slide Deck in Obviousness Panel tab of course notebook of Sedona PL08.
183 Many argue that the mere existence of all of the claim elements in the prior art as the basis for a finding of obviousness turns this test of patentability on its head since most inventions are "combination of old elements." This applies to some of the most important inventions of all time.
184 *See Sundance, Inc. v. Demonte Fabricating LTD.*, 550 F.3d 1356 at 4-7 (Fed. Cir. 2008).
185 *Id.* (disallowing a patent expert's testimony stating that "[d]espite the absence of any suggestion of relevant technical expertise, [the patent expert] offered expert testimony on several issues which are exclusively determined from the perspective of ordinary skill in the art.").

One crucial consideration is how to get factual evidence into the record during reexamination. Factual evidence can be in the form of trial evidence or testimony, publication, award, sales information, product reviews, etc. Should it be from an expert or at least from a witness considered to be a POSITA? The answer often is yes, although it is a tactical decision to have only a POSITA qualification since a qualified technical expert can typically also opine as a POSITA. Reexamination counsel for the patent owner is ill advised to assume either role explicitly or through attorney argument unless she can be qualified as if she is testifying as such in court.[186]

How should this factual evidence from the expert or POSITA be provided to the CRU? Probably it is best if it is in the form of an affidavit or declaration. But such a submission raises several concerns.

First is the specter of inequitable conduct. Reexamination counsel for the patent owner is particularly vulnerable because the law is somewhat confused in this area of what constitutes sufficient disclosure of pecuniary benefit between affiant/declarant and the patent owner. Future versions of this paper will address this more fully. It behooves the drafter to err on the side of comprehensive disclosure, although such approach increases the size of the administrative record, something the PTO has indicated it would like to avoid, all things being equal.[187]

The second concern is that the reexamination examiners have no mechanism and little experience in assessing the competency and veracity of the information and analysis presented in written submissions.[188] The third concern is the strict page limits imposed on responses to Office actions. We note however, that if the submission is denominated as "factual" as opposed to "argument" it is NOT counted in the page limit.

We expect that the obviousness area of patent practice will experience extensive attention in the next year as applicants, patent owners, and challengers grapple with the practical implications of *KSR* in PTO examinations, in the CRU, at the BPAI, in the federal courts and at the ITC.

D. **Ex Parte** ***Reexamination***

Ex parte reexamination can be requested by a patent owner or any third party requester at any time during the enforceability of a patent.[189] Subsequent requests for *ex parte* reexamination by a third party requester are permitted, provided the prior art raising the new SNQ is not cumulative to prior art previously considered. Co-pending reexamination proceedings may be merged.[190] The patent owner is not permitted to broaden the scope of claims during *ex parte* reexamination.[191] A third party requester can petition the

---

186 *Id.*

187 Since the duty of disclosure does not apply to reexamination counsel for third party requester, can "reverse" inequitable conduct be found by the PTO, BPAI or the courts for requester submissions clearly hiding the ball from the reexamination examiners? Or is the only possible violation that of 37 C.F.R. § 11.18? And if the latter is the case, how would such a violation be raised? Would it be done by OPLA based on a Petition from the patent owner that would be referred to The Office of Enrollment and Discipline (OED)? We know the PTO is thinking about these issues but we do not know of any public information from OED showing such a violation has been successfully prosecuted. This different standard of care between reexamination counsel for patent owner and reexamination counsel for third party requester troubles many people and we expect it to be addressed by some tribunal soon.

188 This is a broader problem than the mechanisms available to and experience possessed by reexamination examiners in that in *ex parte* prosecution there traditionally has been little use of affidavits and declarations. But post-*KSR,* this could change dramatically especially in light of recent BPAI and Federal Circuit decisions on obviousness.

189 35 U.S.C. § 302; 37 C.F.R. § 1.510.

190 37 C.F.R. § 1.565.

191 *Id.* § 1.552(b).

PTO Director to review a determination refusing *ex parte* reexamination.[192] The Director's decision on the petition is non-appealable, but can be challenged via a district court action.

In response to a grant of *ex parte* reexamination, the patent owner is entitled to file a statement on the new question of patentability, including any proposed amendments the patent owner wishes to make.[193] Where the *ex parte* reexamination was requested by a third party, the third party is entitled to respond to the patent owner's statement.[194] This may be one reason why patent owners rarely submit a statement prior to receipt of an Office action.

The patent owner is entitled to appeal to the BPAI and to the Federal Circuit with respect to any decision adverse to the patentability of any original, proposed amended or new claim of the patent.[195]

E. Director-Initiated **Ex Parte** *Reexamination*

The PTO Director can institute *ex parte* reexaminations *sua sponte*.[196] Under current PTO practice, the Director can institute a reexamination only upon a finding that a patent brings disrepute on the PTO or that a significant procedural error occurred during examination. For example, where a patent examiner failed to consider references submitted in an information disclosure statement during the examination process, despite numerous requests from the applicant, the PTO may independently determine whether the references raise a SNQ. Some argue that, since the creation of the CRU, the use of Director-initiated *ex parte* reexaminations has ceased based on the belief that the affected public can assume the burden of policing patents that are adverse to them. However, others argue that the lack of Director-initiated *ex parte* reexaminations since creation of the CRU is simply a coincidence.

F. Inter Partes *Reexamination*

1. Generally

Inter partes reexamination can be requested by any party other than the patent owner and its privies, at any time during the period of enforceability of a patent.[197] *Inter partes* reexamination is only available for patents that issued from an original application filed in the United States on or after November 29, 1999.[198] Until recently, the meaning of "an original application" was not fully settled. The issue was whether an *inter partes* reexamination can be filed on a patent from a continuation application having a filing date on or after November 29, 1999, but which claims priority to a filing before November 29, 1999. The PTO's position was that "an original application" includes any application with an actual filing date on or after November 29, 1999, regardless of whether that application claimed priority to an application filed before that date. The PTO's interpretation was upheld in district court[199] and was subsequently affirmed by the Federal Circuit in *Cooper Techs. Co. v. Dudas*.[200] The practical effect is that only patents whose actual filing date is on or after November 29, 1999, is eligible for *inter partes* reexamination, irrespective of whether the patent's effective filing date is earlier.

---

192 *Id.* § 1.515(c).
193 *Id.* § 1.530.
194 *Id.* § 1.535.
195 35 U.S.C. § 306.
196 35 U.S.C. § 303; 37 C.F.R. § 1.520.
197 37 C.F.R. § 1.913.
198 *Id.*
199 *Cooper Techs. Co. v. Dudas*, 85 U.S.P.Q.2d 1465 (E.D. Va. Nov. 30, 2007).
200 *See Cooper Techs. Co. v. Dudas*, 536 F.3d 1330 at 2-7 (Fed. Cir. 2008).

As with *ex parte* reexaminations, a third party requester can petition the Director to review a determination refusing *inter partes* reexamination. The Director's decision of the petition is non-appealable.[201]

A patent owner is entitled to appeal to the BPAI and to the Federal Circuit with respect to any decision adverse to the patentability of any original or proposed amended or new claim of the patent.[202] A third party requester is entitled to appeal to the BPAI and to the Federal Circuit with respect to any *final* decision favorable to the patentability of any original or proposed amended or new claim of the patent.[203]

2. Estoppels in **inter partes** *reexamination*

Estoppels in an *inter partes* reexamination flow two ways—from the *inter partes* reexamination to the civil action and from the civil action to the *inter partes* reexamination. In a subsequent civil action, a party cannot argue invalidity of a claim finally determined to be valid or patentable on any ground which that party raised or could have raised during the *inter partes* reexamination.[204] The third party is not, however, precluded from asserting invalidity based on newly discovered art that was unavailable to the third party requester and the PTO at the time of the *inter partes* reexamination.[205] Similarly, once a final decision has been entered against a party in a civil action that the party has not sustained its burden of proving the invalidity of any patent claim in suit, the party and its privies cannot request *inter partes* reexamination on the basis of issues that the third party raised or could have raised in the civil action.[206] It appears from the statute that these estoppels in *inter partes* cases apply only to civil actions brought in the district court under 28 U.S.C. § 1338 and not to Section 337 ITC investigations. Whether this was an oversight is unknown, but the authors are not aware of any cases in which estoppel has been applied in an ITC investigation.

Once a request for *inter partes* reexamination has been granted, the third party requester cannot file a subsequent request for *inter partes* reexamination while the prior *inter partes* reexamination is pending.[207] Further, once a final decision is granted favorable to patentability of any original, proposed amended, or new claim, the third party requester cannot thereafter request *inter partes* reexamination of such claim on the basis of issues that the third party raised or could have raised in the prior *inter partes* reexamination.[208]

Finally, after an *inter partes* reexamination has been instituted, the third party requester is precluded from citing any additional prior art unless it rebuts a finding of the examiner or a response by the patent owner, or if it became known or available after filing the request.[209] There are no estoppels, however, that prevent the third party requester from filing subsequent *ex parte* reexaminations. But the CRU closely examines such subsequent requests with a close eye to whether there is, in fact, a true SNQ.

---

201 37 C.F.R. § 1.927.
202 35 U.S.C. § 315(a).
203 *Id.* § 315(b).
204 35 U.S.C. § 315(c)..
205 *Id.* § 315(c).
206 35 U.S.C. § 317(b).
207 *Id.* § 317(a).
208 *Id.* § 317(b).
209 *Id.* § 1.948.

3. Real Party in Interest

While *ex parte* reexamination requests may be filed anonymously by any party, at any time,[210] *inter partes* reexamination requests must identify the real party in interest filing the request. The real party in interest requirement is closely tied to the *inter partes* reexamination estoppel provisions.[211] The estoppel provisions of 35 U.S.C. § 317(b) apply to the "third party requester" and "its privies." The real party in interest disclosure therefore must be "to the extent necessary for a subsequent person filing an *inter partes* reexamination request to determine whether that person is a privy."[212]

The PTO currently is struggling with the real party in interest rule where there is concurrent district court litigation with multiple defendants. Typically, in a multi-defendant litigation, the group of defendants will create a joint defense group ("JDG"). Issues occur when one or more defendants, but not all, file an *inter partes* reexamination request. For instance, the filing defendant (or defendants) may have minimal infringement exposure but is alleged or believed to be acting as a surrogate for the other defendants who are not officially part of the *inter partes* reexamination and thus not bound by the estoppel provisions.[213] An alternative scenario is that the defendant filing the *inter partes* reexamination request is doing so on its own initiative, and perhaps against the wishes of one or more co-defendants.

In such situations, who is the real party in interest? Just the third party requester or the entire JDG? Just those JDG members who have concurred with the reexamination strategy? Just those JDG members who have provided prior art, research, review, analysis, drafts, staffing support, financial backing, concurrence on actual filings, approval of filings, etc.? The authors are aware of several cases involving various flavors of this scenario where the patent owner has filed a petition to vacate the reexamination order, or suspend the reexamination, on the grounds that the real party in interest has not been identified and the PTO therefore lacks jurisdiction to continue reexamination proceedings.[214] Where the parties disagree on the facts, the PTO takes the position that it has not been vested with the tools, such as a subpoena power, statutory authority, or a discovery process in reexamination proceedings, necessary to make a proper factual determination. Nor is there clear guidance in the *inter partes* reexamination statute, its legislative history, the PTO rules, or from the courts as to how to resolve such real-party-in-interest issues which are real but are often ignored in practice.

Perhaps most importantly, petitions challenging the real party in interest are not handled by the CRU but are referred to the Office of Patent Legal Administration ("OPLA"). These petitions take time and often significant resources to resolve and are perceived by some as being directly contrary to the statutory requirement that the PTO act with "special dispatch." As one OPLA official has stated, such petitions act as an "anchor" on reexamination proceedings that bring it to a standstill or even prevent the reexamination from getting underway.

---

210 *See* 37 C.F.R. § 1.510(b) (setting forth information required for an *ex parte* reexamination request and not including the identity of that requester); *see also* 37 C.F.R. § 1.510(f) (requiring attorney or agent to have power of attorney if "a request is filed by an attorney or agent identifying another party on whose behalf the request is filed").
211 *See* M.P.E.P. § 2612 ("note that it is the real party in interest that is subject to the estoppel provisions and not the party who actually filed the request.").
212 37 C.F.R. § 1.915(b)(8).
213 35 U.S.C. §§ 315(b) and 316(b).
214 The real party in interest issue arises in *inter partes* Reexamination Control Nos. 95/000,227 and 95/000,229.

There is at least one case where the PTO has dismissed an *inter partes* reexamination where the real party in interest was not identified to its satisfaction.[215] The request was filed by an entity calling itself "Troll Busters." The requester's website describes its service as completely anonymous: "Troll Busters takes aim and fires in our own name. The Patent Troll will never know who or how many are behind the 'hit'."[216] In practice, the PTO generally will not look beyond the required statement identifying the real party in interest unless it is not facially accurate or is ambiguous. In the Troll Busters case, based on the information posted on the Troll Busters web pages, the PTO issued a show cause order to establish the identity of the real party in interest. The PTO was not persuaded by the response and terminated the reexamination stating that "Troll Busters cannot act as a 'shill' in an *inter partes* reexamination request to shield the identity of the real party or parties in interest."[217] This is the only case of which the authors are aware in which the PTO has terminated an *inter partes* reexamination request based on a finding of a violation of the real party in interest requirement.

The Troll Busters case establishes several important considerations for challenging the real party in interest. First, "extrinsic evidence may be submitted by the patent owner to support a petition to vacate the filing date or the Office may use extrinsic evidence to, *sua sponte*, order the requester/real party in interest to show cause."[218] Second, the PTO stated that "[a]n entity named as the sole real party in interest may not receive a suggestion from another party that a particular patent should be the subject of a request for *inter partes* reexamination and be compensated by that party for the filing of the request … without naming the party who suggested and compensated the entity for the filing of a request."[219] Finally, the PTO explicitly noted that *ex parte* reexamination was still an option for Troll Busters.

In sum, the PTO has a limited ability and appetite to resolve real party in interest disputes in the context of *inter partes* reexamination proceedings because the PTO does not have the discovery mechanisms in reexamination proceedings and resources to mediate or decide such disputes. However, the Troll Busters case shows that an insufficient response to a show cause order may result in dismissal of the *inter partes* reexamination.

When a final decision in an *inter partes* reexamination is favorable to patentability, the estoppel provisions of 35 U.S.C. § 315(c) attaches to the civil action. Although the PTO has limited ability to investigate and resolve real party in interest issues for PTO proceedings, some practitioners argue that a district court should investigate the true interested party prior to instituting the estoppels in the litigation, particularly in multi-party litigation. This is a hot area of dispute, and the authors expect to see more developments over the next year.

G. Mergers of Concurrent Proceedings

1. Merger of Co-Pending Reexaminations

Multiple *ex parte* reexaminations on the same patent will generally be merged.[220] However, the decision is at the sole discretion of the Office, which will take into account its statutory mandate of "special dispatch."[221]

215 *See* Reexamination Control No. 95/001,045.
216 See http://www.troll-busters.com/.
217 *See* Control No. 95/001,045, "Decision Vacating Filing Date," at 7 (mailed Aug. 25, 2008).
218 *Id.*
219 *Id.*
220 37 C.F.R. § 1.565.
221 MPEP § 2283.

If one of the multiple co-pending reexamination requests is *inter partes*, then the merger decision proceeds under the *inter partes* rules.[222] According to the MPEP, merger decisions are made by the OPLA. After multiple reexaminations have been ordered, the CRU will deliver the multiple orders to the OPLA, and OPLA will determine whether and how the reexamination should be merged. OPLA will thereafter issue a merger order that will govern the merged proceedings. The merged proceedings will then proceed under the *inter partes* rules, except that the third party requester will maintain its rights under the *ex parte* rules—e.g., the third party requester would maintain its right to respond to any patent owner's statement.[223] No *ex parte* rights appear to remain with the patent owner and the right to interview any *ex parte* issues appears to be foreclosed as a general rule. If a party feels its *ex parte* rights have been unfairly terminated by a merger, then it may consider a Rule 181 petition to preserve ex parte rights.

2. Merger of Co-Pending Reissue Applications and Reexaminations

The authors are aware of multiple situations where a patent owner has, upon being subjected to a reexamination request, proceeded to file a reissue application.[224] It is also not unusual for a patent owner to file a reissue application in advance of a lawsuit to clear up any errors or to put a claim set in better condition for litigation. If the reissue application is filed within the two years of the issue date of the patent, a broadening reissue is available.[225] While a patent owner may not broaden claims during reexamination, a broadening reissue application may preserve that right.

Where a reissue application and a reexamination are co-pending, the PTO may merge the proceedings or suspend one of the two proceedings.[226] It is the general policy of the PTO that the two proceedings will not be allowed to proceed simultaneously without merger or suspension of one or the other.[227] The reason for the policy is to permit timely resolution of both the reissue and the reexamination, and to prevent inconsistent and possibly conflicting amendments. Therefore, even if the parties do not make a specific petition for merger or suspension, the PTO will take action *sua sponte* to prevent parallel proceedings.[228] Because of the statutory mandate to treat reexaminations with "special dispatch,"[229] the PTO should not suspend the reexamination unless there were exceptional circumstances. This is especially true due to the potential for extending prosecution of reissue applications through continuation applications or requests for continued examination.

As with the merger of multiple reexamination requests, the determination to merge or suspend is made by the OPLA, and the decision is made on a case by case basis.[230] Factors include: (i) timing (whether reissue was filed first); (ii) the statutory mandate to treat reexaminations with "special dispatch"; (iii) the fact that the reissue could continue indefinitely via continuation applications; and (iv) whether the patent owner consented to a stay of the reissue application. A merger order will typically lay out the ground rules for the merged proceeding to proceed simultaneously. Importantly, jurisdiction of a merged proceeding stays with the CRU, not with the Technology Center reissue examiner.

---

222 37 C.F.R. § 1.989.
223 *Id.*
224 *See, e.g.*, Reissue application no. 11/513,425 and *inter partes* reexamination control no. 95/000,155.
225 *See* 35 U.S.C. § 251.
226 *See* 37 C.F.R. § 1.565(d) for reissue merger with *ex parte* reexamination and 37 C.F.R. § 1.991 for reissue merger with *inter partes* reexamination.
227 MPEP § 2686.03.
228 *See, e.g.*, merger decision for *inter partes* reexamination control number 95/000,270, which was merged (sua sponte) with broadening reissue application serial number 11/703,606; and merger decision for *inter partes* reexamination control number 95/000,269, which was merged (sua sponte) with reissue application serial number 11/636,350. Both of these cases involved concurrent district court litigation.
229 35 U.S.C. §§ 305, 314(c).
230 *In re Allan*, 71 U.S.P.Q.2d 1751 (Com'r Pat. & Trademarks Apr. 26, 2004) (non-precedential).

Where the merger is with an *inter partes* reexamination, the OPLA's merger order strictly limits the third party requester's participation to those issues unique to the reexamination. For example, a third party requester will be strictly prohibited from commenting on the reissue claims or other issues unique to reissue, such as recapture. Further, the merged proceeding will not allow for interviews or extensions of time (except for good cause). Finally, in merged proceedings, the reexamination is terminated upon the issuance of a reexamination certificate.

H. Extensions of Time

Reexaminations must be carried out with "special dispatch."[231] For this reason, patent owners are uniformly given shortened periods of time to respond to Office actions on the merits—typically two months instead of the usual three allowed during original prosecution. Further, because reexaminations are not "applications," the ability to obtain extensions of time under 37 C.F.R. § 1.136 is foreclosed. Instead, patent owners seeking an extension of time are bound by reexamination Rules 1.560 and 1.956. Under those rules, extensions of time are only given if sufficient cause is demonstrated. Even then, extensions are typically available for only one month barring a showing of extraordinary circumstances.[232]

That said, well supported requests are often granted. Factors that appear to warrant extensions of time include, for example, ongoing parallel discovery in a co-pending district court or ITC action. Discovery such as expert reports or inventor depositions that are due during the two month window for responding to the Office action may necessitate added time for review and consideration.[233] Other factors include the availability of declarants, the need to obtain evidence in support of patentability arguments and whether reexamination is new to the proceedings. For instance, extensions of time have been granted when reexamination counsel needs to investigate obtaining evidence to support a declaration under Rule 131 to swear behind the date of a reference, or needs to obtain evidence to support a declaration of non-obviousness under Rule 132. The same is true when the need arises to obtain evidence of secondary considerations in rebutting an obviousness rejection.

An important aspect of any request to extend time is an affirmative showing of what steps the patent owner has taken in responding to the Office action. *See* MPEP § 2265. A detailed showing of the steps being taken to respond to the Office action is required. Only with the this showing may the CRU properly balance the need to proceed with special dispatch against the ability of the patent owner to fully respond to the rejections in the Office action and to adequately defend its patent right.

It should also be noted that filing a request for an extension of time does NOT toll the time to respond. Therefore, any requests should be made well in advance of the deadline. The CRU responds to requests in a fairly prompt manner. One recent evaluation shows a mean decision time of 14 days, with a median time of 11 days. The authors have seen the CRU act in as little as five days from the request.

If an extension of time is secured for responding to a particular action, any subsequent request to extend for the same action will only be granted if there are extraordinary circumstances, such as incapacitation of reexamination counsel.[234] For

231 35 U.S.C. §§ 305 and 314(c).
232 *See* MPEP §§ 2265 and 2665.
233 *See* Control Nos. 95/000,285, 95/000,310, and 95/001,018.
234 *See* MPEP §§ 2265 and 2665.

responses to subsequent actions, the standard reverts to sufficient cause. Put differently, the term "second or subsequent requests" refers to a second or subsequent request in the context of the particular deadline for which an extension is sought, and not to second or subsequent requests over the course of the entire reexamination. This is important because "second or subsequent" requests are only granted under "extraordinary circumstances."

However, based on conversations with two CRU Supervisory Patent Examiners ("SPEs"), if the PTO perceives a pattern of delay, such as where an applicant requests an extension of time for every deadline, subsequent requests over the course of an entire reexamination receive closer and closer attention. While not reviewed under the "extraordinary circumstances" standard, these subsequent extension of time requests will get closer inspection and may require additional proof than would have been required for a first request.

I. Page Limits For **Inter Partes** *Reexamination Papers*

The *inter partes* reexamination rules impose strict page and word limits on the length of patent owner and third party requester submissions during reexamination at the CRU and on appeal to the BPAI. These limits can have a significant impact on a party's ability to effectively advocate its positions.

The rules impose a 50-page limit on Office action responses filed by the patent owner and written comments filed by the third party requester.[235] There is no corresponding limit set on the length of the reexamination request or the Office action issued by the examiner.[236] It is therefore not uncommon for the first Office action itself to exceed the 50-page limit.

The motivation behind the page limit restriction appears to be efficiency. The page limits are viewed by the PTO as forcing prosecuting attorneys to economize their arguments, thereby reducing the number of pages that an examiner must review in each Office action cycle. However, many argue that divorcing the page limit requirement from the size of the request and the extent of the Office action prejudices both the patent owner and the third party requester. In cases having lengthy and complex Office actions, these page limits may not afford the patent owner or third party requester the ability to argue adequately all the issues raised in the reexamination request or in an Office action.

The patent owner and the third party requester can petition OPLA for an enlargement of the page limits. However, the grant of an extension is not automatic. Many practitioners also argue that the parameters for when to request an extension and what grounds are sufficient to obtain an extension are unclear and inconsistent across reexamination proceedings.

The examiner makes the determination of whether a filing exceeds the 50-page limit. When expert declarations are used by either party, the examiner must determine whether a page in the declaration counts towards the 50-page limit. Practitioners argue that the standards for making this determination are unclear and are also not consistent across CRU examiners. A need exists for guidance from the PTO on this issue.

Limits also are applied to Appeal Briefs. On appeal to the BPAI, appellants are limited to thirty pages or 14,000 words, excluding appendices of claims and reference

235 37 C.F.R. § 1.943(b).

236 The trend in *inter partes* reexamination appears to be the filing of lengthy reexamination requests. In many cases, the reexamination request exceeds 500 pages. It is not uncommon for the request to exceed 250 pages.

materials such as prior art references. All subsequent briefs are further limited to 15 pages or 7,000 words. The MPEP waives the pages limits for the table of contents pages, the table of case law pages, and the pages of the claims (but not claim charts applying the art to the claims). OPLA has recently ruled that the exclusion of reference material from the word or page limits applies to the following components of briefs: Real Party in Interest, Related Proceeding Appendix and Certificate of Service.[237]

The differing limits on opening briefs in an appeal will, in some cases, hamper patent owner relative to requester. Where a patent owner is not cross-appealing, such as in the case of it having all claims confirmed in reexamination, the only brief a patent owner may file is the response brief, which is entitled to half of the length of the opening brief.

Complicating matters, the MPEP provides that expert opinion as to whether the claims are anticipated or rendered obvious in view of the prior art are counted in the page limits.[238] In addition, parties attempting to overcome the page limit hurdle by incorporating by reference more detailed argument do so at their own peril. The rules provide that "Any arguments or authorities not included in the [Appellant's] brief …will be refused consideration by the Board." 37 C.F.R. § 41.67(c)(1)(vii). In *Ex Parte Fleming*, the BPAI refused to consider arguments that appellant incorporated by reference. [239] In Fleming, the appellants sought to raise an argument on appeal that it raised during prosecution, stating: "[a] detailed presentation of this argument cab [sic] be found in Applicant's response dated May 8, 2007, and is not reproduced herein."[240] The BPAI responded: "we decline to consider this argument."[241]

How can the PTO balance the need for economy while protecting the right of the patent owner to protect its property right? Can the PTO *sua sponte* waive the page limit rule when an Office action exceeding a specific length is issued?

This issue is exacerbated by the need to create a full and complete factual record of the story of the invention and the relevant prior art to overcome obviousness rejections. In the patent interference context, an Evidence Appendix is used to create the factual record and is not counted as part of the page limits for attorney argument. The PTO in the interference context takes a liberal view of what is allowed in the Appendix that often runs thousands of pages in length. Many argue that the CRU should adopt a corresponding approach to the page limit in reexaminations.

J. Evidence Considerations

The use of declarations or affidavits in a reexamination proceeding is critical. Since the *KSR* decision, the number of obviousness-based challenges to patentability has increased. Expert declarations in support of validity or in support of invalidity of the claims undergoing reexamination are presented in declarations under 37 C.F.R. 1.132 (referred to herein as "Rule 132 declarations"). The need for Rule 132 declaration is especially important in defending against obviousness challenges. Declarations are often the only way to get expert evidence as well as evidence of secondary consideration before the examiner. To ensure declarations are accepted by the CRU and given persuasive weight, practitioners should be mindful of the critical aspects of declaration practice. PTO guidance provides that evidence

237 In re Farmwald, Inter Partes Reexamination Control No., 95/001,905 (May 27, 2010 Decision).
238 MPEP 2667.
239 2010 WL 964711 (Appeal 2009-5123, Mar. 15, 2010).
240 *Id.*
241 *Id.*

of fact is given the greatest weight, depending on its relevance. Opinion testimony is given less weight, depending on the underlying factual basis for the opinions. Allegations are not entitled to any weight. Factual evidence can include comparative test data to show unexpected results or inoperability of an applied reference, sales figures to show commercial success, and publications or declarations of skilled artisans to long-felt need. Because factual evidence is preferred by the PTO, is important to state a detailed factual basis for any opinion offered by a declarant. Opinions on the ultimate legal issue will not be given weight but other opinions supported by evidence will be given weight.[242] It is also critical that any evidence of secondary considerations be shown to be related to the claimed inventions. These considerations are also relevant to a party attacking another's declaration.

There is little or no guidance on how the CRU is to weigh competing evidence of equal weight, such as conflicting factual evidence or competing factually supported opinions. Reexaminations are decided on a cold record with no live testimony and no cross-examination. The CRU therefore cannot rely on things such as witness demeanor. Without explicit guidance, deciding between competing declaration appears to be a judgment call for the CRU.

VI. Central Reexamination Unit (CRU) Improves Quality and Reduces Pendency

A. Dedicated Examiners

In 2005, the PTO created a dedicated Central Reexamination Unit (CRU) to ensure quality and consistency of reexamination proceedings, and to reduce pendency. The CRU handles all reexaminations regardless of technology, and all legacy reexaminations have been transferred to the CRU from their respective technology centers. The CRU is currently headed by CRU Director, Mr. Greg Morse, who assumed his new duties on June 23, 2008.

In 2006, the CRU included 31 patent examiners, 3 supervisory patent examiners, and 10 support staff. When Mr. Morse took over in June 2008, the CRU ranks had grown to a total of 53 patent examiners with 10 in the biological/chemical group, 12 in the mechanical arts, and 31 in the electrical arts. There are now 58 examiners and six supervisors. A job announcement is outstanding seeking additional CRU examiners who must have been a primary examiner at the PTO and thus excludes people outside the PTO from joining the CRU.

The examiners in the CRU are selected from the ranks of the PTO examining corps and have an average of approximately 17 years of examining experience. Many of the CRU examiners also have advanced technical degrees and/or law degrees. The CRU examiners are evaluated on work quality and workflow, not on production quotas as is the case with the regular examining corps. Assignment to the CRU is recognized throughout the PTO as professional advancement; morale throughout the CRU is high, and turnover is low. The examiner position in the CRU is currently "temporary," but the recent announcement for job openings indicates that the position is "transitional." It remains to be seen whether the position will become "permanent" and thus subject to performance metrics as is the case with examiner positions in original prosecution.

---

242 *Id.,* MPEP 716.01(c).

Preparation of each Office action involves two patent examiners and one supervisory patent examiner. One of the patent examiners prepares the Office action. The second patent examiner and the supervisory patent examiner review the Office action before the Office action is mailed. The three examiners confer to resolve any disagreements.

B. Interaction Between CRU and OPLA

As noted above, the CRU works closely with the Office of Patent Legal Administration ("OPLA") to resolve petitions to the director, disputes involving the "real party in interest," petitions to merge proceedings, and the like. The CRU also works closely with the OPLA in examining the practical effects of its own rules and in examining ways to streamline reexamination proceedings to reflect today's realities.

At least one OPLA official has acknowledged that, while the present rules worked well in the past when there were relatively few reexaminations, they are "clunky" in view of today's realities. The CRU and OPLA are therefore currently looking to "streamline" some of the rules.

C. Practice Suggestions

Under its new practice, the CRU will only reexamine claims for which a SNQ is alleged and found to exist. In the past, a SNQ for even a single claim would typically trigger the reexamination of all claims.

Under the new rules, every submitted reference must be applied to at least one claim. Many requests are being denied a filing date for failing to apply every cited reference. The filing date is important because, as mentioned above, the PTO strives to issue a final office within two years of the filing date.

Litigation docket numbers are not to be cited in PTO disclosure forms (i.e., PTO Form 1449, PTO/SB/08A, or PTO/SB/08B). Affidavits and/or testimony transcripts can be cited in PTO disclosure forms.

Extensions of time in reexamination proceedings must be obtained in advance and will not be granted without a substantial reason. A request for an extension of time must include a description of relevant activities to date, reasons necessitating an extension of time, and relevant actions that will be taken during the requested extension period. Potentially valid reasons include the death or unavailability of an inventor or a need for trial testimony or exhibits. Extensions of time will generally not be granted for holidays or vacations.

If a practitioner has a concern about a reexamination, he or she should contact the Office of the CRU Director before petitioning the Patent Commissioner. A petition to the Commissioner triggers the transfer of the reexamination file from the CRU to the Commissioner, which may delay the reexamination proceeding.

D. Recommendations That Are Circulating

Substantial discussion exists among reexamination experts about ways the current reexamination process can be improved. While the topics discussed below are not meant to be exhaustive, the ones presented are "hot button" issues and deserve immediate and special attention.

1. Extensions of Time

In reexamination proceedings, deadlines for filing responses or third party papers are generally extremely short, particularly in cases involving a concurrent litigation. However, the PTO is cognizant that a patent owner must be provided with a fair opportunity to present an argument against any attack on the patent[243] and has granted requests for extension of time in a number of cases. If a patent owner is faced with an action presenting extensive and complex rejections or an action requiring development of factual affidavits, the patent owner should consider filing a request for an extension of time. Such a request should be detailed and provide sufficient cause to justify the extension. When deciding such requests, the PTO must balance the interest of the parties against the statutory mandate that reexaminations be conducted with special dispatch.

The *inter partes* reexamination rules allow for the patent owner to request an extension of time to reply to an Office action (or any action) for sufficient cause.[244] However, the rules explicitly prohibit the third party from receiving an extension of time for submitting written comments to the patent owner's Office action response.[245] Many practitioners argue that this rule is unfair to the third party requester, particularly in cases where the patent owner has received an increase in the page limit for reply or submits voluminous factual evidence. One recommendation being circulated is to remove the prohibition against the third party requester and allow the third party requester to request extensions of time using the same standards applied to the patent owner.

2. Page Limit Waivers

As discussed above, the *inter partes* reexamination rules impose a 50-page limit on Office action responses filed by the patent owner and written comments filed by the third party requester. One recommendation circulating is that the PTO remove the 50-page limit. Opponents argue that removal of any page limit would invite spurious and unfocused arguments, increasing the burden on the examiners. Another recommendation circulating is that the PTO *sua sponte* waive or increase the 50-page limit when issuing extensive Office actions.

Another factor that directly impacts page limits is whether the responder is presenting facts or argument. A simple presentation of facts does not count against the page limit, while attorney (or applicant) argument does. For this reason, the authors recommend that practitioners take a lesson from district court litigation and clearly separate the facts from the argument. This can be done in the body of the response by delineating factual sections from arguments, much the same way a summary judgment motion or opinion will have separate sections for "findings of fact" and "conclusions of law." Alternatively, practitioners should consider a separate Evidence Appendix. Both of these devices will assist the responder in separating facts surrounding the story of the invention and the prior technology from the arguments in favor (or against) patentability.

One complicating factor worth noting is that the line between "facts" and "argument" may be a blurry one. Discussion of teachings of a reference is factual. The ultimate conclusion of obviousness is legal. In between, however, could exist gray areas. Nonetheless, practitioners can only help themselves if clear distinctions are made in the body of any response.

---

243 *See* MPEP §§ 2265 and 2665.
244 37 C.F.R. § 1.956.
245 37 C.F.R. § 1.947.

3. Adopt an "Interference-type" Approach

Given the concern regarding the pendency of reexamination proceedings, particularly *inter partes* proceedings, some experts have argued that the reexamination process should be modeled after the interference process currently instituted by the PTO.

E. CRU Criticisms

Some practitioners have nicknamed the CRU as the "Central Rejection Unit." Is there any practical truth to this moniker? With the much greater resources devoted to the reexamination process by the CRU, coupled with the impact of *KSR* on how printed publications and patents will be treated by the CRU, and in *inter partes* reexaminations the presence of the third party requester throughout the process, speculation abounds that the chance of a claim surviving the reexamination process is becoming smaller and smaller as time goes on. Is this speculation correct in practice?

Some critics argue that the CRU is unfairly inclined to find claims invalid in reexaminations especially where these claims have been subjected to extensive prior litigation and have survived significant validity challenges in the courts. They contend that, because the CRU essentially does no searching but instead relies on the parties to do this task, the examiners do not appreciate the innovation captured by the claimed invention because they are "spoon fed" the prior art by the requester.

Moreover, critics contend that the CRU tends to "rubber stamp" the reexamination request filed by the requester in *inter partes* proceedings and that the first Office action almost always rejects all of the claims. Recently, however, the CRU has granted several requests for reexamination where the order did not adopt all SNQs and or proposed rejections. Some contend that any "rubber stamp" improperly shifts the presumption to the patent owner that the claims are *prima facie* invalid. They also contend that the patent owner essentially can only put forth evidence of nonobviousness in response to the first Office action and that trying "to prove a negative" is daunting at best, and a practical impossibility at worst. Critics argue that this approach subverts the constitutional basis for patents.

Finally, some critics contend that the PTO is pressuring the CRU to be negative towards claims in reexamination to perhaps mollify powerful political forces at work. It is argued that those political forces seek to weaken valuable patent rights owned by non-practicing entities, which happen to encompass entities such as universities and R&D innovation companies. Some perceive that these entities are being ganged up on and attacked unfairly in the courts. The employees of the PTO interviewed by the authors believe that this criticism is unfounded. These employees counter that any pressures from the PTO are limited to meeting deadlines and timelines.

F. Practitioner Criticisms

Some critics argue that practitioners are abusing the reexamination process by filing papers for intentional delay or by turning the petition process into a form of motion practice before OPLA.

The authors interviewed high level OPLA, CRU and BPAI officials as well as several judges. These individuals provided extensive criticisms and suggestions on how practitioners could improve in terms of dealing with the PTO and the courts in reexaminations. The authors have provided these criticisms and suggestions in the applicable sections of this paper rather than providing a long laundry list here. Suffice it to say, the bar needs to improve its practices and procedures in a manner commensurate with what the PTO and the courts need to do.

VII. REEXAMINATION STATISTICS

We conclude this paper with a discussion of reexamination statistics. The authors have done their own independent data gathering and analysis of reexaminations and offer the following summary. To provide further insight into reexamination practice, the authors have conducted a thorough review of all 768 reexamination requests noticed in the Official Gazette and filed between January 1, 2009, and January 1, 2010.

Review of these requests showed that 219 *inter partes* reexamination requests were filed involving 119 unique patent owners. The owners included 97 U.S.-based companies, four German-based companies, four Japanese-based companies, three Korean-based companies, and two Taiwanese-based companies with the remainder based in eight other countries.

The 550 *ex parte* reexamination requests filed involved 392 unique patent owners, The owners included 307 U.S.-based companies, 14 Japanese-based companies, 10 Canadian-based companies, nine Taiwanese-based companies, eight German-based companies, six Korean-based companies, five United Kingdom-based companies, and five Swiss-based companies with the remainder of the patent owners based in 16 other countries. Of the *ex parte* reexaminations only 11% (60) were requested by patent owners. The remaining 89% (490) were initiated by third party requesters.

The technologies involved were similarly diverse. The top technology classes representing 41% of all reexaminations included "Surgery" (49), "Drug, bio-affecting and body treating compositions" (42), "Telecommunications" (28), "Electrical computers and digital data processing systems: input/output" (26), "Chemistry: molecular biology and microbiology" (22), "Communications: electrical" (22), "Electrical computers and digital processing systems: multicomputer data transferring" (21), "Multiplex communications" (21), "Data processing: financial, business practice, management, or cost/price determination" aka business methods (18), "Image analysis" (18), "Television" (17), "Active solid-state devices (e.g., transistors, solid-state diodes)" (16), and "Data processing: presentation processing of document, operator interface processing, and screen saver display processing" (15) to name only a few.

Requests for reexamination were assigned to one of three art units ("AU") in the CRU supervised by Supervisory Primary Examiners ("SPEs"). Specifically, AU 3992 (specializing in the Electrical and Processing arts, staffed by 35 examiners including four SPEs) was assigned 296 *ex parte* and 153 *inter partes* proceedings, thus receiving over 58% of all reexaminations instituted by the CRU in the past year. AU 3993 (specializing in the Mechanical and Material arts, staffed by 16 examiners including a SPE) was assigned 150 *ex parte* and 32 *inter partes* reexaminations or 24% of the total instituted in 2009. Finally, AU 3991 (specializing in the Biotechnology and Pharmaceutical arts, staffed by 13 examiners including a SPE) was assigned 103 *ex parte* and 104 *inter partes* reexaminations. While only assigned 18%, this unit received the largest relative proportion of *inter partes* proceedings.

Attached are the most recent reexamination statistics from the PTO dated June 30, 2010. The most recent official reexamination filing data and statistics, including methods for searching by control number and links to historical and operational statistics, are made available to the public by the Central Reexamination Unit (CRU).[246]

246 http://www.uspto.gov/patents/stats/Reexamination_Information.jsp.

2010 CRU REEXAMINATION STATISTICS

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

Ex Parte Reexamination Filing Data - June 30, 2010

1. Total requests filed since start of *ex parte* reexam on 07/01/81 .... 10782[1]

a. By patent owner	3680	34%
b. By other member of public	6937	64%
c. By order of Commissioner	165	2%

2. Number of filings by discipline

a. Chemical Operation	3019	28%
b. Electrical Operation	3895	36%
c. Mechanical Operation	3697	34%
d. Design Patents	171	2%

3. Annual *Ex Parte* Reexam Filings

Fiscal Yr.	No.	Fiscal Yr.	No.	Fiscal Yr.	No.	Fiscal Yr.	No.
1981	78 (3 mos.)	1989	243	1997	376	2005	524
1982	187	1990	297	1998	350	2006	511
1983	186	1991	307	1999	385	2007	643
1984	189	1992	392	2000	318	2008	680
1985	230	1993	359	2001	296	2009	658
1986	232	1994	379	2002	272	2010YTD	539
1987	240	1995	392	2003	392		
1988	268	1996	418	2004	441		

4. Number known to be in litigation.. .3443...............32%

5. Decisions on requests...10296

a. No. granted... ...9467................92%

(1) By examiner	9353
(2) By Director (on petition)	114

b. No. denied ....829.................8%

(1) By examiner	793
(2) Reexam vacated	35

[1]Of the requests received in FY 2010, 16 requests have not yet been accorded a filing date, and preprocessing of 55 requests was terminated for failure to comply with the requirements of 37 CFR 1.510. See Clarification of Filing Date Requirements for *Ex Parte* and *Inter Partes* Reexamination Proceedings, Final Rule, 71 Fed. Reg. 44219 (August 4, 2006).

6. Total examiner denials (includes denials reversed by Director)....................................907

a. Patent owner requester	451	50%
b. Third party requester	456	50%

7. Overall reexamination pendency (Filing date to certificate issue date)

a. Average pendency	25.4 (mos.)
b. Median pendency	20.0 (mos.)

8. Reexam certificate claim analysis:

	Owner Requester	3rd Party Requester	Comm'r Initiated	Overall
a. All claims confirmed	22%	25%	12%	23%
b. All claims cancelled	8%	13%	23%	12%
c. Claims changes	70%	62%	65%	65%

9. Total *ex parte* reexamination certificates issued (1981 – present) 7586

a. Certificates with all claims confirmed	1786	23%
b. Certificates with all claims canceled	877	12%
c. Certificates with claims changes	4923	65%

10. Reexam claim analysis – requester is patent owner or 3rd party or Commissioner initiated.

a. Certificates – PATENT OWNER REQUESTER 2910

(1) All claims confirmed	634	22%
(2) All claims canceled	250	8%
(3) Claim changes	2026	70%

b. Certificates – 3rd PARTY REQUESTER 4524

(1) All claims confirmed	1134	25%
(2) All claims canceled	592	13%
(3) Claim changes	2798	62%

c. Certificates – COMMISSIONER INITIATED REEXAM 152

(1) All claims confirmed	18	12%
(2) All claims canceled	35	23%
(3) Claim changes	99	65%

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

Inter Partes Reexamination Filing Data – June 30, 2010

1. Total requests filed since start of *inter partes* reexam on 11/29/99 .... 923[1]

2. Number of filings by discipline

a. Chemical Operation	179	20%
b. Electrical Operation	443	48%
c. Mechanical Operation	289	31%
d. Design Patents	12	1%

3. Annual Reexam Filings

Fiscal Yr.	No.	Fiscal Yr.	No.	Fiscal Yr.	No.	Fiscal Yr.	No.
2000	0	2003	21	2006	70	2009	258
2001	1	2004	27	2007	126	2010YTD	189
2002	4	2005	59	2008	168		

4. Number known to be in litigation .... 638 .... 69%

5. Decisions on requests 816

   a. No. granted .... 781 .... 96%

      (1) By examiner 779
      (2) By Director (on petition) 2

   b. No. not granted .... 35 .... 4%

      (1) By examiner 31
      (2) Reexam vacated 4

6. Overall reexamination pendency (Filing date to certificate issue date)
   a. Average pendency 35.9 (mos.)
   b. Median pendency 31.4 (mos.)

7. Total inter partes reexamination certificates issued (1999 - present) 167

a. Certificates with all claims confirmed	14	8%
b. Certificates with all claims canceled (or disclaimed)	82	49%
c. Certificates with claims changes	71	43%

[1] Of the requests received in FY 2010, 6 requests have not yet been accorded a filing date, and 19 requests have had preprocessing terminated, for failure to comply with the requirements of 37 CFR 1.915. See Clarification of Filing Date Requirements for *Ex Parte* and *Inter Partes* Reexamination Proceedings, Final Rule, 71 Fed. Reg. 44219 (August 4, 2006).

UNCHARTED WATERS: DETERMINING ONGOING ROYALTIES FOR VICTORIOUS PATENT HOLDERS DENIED AN INJUNCTION

*Ronald J. Schutz & Patrick M. Arenz**
Robins, Kaplan, Miller & Ciresi L.L.P.
Minneapolis, MN

INTRODUCTION

The Supreme Court's *eBay* decision, of course, caused a sea change. Before *eBay*, courts awarded victorious patent holders an injunction preventing the adjudged infringer from future infringing acts unless the case was "exceptional." Since *eBay*, courts reluctantly grant injunctions for at least victorious non-practicing entity patent holders. The Federal Circuit in *Paice v. Toyota Motor Corp.*, moreover, determined that district courts may issue an ongoing royalty for future infringing acts, if a victorious patent holder is denied an injunction. But the *Paice* Court did not specify what factors should be considered in determining an ongoing royalty. Nor was the issue of who—the district court or the jury—properly at issue on appeal. A discussion of this decision is set forth below.

In light of *eBay* and *Paice*, therefore, patent holders who are denied an injunction face uncertain waters for future compensation. This paper addresses those patent holders' options for proper, adequate, and fair compensation after a jury finds infringement of a valid patent, and an infringer willingly decides to continue infringing. In particular, the paper addresses, and ultimately concludes, that an ongoing royalty must be decided by a jury, as required by the Seventh Amendment, and that the jury should decide the ongoing royalty by assessing what the infringer would pay to avoid an injunction, which should in most cases be higher than the royalty found for past infringement as a result of the change in the parties' legal relationship.

---

* Ronald J. Schutz is the Chair of the Intellectual Property Litigation Group and a member of the Executive Board of Robins, Kaplan, Miller & Ciresi L.L.P. He is a Fellow of the American College of Trial Lawyers and is listed in *The Best Lawyers* in America. In 2008 the *National Law Journal* named Mr. Schutz to its annual list of the Top 10 Winning Litigators in the United States. He can be reached at RJSchutz@rkmc.com. Patrick M. Arenz is an associate at Robins, Kaplan, Miller & Ciresi L.L.P. He focuses his practice on intellectual property litigation and can be reached at PMArenz@rkmc.com.

The content of this article should not be taken as legal advice or as an expression of the views of the firm, its attorneys, or any of its clients. We hope the article spurs discussion in the legal community with insight into the experience of the authors.

I. The Federal Circuit's *Paice v. Toyota Motor Corp.* Decision Permits Ongoing Licenses.

The seminal case involving—and permitting—ongoing royalties is *Paice v. Toyota Motor Corp.*[1] The patented technology in *Paice* dealt with an automobile drive train that utilized a microprocessor and a controllable torque transfer unit, so that wheels in hybrid vehicles use torque from an internal combustion engine, an electric motor, or a combination of both.[2] The accused products were Toyota's Prius hybrid cars.[3]

At the district court, venued in the Eastern District of Texas in front of Judge David J. Folsom, a jury concluded that Toyota infringed one of Paice's patents.[4] The jury awarded Paice a reasonable royalty of $4,269,950.00.[5] This amounted to a royalty of $25 per infringing car.[6] During post-trial motions, the court rejected Paice's motion for a permanent injunction.[7] The court determined that Paice—a non-practicing entity—would not suffer irreparable injury and that monetary damages were adequate.[8] In conjunction with this denial, the court imposed an ongoing royalty of $25 per infringing vehicle for the remaining life of the patent.[9] Both parties appealed this decision.

On appeal, the Federal Circuit rejected Toyota's appeal and Paice's cross-appeal on infringement issues.[10] Of most relevance, however, the Federal Circuit examined Paice's two other cross-appeal issues. First, Paice challenged the district court's statutory authority to issue an ongoing royalty.[11] Second, Paice alleged, even if a court could issue an ongoing royalty, the district court denied Paice its right to a jury determining the amount of the ongoing royalty rate.[12]

With respect to the first issue, the Federal Circuit started its analysis with 35 U.S.C. §283.[13]

> The several courts having jurisdiction of cases under this title may grant injunctions in accordance with the principles of equity to prevent the violation of any right secured by patent, on such terms as the court deems reasonable.

The court reviewed this statute to determine "whether an order *permitting* use of a patented invention in exchange for a royalty is properly characterized as *preventing* the violation of the rights secured by the patent."[14] Without much analysis, the Federal Circuit cited an earlier Federal Circuit case, *Shatterproof Glass Corp. v. Libbey-Owens Ford Corp.*,[15] which allowed a "compulsory license,"[16] noting that neither party objected to the district

---

1 504 F.3d 1293 (Fed. Cir. 2007), *reh'g en banc denied*, [hereinafter *Paice I*]. The panel included Circuit Judges Lourie, Rader, and Prost. Judge Prost authored the opinion. Judge Rader filed a concurrence.
2 *See id.* at 1296.
3 *Id.* at 1299.
4 *Id.* at 1302.
5 *Id.*
6 *Id.* at 1303.
7 *Id.*
8 *See Paice LLC, v. Toyota Motor Corp.*, No. 2:04-cv-211, 2006 WL 2385139, at *5-6 (E.D. Tex. Aug. 16, 2006).
9 *Paice I*, 504 F.3d at 1303.
10 *See id.* at 1312-13.
11 *Id.* at 1314.
12 *Id.*
13 *Id.*
14 *Id.*
15 758 F.2d 613, 628 (Fed. Cir. 1985).
16 Notably, the court used "ongoing royalty" and distinguished it from "compulsory royalty." *Id.* at 1313 n.13. "We use the term ongoing royalty to distinguish this equitable remedy from a compulsory license. The term "compulsory license" implies that anyone who meets certain criteria has congressional authority to use that which is licensed." *Id.* Judger Rader, nonetheless, recognized less of a difference. "Nonetheless, calling a compulsory license an 'ongoing royalty' does not make it any less a compulsory license." *Id.* at 1316 (Rader, J., concurring).

court's authority to impose that remedy.[17] The court also relied on antitrust law. "In the context of an antitrust violation, 'mandatory sales and reasonable-royalty licensing' of relevant patents are 'well-established forms of relief when necessary to an effective remedy, particularly where patent have provided the leverage for or have contributed to the antitrust violation adjudicated.'"[18]

The Federal Circuit, therefore, concluded that an ongoing royalty may be appropriate when an injunction is denied.[19] But the court cautioned that, in most cases, when a district court denies a permanent injunction, the court may wish to allow the parties first the opportunity to negotiate a license.[20] Only if those negotiations fail should the court step in to "assess a reasonable royalty in light of the ongoing infringement."[21]

Based on the record on appeal, the Federal Circuit remanded the case to the district court because the district court's ongoing royalty order provided no reasoning to support the $25 rate.[22] On remand, the Federal Circuit required the district court to "take additional evidence if necessary to account for any additional economic factors arising out of the imposition of an ongoing royalty."[23] Judge Rader, in his concurrence, further elaborated that "pre-suit and post-judgment acts of infringement are distinct, and may warrant different royalty rates given the change in the parties' legal relationship and other factors."[24]

The Federal Circuit then addressed Paice's argument that it is entitled to have a jury determine the ongoing royalty. In this respect, the court presented Paice's argument as "merely stat[ing] that '[i]t is well settled that the determination of damages is a legal question which carries a Seventh Amendment right to a jury trial.'"[25] While the court acquiesced with Paice as a general matter, the Federal Circuit cited case law that money damages does not always carry a jury trial right, but did not analyze ongoing royalties specifically.[26] Therefore, the court concluded that "the fact that monetary relief is at issue in this case does not, standing alone, warrant a jury trial."[27] Quite simply, little analysis went into the argument, which required little analysis to decide.

II. How and Who Determines an Ongoing License.

The *Paice* decision leaves open more questions than it answers. Specifically, the law is not clear who—the court or the jury—determines the ongoing royalty, or how the fact-finder determines an appropriate ongoing license. Each of these issues is discussed below.

---

17 *Paice*, 504 F.3d at 1314.
18 *Id.* (citing *United States v. Glaxo Group, Ltd.*, 410 U.S. 52, 59 (1973)).
19 *Id.* at 1314-15.
20 *Id.* at 1315.
21 *Id. See also id.* at 1316-17 (Rader, J., concurring) (opining that "this court should require the district court to remand this issue to the parties, or to obtain the permission of both parties before setting the ongoing royalty rate itself").
22 *Id.* at 1315 ("The district court may determine that $25 is, in fact, an appropriate royalty rate going forward. However, without any indication as to why that rate is appropriate, we are unable to determine whether the district court abused its discretion.").
23 *Id.*
24 *Id.* at 1316 (Rader, J., concurring).
25 *Id.* at 1316.
26 *Id.* at 1316. The court cited *Root v. M.S. Ry. Co.*, 105 U.S. 189, 207 (1881), for the proposition that "[w]hen, … relief was sought which equity alone could give … in order to avoid multiplicity of suits and to do complete justice, the court assumed jurisdiction to award compensation for the past injury, not, however, by assessing damages, which was the peculiar office of the jury, but requiring an account of profits….", and *Bowen v. Massachusetts*, 487 U.S. 879, 910 (1988) explaining "even if the District Court's orders are construed in part as orders for the payment of money by the Federal Government to the State, such payments are not 'money damages' …. That is, since the orders are for specific relief (they undo the Secretary's refusal to reimburse the State) rather than for money damages (they do not provide relief that substitutes for what which ought to have been done) they are within the District Court's jurisdiction….").
27 *Id.* at 1316.

A. Who Decides an Ongoing Royalty—the Court or the Jury?

The issue of whether a patent holder has a Seventh Amendment right to an ongoing royalty determination was ostensibly in front of the court in *Paice*. However, as explained above, the Federal Circuit made clear that *Paice* had not sufficiently raised the issue. The following discussion provides a fuller analysis of whether a jury trial right should attach to an ongoing royalty determination, and then how courts have decided ongoing royalties without a jury trial requirement.

1. A Patent Holder Is Entitled to Have a Jury Determine an Ongoing Royalty.

The United States Constitution and Supreme Court precedent establishes a right to jury trial for patent holders denied an injunction and who pursue an ongoing royalty. The Seventh Amendment provides that "[i]n suits at common law, where the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved…."[28] This right is not limited to common law causes of action, however, but also includes actions created by Congress.[29] The right to a jury trial, of course, is one of the most sacred in our democracy. "Trial by a jury of laymen rather than by the sovereign's judges was important to the founders because juries represent the layman's common sense, the 'passional elements in our nature,' and thus keep the administration of law in accord with the wishes and feelings of the community."[30] Indeed, the right to a jury trial was "one of the only rights unanimously included in the original state constitutions."[31]

The United States Supreme Court applies a two-prong test to determine if a jury right attaches to a cause of action.[32] Under the first prong, "a court must compare the action with the analogous action brought in the courts of England during the eighteenth century, prior to the merger of law and equity. Second, a court must look to the remedy sought and determine whether it is legal or equitable in nature."[33] Application of this "historical test"[34] compels the conclusion that a patent holder has a Seventh Amendment right to trial by jury for the issue of ongoing royalties.

a. <u>Patent infringement cases for damages have always been tried in courts of law</u>.

Patent holders have always had the option to sue for patent infringement in courts of law. In fact, "[i]n 1790, before the adoption of the Seventh Amendment in 1791, the patent statute in this country provided only for the award of damages in an action at law, with the right to a trial by jury."[35] Indeed, as the United States Supreme Court explained in

28 U.S. CONST. amend. VII.
29 *Tull v. United States*, 481 U.S. 412, 417 (1987) (citing *Curtis v. Loether*, 415 U.S. 189, 193 (1974)).
30 *Parklane Hosiery Co. v. Shore*, 439 U.S. 322, 343 (1979) (Rehnquist, J., dissenting). *See also Jacob v. New York*, 315 U.S. 752, 752-53 (1942) ("The right of jury trial in civil cases at common law is a basic and fundamental feature of our system of federal jurisprudence which is protected by the Seventh Amendment. A right so fundamental and sacred to the citizen, whether guaranteed by the Constitution or provided by statute, should be jealously guarded by the courts.").
31 *Parklane*, 439 U.S. at 341 (Rehnquist, J., dissenting).
32 *See Tull*, 481 U.S. at 417.
33 *See, e.g., In re Tech. Lic. Corp.*, 423 F.3d 1286, 1287 (Fed. Cir. 2005).
34 *See, e.g., Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 376 (1996).
35 *In re Lockwood*, 50 F.3d 966, 976 n.12 (Fed. Cir. 1995) (citing *Root,* 105 U.S. 189, 191-92; DONALD S. CHISUM *in* 5 CHISUM ON PATENTS § 20.02 (1995)). Even though the Supreme Court vacated the order in *Lockwood* when, after the Court granted certiorari, Lockwood withdrew his request for a jury trial, the Federal Circuit has continually relied on the "relevant and detailed analysis" in *Lockwood*, which contains the Federal Circuit's most extensive discussion of the historical and legal framework for assessing jury trial rights in connection with patent infringement cases. *See, e.g., In re Tech.*, 423 F.3d at 1289 n.1 (citing *Tegal Corp. v. Tokyo Electron America, Inc.*, 257 F.3d 1331, 1340 (Fed. Cir. 2001)). As a result, the Federal Circuit has explained that *Lockwood* "has been neither supplanted nor questioned" and continues to be "pertinent." *Tegal,* 257 F.3d at 1340.

Markman, "[t]here is no dispute that infringement cases today must be tried to a jury, as their predecessors were more than two centuries ago."[36]

In particular, at common law, the patent holder had the option of selecting a court of law or equity, depending on the remedy sought. A patent holder seeking only damages brought an action at law, whereas a patent holder pursuing an injunction brought suit in equity.[37] For patent cases, therefore, the two-prong historical test melds into a one-prong test focusing on the remedy the patent holder seeks.[38] With respect to ongoing royalties, the first prong is satisfied because that is an action for damages, which is analogous to actions at common law brought in courts of law.

b. An ongoing royalty for future acts of infringement is a legal remedy.

An ongoing royalty as compensation for future infringing acts is a legal remedy. Quite simply, "[b]y the law the jury are judges of the damages."[39] Indeed, the Supreme Court has held that almost all forms of money damages are legal remedies for courts of law. Most generally, the Court has explained that money damages, which serve as compensation or punishment are legal remedies.[40] Therefore, statutory damages for copyright actions, punitive damages, treble damages, *inter alia*, are all considered legal remedies.[41] There is even historical evidence of juries determining discretionary monetary relief.[42] Thus, as a "general rule[,] … monetary relief is legal."[43]

But the *Paice* Court correctly noted that not "any award of monetarily relief must necessarily be 'legal' relief."[44] The Supreme Court has identified two exceptions to the "general rule" that monetary relief is legal. But neither apply to an ongoing royalty in patent cases. First, the Court has "characterized damages as equitable where they are restitutionary, such as in 'action[s] for disgorgement of improper profits."[45] As discussed below, however, an ongoing royalty is compensatory, not restitutionary. Second, "a monetary award 'incidental to or intertwined with injunctive relief' may be equitable."[46] Ongoing royalties do not fall within this exception because patent holders may now receive an ongoing royalty when they are in fact denied equitable relief. An ongoing royalty is not, therefore, intertwined with, but rather a replacement of an equitable remedy.

Ongoing royalties are properly characterized as a legal remedy because they are compensation for future acts of infringement. Compensation, as a purpose for money damages, is "traditionally associated with legal relief."[47] On remand, for instance, the district court in *Paice* explained that "an ongoing royalty rate must still adequately *compensate* a patentee for giving up his right under the law to exclude others…."[48] This is the same

36 *Markman*, 517 U.S. at 377.
37 *See, e.g., Lockwood*, 50 F.3d at 976. *In re Tech.* at 1288.
38 *See, e.g., Medimmune, Inc. v. Genentech, Inc.*, 535 F. Supp. 2d 1020, 1026 (C.D. Cal. 2008) (explaining that the *Lockwood* Court did not need to perform a second prong analysis because the remedy sought controlled the first prong, concluding a jury right attached).
39 *See, e.g., Feltner, Jr. v. Columbia Pictures Television, Inc.*, 523 U.S. 340, 352 (1998).
40 *See id.*
41 *See id.* at 353 (citing *Curtis v. Loether*, 415 U.S. 189, 196 (1974); *Tull v. United States*, 481 U.S. 412, 422 (1987)).
42 *See id.* (citing *Coryell v. Colbaugh*, 1 N.L.J. 77 (1791).
43 *See, e.g., id.* at 352 (1998).
44 *Curtis*, 415 U.S. at 196. *See also Paice*, 504 F.3d at 1316 (noting that "not all monetary relief is properly characterized as 'damages'.").
45 *Chauffeurs, Teamsters & Helpers, Local No. 391 v. Terry*, 494 U.S. 558, 570 (1990) (explaining that "restitution is limited to 'restoring the *status quo* and ordering the return of that which rightfully belongs to the purchaser or tenant.'").
46 *Id.* at 571.
47 *See, e.g., Feltner*, 523 U.S. 353; *Curtis*, 415 U.S. at 196 (concluding that compensatory damages was a traditional form of relief in courts of law).
48 *Paice, LLC v. Toyota Motor Corp.*, 609 F. Supp. 2d 620, 630 (E.D. Tex. 2009) (emphasis added).

purpose as a royalty for past infringement. Title 35 of the United States Code, § 284 provides a victorious patent holder "damages adequate to *compensate* for the infringement."[49] Compensation, as the purpose for an ongoing royalty, is further underscored by courts now issuing ongoing royalties when patent holders fail to prove, *inter alia*, that "remedies available at law, such as monetary damages, are adequate to compensate."[50] Thus, an ongoing royalty is a recognition that money damages compensate the patent holder for the infringer's future acts, and even that the ongoing royalty is a legal remedy. Therefore, a compensatory ongoing royalty should be considered legal relief for a Seventh Amendment analysis. The second prong of the historical test is then met and a jury trial right should attach to an ongoing royalty determination.

2. Courts Have Adopted Various Procedures for Addressing Ongoing Royalties Under the Current Law.

Despite this view that a jury trial right should attach to an ongoing royalty determination, under the current law there is no explicit recognition of this right yet. District courts, therefore, have generally adopted three different procedures to determine an ongoing royalty.[51] First, since *Paice*, district courts have most often decided, or suggested they will decide, the ongoing royalty themselves. That is what Judge Folsom did in *Paice* on remand.[52] The Northern District of California followed suit in *Boston Scientific, Corp. v. Johnson & Johnson*.[53] In that case, the district court held a two-hour evidentiary hearing, which consisted of each party putting on one damages expert.[54] Other courts have indicated they will undertake a hearing, but only after the parties first attempt to negotiate an ongoing royalty, as the Federal Circuit suggested in *Paice*.[55]

The second procedure used is to submit the issue to the jury during the underlying trial. Even if the court ultimately decides the ongoing royalty rate, the court may submit the issue to a jury. Courts, for instance, often do this for inequitable conduct determinations.[56] Judge Ron Clark, sitting in the Eastern District of Texas, has established a practice of submitting the ongoing royalty issue to the jury during the trial.[57] Specifically, Judge Clark recommends the following jury instruction: "What rate or sum of money, if any, do you find is adequate as a reasonable royalty to compensate Plaintiff for the conduct you found to infringe that occurs in the future? Answer in a percentage or in dollars and cents."[58] Judge Clark justifies this practice because "time and expense can be saved by having the damages experts testify once, rather than hold a separate mini-trial on the issue of future damages post-verdict. This procedure would encourage the experts

49 35 U.S.C. § 284 (emphasis added).
50 *See generally eBay, Inc. v. Mercexchange, LLC*, 547 U.S. 388, 390 (2006) (requiring patent holder to prove money damages are inadequate). *See also Mercexchange, LLC v. eBay, Inc.*, 500 F. Supp. 2d 556, 582 (E.D. Va. 2007) (denying injunction, *inter alia*, because money damages compensate patent holder); *Paice LLC v. Toyota Motor Corp.*, No. 2:04-CV-211, 2006 WL 2385139, at * (E.D. Tex. Aug. 16, 2006) (denying injunction, *inter alia*, because patent holder's previous post-trial offer to license technology to defendant "further demonstrates the adequacy of monetary relief").
51 *See Creative Internet Adver., Corp., v. Yahoo!, Inc.*, 674 F. Supp. 2d 847, 851 (E.D. Tex. 2009) ("To this end, the approaches used to calculate of [sic] post-verdict royalties has not been uniform amongst various courts.").
52 *Paice, LLC, v. Toyota Motor Corp.*, 609 F. Supp. 2d 620, 622 (E.D. Tex. 2009) (establishing ongoing royalty rate after holding hearing and considering the parties' submitted papers, testimony, evidence, and case law) [hereinafter *Paice II*].
53 No. C 02-00790, 2009 WL 975424, at *2 (N.D. Cal. Apr. 9, 2009).
54 *See id.*
55 *See Hynix Semiconductor, Inc. v. Rambus, Inc.*, 609 F. Supp. 2d 951, 987 (N.D. Cal. 2009); *Telcordia Techs., Inc. v. Cisco Sys., Inc.*, 592 F. Supp. 2d 727, 748 (D. Del. 2009) (declining patent holder's immediate request for an ongoing royalty, and instead ordering "the parties to negotiate the terms of a reasonable royalty rate going forward. Should the parties fail to reach an agreement, the court will permit the filing of competing proposals").
56 *See, e.g., Qualcomm Incorporated v. Broadcom Corp.*, 548 F.3d 1004, 1009 (Fed. Cir. 2008) (explaining that inequitable conduct was submitted to the jury for an advisory opinion).
57 *See, e.g.*, Lynne Marek, *Juries May Take Up Future Damages in Patent Cases*, THE NATIONAL LAW JOURNAL, Vol. 30, No. 47 (Aug. 4, 2008).
58 *See, e.g.*, *Cummins-Allison, Corp. v. SBM Co., Ltd.*, 584 F. Supp. 2d 916, 921 n.6 (E.D. Tex. 2008).

to keep their testimony about past and future damages logically consistent, and to give reasons for any differences."[59] This approach, however, arguably takes away the ability of the parties to negotiate a license themselves in light of the jury's verdict, as the Federal Circuit recommended.[60]

A third and final approach is for the court to decline to consider and determine an ongoing royalty. Instead, as Judge T. John Ward of the Eastern District of Texas did in *Medtronic Vascular, Inc. v. Boston Scientific, Corp.*, the court may sever a patent holder's continuing causes of action for future royalties.[61] In *Medtronic*, thus, Judge Ward required the patent holder to file a new complaint and proceed with a second action.[62]

B. How an Ongoing Royalty Is Decided.

The Federal Circuit has not provided any specific guidance as to how an ongoing royalty should be decided. In *Paice*, for instance, the Federal Circuit remanded "account for any additional economic factors arising out of the imposition of an ongoing royalty."[63] In light of this uncertainty, courts and commentators have undertaken and proposed different approaches. For instance, some courts simply apply the *Georgia-Pacific* factors to determine an ongoing royalty.[64] One law review commentary suggests the following five-factor test: (1) the market for the patent has, or is substantially certain to, expand or contract; (2) the likelihood that the infringer would explore other markets; (3) the availability and cost of noninfringing alternatives; (4) the capacity of the infringer to design around the patented technology, and the cost at which the infringer could do so; and (5) the expert testimony of qualified experts.[65] Another suggestion is to treat the patent like a stock option, or to recalculate the royalty each year based on the value of the patent during that year.[66] This, of course, is impractical and inefficient for the courts and the parties. Yet another option is to simply do nothing, and allow the patent holder to file additional lawsuits alleging willful infringement.[67]

Despite these diverse approaches, the analysis performed by the district court in *Paice* on remand most fairly compensates a patent holder for losing its right to exclude the infringer from practicing its invention. The court's analysis on remand is set forth below, followed by a recommendation for supplementing that analysis in future cases.

1. The district court's analysis in *Paice* on remand.

On remand in *Paice*, the district court set forth and analyzed the following "modified *Georgia-Pacific* framework" to determine an ongoing royalty: "what amount of money would reasonably compensate a patentee for giving up his right to exclude yet allow an ongoing willful infringer to make a reasonable profit?"[68] The court recognized that "the law must ensure that an adjudged infringer who voluntarily chooses to continue his

---

59 *Ariba, Inc. v. Emptoris, Inc.*, 567 F. Supp. 2d 914, 916 (E.D. Tex. 2008).
60 *But see Cummins-Allison, Corp. v. SBM Co., Ltd.*, 669 F. Supp. 2d 774, 779 (E.D. Tex. 2009) (explaining that "if the parties are given an opportunity to confer on the issue of future damages post-verdict and before final judgment, the court's submission of this question to the jury would not run afoul of the Federal Circuit's statement in *Paice*" with respect to negotiations about an ongoing royalty).
61 No. 2-06-CV-78, 2009 WL 175696, at *1 (E.D. Tex. Jan. 23, 2009).
62 *Id.*
63 *Paice*, 504 F.3d at 1315.
64 *See Boston Scientific*, 2009 WL 975424, at *5 (citing *Orion IP, LLC v. Mercedes-Benz USA, LLC*, No. CV 05-322, at *9 (E.D. Tex. Mar. 28, 2008)).
65 Stephen M. Ullmer, *Paice Yourselves: A Basic Framework for Ongoing Royalty Determinations in Patent Law*, 24 BERKELEY TECH. L.J. 75, 97 (2009).
66 *See* Tim Carlton, The *Ongoing Royalty: What Remedy Should a Patent Holder Receive When a Permanent Injunction Is Denied*? 43 GA. L. REV. 543, 572 (2009).
67 *See* Bernard H. Chao, *After eBay, Inc. v. Mercexchange: The Changing Landscape for Patent Remedies*, 9 MINN. J.L. SCI. & TECH. 543, 565 (2008).
68 *Paice II*, 609 F. Supp. 2d at 624.

infringing behavior must adequately compensate the patent holder for using the patent holder's property. Anything less would be manifestly unjust and violate the spirit, if not the letter, of the U.S. Constitution and the Patent Act."[69] Furthermore, the court explained that an ongoing royalty rate significantly impacts the patent holder's ability "to license its technology and effectively precludes an exclusive licensing agreement."[70]

Applying this framework to a supplemented evidentiary record, the court raised the royalty from $25 per car to $98. Specifically, the court found the following additional economic factors increased the royalty: (1) the change in legal relationship between the parties, as Toyota was an adjudged infringer; (2) the increased cost of oil since the jury verdict, (3) changes in federal fuel efficiency laws, (4) Toyota's dominance in the hybrid industry and its increased reputation as a "green" company, and (5) Toyota has introduced more hybrid cars. The court agreed with Toyota that the availability of non-infringing alternatives is a consideration, but rejected Toyota's argument that the cost of switching to an alternative design should not be considered.[71] Furthermore, the court gave little and perhaps no weight to licenses Toyota argued should lower the royalty. The court explained that neither of licenses accounted for the change in legal relationship between the parties, as neither involved an adjudged infringer.[72]

The court, therefore, started with Paice's expert's application of the 25 percent Rule of Thumb to Toyota's profit margin of nine percent, which resulted in a royalty of 2.25 percent. The court then lowered the royalty to 1.5 percent taking into account the modest earlier damages award and due to the fact that Toyota makes less profit on hybrid than non-hybrid cars. This resulted in a $98 per car royalty.[73]

2. A recommendation for supplementing the *Paice* analysis.

A victorious patent holder denied a permanent injunction must be compensated for losing its right to exclude the infringer from practicing its invention and allowing the infringer to willfully infringe the patent going forward. The *Paice* district court's modified *Georgia-Pacific* framework[74] generally accomplishes this, as long as it includes two concepts discussed below. These concepts contextualize the fact-finder's appreciation of the patent holder's right to exclude, and the willfulness of the infringer's actions.

First, expert witnesses should be able to consider, and counsel should be able to argue, that the jury—see section II(A)(I)—may consider what the infringer would pay to avoid a future injunction. Whether or not a patent holder may later obtain an injunction under *eBay* is irrelevant. The relevance is providing a context for what the right to exclude means. The fundamental nature of a right to exclude is to prevent another from infringing on a right unless on terms agreeable to the entity possessing the right. If the jury is unable to consider what an infringer would pay to avoid a future injunction, then the patent holder's statutory right to exclude[75] is completely eviscerated. Furthermore, this context

69 *Id.* at 630 (citing U.S. CONST. art. I, §8).
70 *Id.*
71 *Id.* at 627.
72 *Id.* at 628.
73 *Id.* at 630. *See also Creative Internet Adver.*, 674 F. Supp. 2d at 860-62 (increasing jury's 20% royalty for past damages to a 23% ongoing royalty).
74 *Id.* at 624 ("what amount of money would reasonably compensate a patentee for giving up his right to exclude yet allow an ongoing willful infringer to make a reasonable profit?").
75 The Court in *eBay* determined the Patent Act does not provide a remedy of automatic injunctive relief, but it did recognize a patent holder's statutory right to exclude others from practicing its invention. "To be sure, the Patent Act also declares that 'patents shall have the attributes of personal property,' §263, including 'the right to exclude others from making, using, offering for sale, or selling the invention,' §154(a)(1)." *eBay Inc. v. Mercexchange, LLC*, 547 U.S. 388, 392 (2006). *See also* U.S. CONST. art. I, § 8 ("Congress shall have the power to … promote the progress of science and useful arts, by securing for limited times to authors and inventors the *exclusive* right to their respective writing and discoveries") (emphasis added). "But the creation of a right is distinct from the provision of remedies for violations of that right." *Id.*

properly frames the issue for the jury. Just as a reasonable royalty for past infringement is based on a hypothetical negotiation for a license at the time of infringement, the ongoing royalty is a hypothetical negotiation to prevent an injunction of an adjudged infringer. Juries will understand this.

Second, as a general matter, the ongoing royalty rate should be higher than it was for the past infringement. This is a result of the changed legal relationship between the patent holder and the adjudged infringer from the jury verdict. The Federal Circuit recognized this fact in *Amado v. Microsoft Corp.*, explaining that "[t]here is a fundamental difference [] between a reasonable royalty for pre-verdict infringement and damages for post-verdict infringement. Prior to judgment, liability for infringement, as well as the validity of the patent, is uncertain, and damages are determined in the context of that uncertainty. Once a judgment of validity and infringement has been entered, however, the calculus is markedly different because different economic factors are involved."[76] District courts since *Paice* and *Amado* have reached the same conclusion.[77]

In light of this changed legal relationship, the patent holder will generally have an increased bargaining position after a jury finds infringement of a valid patent that was not present at the time the hypothetical negotiation began on the eve of infringement. This should generally result in a larger ongoing royalty than the royalty for past infringement. Nonetheless, there could be circumstances where, due to economics and equity, the ongoing royalty may not be as much as or more than the past royalty. The exact contours of any framework for determining an ongoing royalty, however, is best left to be decided by the experience of juries deciding, and courts overseeing, ongoing royalties on a case-by-case basis.

CONCLUSION

In recent years, patent holders have been under attack. Patent "reform" is routinely discussed in Congress and "patent troll" is now part of the vernacular. The *eBay* decision, more significantly, rejected the notion that the Patent Act's "right to exclude" provides an injunction as a remedy. As a result, and following the *Paice* decision, many patent holder's "right to exclude" now means "right to receive an ongoing royalty if an adjudged infringer unilaterally decides to continue making, using, or selling the invention." Fortunately, open issues remain as to how and who determines the ongoing royalty. In order to stop any further erosion of patent rights, the ongoing royalty should be decided by a jury, and the jury should determine the ongoing royalty by considering what an infringer would pay to avoid an injunction.

76 517 F.3d 1353, 1361-62 (Fed. Cir. 2008) (citing *Paice* (Rader, J., concurring)). *Amado*, notably, dealt with a slightly different issue than an ongoing royalty; the issue on appeal was establishing damages for infringing acts from judgment through appeal, while an injunction is stayed.

77 *See also Boston Scientific v. Johnson & Johnson*, No. C 02-00790, 2009 WL 975424, at *5 (N.D. Cal. Apr. 9, 2009); *Paice II*, 609 F.3d at 626-27 ("Once judgment is entered, ongoing infringement by the adjudged infringer is willful; that factor, along with the potential for enhancement, the potential impact of *res judicata*, and many additional factual factors significantly change the ongoing royalty calculus."); *id.* at 628 ("Failing to take into account the change in legal relationship between the parties would be manifestly unjust to Paice."); *Creative Internet Adver.*, 674 F. Supp. 2d at 860. *But see Ariba*, 567 F. Supp. 2d at 918 ("[G]iven the role of assumptions in economic analysis of damages, it is logically inconsistent to argue that a calculation based upon assumptions of infringement and validity would change when those assumptions are replaced by jury findings of the same facts.").

RECENT DEVELOPMENTS IN THE LAW OF OBVIOUSNESS

Rachel Krevans and Matthew Chivvis
Morrison & Foerster LLP
San Francisco, CA

I. STATISTICS

From October 2008 through December 2009, the Federal Circuit heard approximately 51 appeals involving dispositive issues of obviousness. In 21 of those, the Federal Circuit affirmed or affirmed-in-part determinations that one or more claims of a patent was obvious, compared with one where it reversed and six where it vacated. This reversal rate contrasts that for underlying determinations of nonobviousness as only nine were outright affirmed or affirmed-in-part, whereas eight were reversed. Again, six of such determinations were vacated. The Federal Circuit ultimately found claims obvious in the majority of the cases—57 percent. The court found claims nonobvious in only 20 percent of the cases. The rest of the underlying obviousness determinations—about 24 percent—were vacated. These statistics are set out in the chart below. The remainder of this article summarizes some of the more notable cases.

Federal Circuit Obviousness Determinations	Number of Opinions	Percent of Total
Total Opinions 10/08-12/09	≈ 51	—
Affirmed Obviousness Finding	21	41%
Affirmed Nonobviousness Finding	9	18%
Reversed Obviousness Finding	1	2%
Reversed Nonobviousness Finding	8	16%
Vacated Obviousness Finding	6	12%
Vacated Nonobviousness Finding	6	12%
Claims Found Obvious	29	57%

II. NOTABLE CASES

Perfect Web Techs., Inc. v. InfoUSA, Inc., 587 F.3d 1324 (Fed. Cir. 2009) (affirming summary judgment of obviousness)

Judges: Linn, J.; Dyk, J.; Prost, J.

Opinion: Judge Linn filed the opinion of the court.

Summary: Perfect Web filed suit against InfoUSA alleging infringement of a patent claiming methods of managing bulk email. Contemporaneous with claim construction, InfoUSA moved for summary judgment of invalidity. After conducting a claim construction hearing, but without issuing a formal claim construction order, the district court granted InfoUSA's motion. It assumed for purposes of summary judgment that Perfect Web's claim constructions were correct. The district court found that the first three elements of the key claim at issue were disclosed by a single reference and that the fourth and final element would have been obvious. This last element consisted of repeating the first three elements "until said calculated quantity [of emails] exceeds said prescribed minimum quantity."

On appeal, the parties narrowed their argument primarily to the single issue of whether the last step of the key claim would have been obvious. In particular, InfoUSA contended that this step was the "common sense" addition to the first three steps. In analyzing this argument, the Federal Circuit noted "Common sense has long been recognized to inform the analysis of obviousness if explained with sufficient reasoning." Accordingly, the Federal Circuit stated

> We therefore hold that while an analysis of obviousness always depends on evidence that supports the required *Graham* factual findings, it also may include recourse to logic, judgment, and common sense available to the person of ordinary skill that do not necessarily require explication in any reference or expert opinion.

The court found that evidence to support the district court's "common sense" reasoning appeared in the record: the fact that the last step of the claim recited repeating earlier steps. Separately, InfoUSA argued and the Federal Circuit found that the last step would have been obvious to try. The judgment of invalidity was affirmed.

Amgen Inc. v. Hoffman-La Roche, Ltd., 580 F.3d 1340 (Fed. Cir. 2009) (affirming-in-part and vacating-in-part findings that claims were patentably distinct)

Judges: Mayer, J.; Clevenger, J.; Schall, J.

Opinion: Judge Schall filed the opinion of the court.

Summary: Amgen brought a declaratory judgment action against Roche alleging that Roche's MIRCERA® would infringe Amgen's five patents covering human erythropoietin ("EPO") if imported into the United States. Following summary judgment proceedings, a jury trial, and JMOL, the district court entered judgment that four of Amgen's patents were infringed and not invalid, and that the remaining patent was neither invalid nor infringed. The court further granted Amgen a permanent injunction, enjoining Roche from marketing MIRCERA® in the United States. Roche appealed on multiple grounds, including anticipation, indefiniteness, double patenting, and non-infringement. Amgen cross-appealed.

One of the key issues on appeal was obviousness-type double patenting. Amgen contended that three of its later-filed patents were protected by the § 121 safe harbor against double patenting for divisional applications. Roche argued that these patents were not entitled to safe harbor treatment because they did not issue from true divisionals; rather they issued from continuations of an earlier-filed parent application. While the PTO had issued a restriction requirement in the parent application, it was undisputed that the applications at issue were continuations. Nevertheless, Amgen posited that the applications qualified as divisionals under MPEP § 201.06 because they were (1) carved out of a pending application, (2) contained claims to distinct and independent inventions, and (3) disclosed and claimed subject matter disclosed in the parent application. In essence, Amgen argued that since the applications could have been filed as divisionals, they should be treated as such. The Federal Circuit was not persuaded. It found that because Amgen "checked the continuation application box," it could not now argue that the applications were divisionals. Accordingly, the district court's findings with respect to obviousness-type double patenting for Amgen's product patents were vacated and remanded.

Next, the Federal Circuit addressed the time frame for double patenting analysis of product claims. The court's earlier decision in *Takeda Pharmaceutical Co. v. Doll*, 561 F.3d 1372 (Fed. Cir. 2009), had allowed a patentee to rely on developments in the art as of the date of a later filed application to show that the product claims in that application were patentably distinct over earlier filed product-by-process claims. Roche argued the same rule should apply for any intervening art that shows a claim is not patentably distinct. The Federal Circuit found that such a rule would run contrary to § 120, by denying the patents at issue (all continuations) the benefit of their parent's filing date. In the end, the court limited the holding of *Takeda*:

> Because of § 120, we read *Takeda* to stand for the limited proposition that an applicant can only rely on subsequent developments in the art up to the filing date of the "secondary application" in order to show that alternative processes to make the product render the product and the process for making that product patentably distinct.

The Federal Circuit, however, noted that *Takeda* is a two way street. If Amgen were to rely on intervening art to establish patentability on remand, Roche would be able to rely on its own intervening art to show that any alternative processes for making the claimed products do not render the claims patentably distinct.

The Federal Circuit then turned to Amgen's process patents. The district court had engaged in the proper analysis with respect to these patents by: (1) construing the claims and identifying any differences, and (2) considering whether the differences would have been obvious to one of ordinary skill in the art. The earlier filed patent had claimed transfecting Chinese hamster ovary ("CHO") cells with DNA coding for human EPO. The patents at issue claimed using these cells to produce therapeutically effective amount of EPO. Using an analysis analogous to that conducted under § 103, the Federal Circuit found that one of ordinary skill would not have had a reasonable expectation of success in isolating EPO after transfecting the CHO cells. Thus, it affirmed the district court's finding that the process claims were patentably distinct. After affirming and vacating various other aspects of the district court's judgment, the Federal Circuit ultimately decided not to disturb the district court's injunction.

In re Fallaux, 564 F.3d 1313 (Fed. Cir. 2009) (affirming BPAI finding that the two-way test for obviousness-type double patenting did not apply)

Judges: Schall, J.; Archer, J.; Moore, J.

Opinion: Judge Moore filed the opinion of the court.

Summary: The inventors of the Fallaux application appealed a decision by the BPAI affirming final rejections of various claims for obviousness-type double patenting. The rejections had found that the claims were unpatentable over certain patents to Vogels et al. by way of a single common inventor. In a footnote, the Federal Circuit noted:

> Neither party raised or argued the question of whether a patent may be used as a reference for an obviousness-type double patenting rejection where the patent shares only a common inventor with the application, rather than an identical inventive entity or a common assignee. The Manual of Patent Examination Procedure (MPEP) allows such a rejection. This opinion should not be read to decide or endorse the PTO's view on this issue.

Despite this comment, the court analyzed double patenting over the Vogels patents, and ultimately upheld the rejection.

The key issue on appeal was whether the appropriate double patenting test had been used. In prosecution, the examiner had applied the one-way test to reject the Fallaux claims; the inventors argued that the two-way test should have been applied. Under settled law, the two- way test is only appropriate where: (1) a later filed application that has issued as a patent is used to frame rejections against an earlier filed application, and (2) the PTO is "solely responsible for the delay in causing the second-filed application to issue prior to the first." Substantial evidence supported the BPAI's finding that the inventors chose to file the Fallaux application after the Vogels patents had already issued. Hence the Federal Circuit found that the PTO was not solely responsible for the delay in examining the Fallaux application because the inventors could have included the claims in a parent application.

Perhaps more notably, the inventors further argued that the two-way test should apply because they were not seeking a time-wise extension of the patent right. Indeed, they had attempted to file a terminal disclaimer. This disclaimer was ineffective because the Vogels patents had been assigned to different entities and therefore were not "commonly owned." The court made no further reference to whether the lack of common ownership undermined the double patenting rejection itself. Rather, in addressing this issue, the Federal Circuit stated a second justification for obviousness-type double patenting: "harassment by multiple assignees," and affirmed the decision of the BPAI.

Fresenius USA, Inc. v. Baxter Int'l, Inc., 582 F.3d 1288 (Fed. Cir. 2009) (reversing-in-part JMOL of non-obviousness)

Judges: Newman, J.; Gajarsa, J.; Dyk, J.

Opinions: Judge Gajarsa filed the opinion of the court. Judges Newman and Dyk filed concurring opinions.

Summary: Fresenius appealed and Baxter cross-appealed a final judgment entered by the district court. The patented technology involved touch screen hemodialysis machines. Fresenius initially filed suit seeking a declaration that certain of Baxter's patents were invalid and not infringed. Baxter counterclaimed for infringement. Following claim construction, Fresenius admitted infringement, leaving validity as the only issue remaining for trial. The jury found that some of the claims were anticipated and that they all were obvious. But the district court granted Baxter's motion for JMOL and issued a permanent injunction, finding that substantial evidence did not support the jury's verdict.

The Federal Circuit addressed the issue of obviousness, analyzing the "implicit findings necessary to support the verdict" because the jury did not make explicit factual findings. For one claim, the issue was whether the jury could implicitly use anticipation evidence and opinion offered as to an independent claim to support its finding that a dependent claim was obvious. The Federal Circuit held that it could. The court also addressed the breadth of Markush claims, the use of demonstratives, and motivation to combine the prior art, finding support for the jury's obviousness findings. The Federal Circuit thus reversed the JMOL of non-obviousness for most of the claims. The court affirmed the JMOL with respect to six means-plus-function claims, however, because Fresenius had neither identified the corresponding structure for the claimed means, nor compared it to structures in the prior art. The permanent injunction was therefore vacated, so that it could be reconsidered in light of the court's holding.

While this case is an interesting example of how the court reviews obviousness decisions by both juries and judges, it is perhaps more interesting for the two concurring opinions, in which Judges Newman and Dyk debated the merits of stays pending reexamination. Judge Dyk joined the opinion "on the understanding that it does not foreclose the district court in its discretion from staying further proceedings pending the outcome of the reexamination before the U.S. Patent and Trademark Office." Judge Newman wrote separately to respond to Judge Dyk's proposal, noting that "at this stage in this protracted litigation, after full trial and decision in the district court, after full appeal and ruling of the appellate court and with only a modified remedy remaining on remand, such an action would be inappropriate." While she noted she is a "strong supporter of the principle of reexamination," Judge Newman analyzed the long pendency of stays pending reexamination, and concluded that a stay would be "contrary to the precepts of expeditious and just resolution of disputes" in the case at hand.

Callaway Golf Co. v. Acushnet Co., 576 F.3d 1331 (Fed. Cir. 2009) (vacating jury verdict on obviousness and remanding)

Judges: Linn, J.; Dyk, J.; Prost, J.

Opinion: Judge Dyk filed the opinion of the court.

Summary: Callaway brought suit against Acushnet, alleging that it infringed four patents with claims covering golf balls. Acushnet stipulated that it infringed, leaving anticipation and obviousness as the issues for trial. A jury determined that all of the asserted claims were not invalid except for a single dependent claim, which the jury found invalid for obviousness. Notably, the jury found that an independent claim (from which the invalid claim depended) was not invalid. The district court entered final judgment on the jury's verdict.

On appeal, one key issue was whether the district court properly entered judgment on what Acushnet contended was an "irreconcilably inconsistent" jury verdict. The Federal Circuit reviewed the issues not unique to patent law under Third Circuit law. First, it found that that verdict was inconsistent because "[a] broader independent claim cannot be nonobvious where a dependent claim stemming from that independent claim is invalid for obviousness." Second, the court found that the verdict could not be read in such a way as to resolve the inconsistency; the evidence at trial supported either outcome and neither party was entitled to JMOL. The Federal Circuit therefore vacated the district court's judgment and remanded for a new trial on obviousness. After remand, a jury found all the asserted claims obvious.

In re Kubin, 561 F.3d 1351 (Fed. Cir. 2009) (affirming BPAI finding of obviousness)

Judges: Rader, J.; Friedman, J.; Linn, J.

Opinion: Judge Rader filed the opinion of the court.

Summary: The inventors of the Kubin application appealed a decision of the BPAI, which found obvious their claimed DNA encoding a protein known as the "Natural Killer Cell Activation Inducing Ligand ('NAIL')." The key piece of prior art was the Valiante patent, which disclosed "a receptor protein called 'p38'" that is the same protein as NAIL. Thus, the prior art established that the protein encoded by the inventor's claimed DNA was previously known to exist. Valiante, however, disclosed "neither the amino acid sequence of p38 . . . nor the polynucleotide sequence that encodes p38." Nevertheless, the BPAI found the claims at issue obvious in light of Valiante and a laboratory manual providing general methods for cloning genetic material.

The Federal Circuit affirmed the BPAI. In so doing, it overruled the longstanding precedent of *In re Deuel*, 51 F.3d 1552 (Fed. Cir. 1995), which had held that the results of cloning a gene of unknown sequence are non-obvious, even where there was a reasonable expectation of success: "Insofar as *Deuel* implies the obviousness inquiry cannot consider that the combination of the claim's constituent elements was 'obvious to try,' the Supreme Court in *KSR* unambiguously discredited that holding." The Federal Circuit noted that Valiante disclosed the protein of appellants' interest and suggested the protocol for cloning it. Thus, the court found that the claimed invention was "'the product not of innovation but of ordinary skill and common sense.'" The Federal Circuit also considered how *In re O'Farrell*, 853 F.2d 894 (Fed. Cir. 1988), limits "obvious to try" analysis, but found the case fell outside the purview of its exceptions.

Altana Pharma AG v. Teva Pharms. USA, Inc., 566 F.3d 999 (Fed. Cir. 2009) (affirming denial of preliminary injunction on the basis of obviousness)

Judges: Newman, J.; Gajarsa, J.; Ward, J.

Opinions: Judge Ward (District Judge of the Eastern District of Texas, sitting by designation) filed the opinion of the court. Judge Newman filed a concurring opinion.

Summary: Altana sued several generic drug manufacturers for infringement of a patent directed to antiulcer medication after they filed ANDAs, seeking approval to market generic versions of Altana's proton pump inhibitor ("PPI") Protonix.® The defendants admitted infringement, but maintained that the asserted claims were invalid for obviousness. Altana sought a preliminary injunction, which was denied. In denying the injunction, the district court found that Altana had not established a strong likelihood of success on the merits because Altana's prior art patent combined with two other references established a prima facie case of obviousness. The court also found that Altana had not shown the requisite irreparable harm. The denial of the injunction was affirmed on appeal.

In affirming, the Federal Circuit focused its analysis on Altana's likelihood of success on the merits. According to the court:

> If the alleged infringer raises a "substantial question" of invalidity, the preliminary injunction should not issue. The burden on the accused infringer to show a substantial question of invalidity at this stage is lower than what is required to prove invalidity at trial. "Vulnerability is the issue at the preliminary injunction stage, while validity is the issue at trial."

Altana challenged the district court's finding that defendants had raised a substantial question of obviousness. Since the case involved a chemical compound, the court noted that the obviousness analysis could turn on the structural similarities between the claimed compound and the prior art: "Obviousness based on structural similarity may be proven by the identification of some motivation that would have led one of ordinary skill in the art to select and modify a known compound in a particular way to achieve the claimed compound." The Federal Circuit found that the district court did not abuse its discretion in determining that the patent at issue was vulnerable to obviousness attack; Altana's prior art patent used a structurally similar compound that one of ordinary skill could have pursued to develop PPIs. The Federal Circuit also found that the district court did not clearly err in its irreparable harm analysis. Judge Newman concurred in the opinion of the court, noting that she did not believe that the evidence established invalidity of Altana's patent but that deference to the district court was warranted because it had engaged in a discretionary action at a preliminary stage in the case.

Ball Aerosol & Specialty Container, Inc. v. Ltd. Brands, Inc., 555 F.3d 984 (Fed. Cir. 2009) (reversing denial of summary judgment of obviousness)

Judges: Lourie, J.; Clevenger, J.; Linn, J.

Opinion: Judge Lourie filed the opinion of the court.

Summary: Ball sued Limited for infringement of a patent directed to a candle tin with a removable cover that also acts as a base for the candle holder. The district court denied Limited's motion for summary judgment of obviousness and *sua sponte* granted summary judgment of validity to Ball. The Federal Circuit reversed. On appeal, the key issue was whether a motivation to combine the various teachings of the prior art existed. The district court interpreted *KSR* to require that Limited's motivation to combine analysis "should have been made explicit." The Federal Circuit noted that this requirement applied to the district court's analysis, not that of the litigants. Since the problem the invention sought to address was known, and it was undisputed that all of the elements of the asserted claims were in the prior art, the court found that *KSR* compelled a grant of summary judgment of obviousness. Ball's showing of commercial success was not sufficient to outweigh the indication of obviousness from the prior art.

Boston Scientific Scimed, Inc. v. Cordis Corp., 554 F.3d 982 (Fed. Cir. 2009) (reversing denial of JMOL of obviousness)

Judges: Lourie, J.; Rader, J.; Prost, J.

Opinion: Judge Lourie filed the opinion of the court.

Summary: Boston Scientific brought suit against Cordis for infringement of a patent relating to drug eluting stents. At trial, the jury found that the asserted claim was infringed and not obvious. Cordis then filed a motion for JMOL and a new trial, which the district court denied, finding the jury verdict was supported by the evidence. On appeal, Cordis argued that the claim at issue was obvious over a single patent assigned to Medtronic. Boston countered that Medtronic failed to develop the claimed step until after the filing date of the patent at issue, suggesting the claimed invention was not obvious.

In assessing the parties' arguments, the Federal Circuit first addressed the standard of review. The court noted that it reviews the jury's conclusions on obviousness "without deference" and the underlying findings of fact for substantial evidence. The court further clarified:

> While a jury may render a decision on a question of obviousness when it is considering any underlying fact questions, obviousness is ultimately a question of law that this court reviews *de novo*. When we consider that, even in light of a jury's findings of fact, the references demonstrate an invention to have been obvious, we may reverse its obviousness determination.

The court then reviewed the prior art references independent of the jury's implicit findings and found the claim at issue obvious over a single reference. That reference contained aspects of the claimed stent in two separate embodiments shown in different figures. According to the court, one of ordinary skill would have been motivated to combine these figures. The court dismissed Boston Scientific's evidence of secondary considerations due to the strength of the prima facie case. The judgment of the district court was therefore reversed.

Sundance, Inc. v. DelMonte Fabricating Ltd., 550 F.3d 1356 (Fed. Cir. 2008) (reversing JMOL of non-obviousness)

Judges: Dyk, J.; Prost, J.; Moore, J.

Opinion: Judge Moore filed the opinion of the court.

Summary: Sundance sued DeMonte for infringement of a patent directed to segmented covering systems. A jury found that the asserted claim was infringed but invalid for obviousness. Sundance moved for JMOL of non-obviousness, which the district court granted, finding that DeMonte had failed to show reasons to combine. The district court also noted that secondary considerations, such "long-felt but unresolved need" and "copying," supported a finding of non-obviousness. It denied DeMonte's request for reconsideration in light of *KSR*. On appeal, the Federal Circuit considered whether the court properly admitted the testimony of DeMonte's expert—a patent lawyer—on various issues, including obviousness. At trial, Sundance had moved in limine to exclude the expert's testimony. The Federal Circuit found the district court abused its discretion in denying the motion in limine, explaining: "Unless a patent lawyer is also a qualified technical expert, his testimony on these types of issues is improper and thus inadmissible." As a result, the court found that there was no expert testimony supporting the jury verdict of obviousness. It decided that expert testimony was not required, however, because the technology at issue was "simple." Indeed, the Federal Circuit concluded that one of ordinary skill would have found it obvious to incorporate the features of the claimed invention into a cover system, barely addressing secondary considerations raised by Sundance. The court noted, "Secondary considerations of nonobviousness—considered here by the district court—simply cannot overcome this strong prima facie case of obviousness." Accordingly, the district court's judgment was reversed.

Cohesive Techs., Inc. v. Waters Corp., 543 F.3d 1351 (Fed. Cir. 2008) (finding a verdict of non-obviousness does not foreclose a finding of anticipation)

Judges: Mayer, J.; Linn, J.; Prost, J.

Opinion: Judge Linn filed the opinion of the court. Judge Mayer filed a dissenting opinion.

Summary: Cohesive brought three related actions against Waters alleging infringement of two patents covering high-performance liquid chromatography columns ("HPLC"), which are used for separating, identifying, and measuring compounds contained in a liquid. In the first action, a jury found that one of Cohesive's patents was not invalid and that Waters's 30 mu products infringed. The court did not charge the jury on anticipation as requested by Waters. In the second action, the district court granted summary judgment that the same products infringed another of Cohesive's patents, and that this patent also was not invalid. In the third action, the district court granted summary judgment that Waters's 25 mu products did not infringe either patent.

On appeal, the Federal Circuit found that the district court was wrong to conclude that granting JMOL on anticipation would not harm Waters in light of the jury's opportunity to consider obviousness. The court explained:

Despite the often quoted maxim that anticipation is the "epitome of obviousness," novelty under 35 U.S.C. § 102 and nonobviousness under 35 U.S.C. § 103 are separate conditions of patentability and therefore separate defenses available in an infringement action. While it is commonly understood that prior art references that anticipate a claim will usually render that claim obvious, it is not necessarily true that a verdict of nonobviousness forecloses anticipation. The tests for anticipation and obviousness are different. For instance, the court noted that "secondary considerations are not an element of a claim of anticipation," so not every anticipated invention will also have been obvious. The Federal Circuit therefore remanded the issue of anticipation back to the district court.
In dissent, Judge Mayer agreed that the district court should have submitted the issue of anticipation to the jury. He disagreed, however, that the case should be remanded in light of the jury's finding on obviousness. According to Judge Mayer, "The majority's assertion that a claim can be anticipated but not obvious flies in the face of a long line of precedent to the contrary."

REFUSALS TO DEAL: IS ANYTHING LEFT; SHOULD THERE BE?

Daniel R. Shulman
Gray, Plant, Mooty, Mooty & Bennett, P.A.
Minneapolis, MN

I. INTRODUCTION

This paper will explore the state of the law on refusals to deal. This was an area of the jurisprudence under Section 2 that used to put some rather stringent limits on the conduct of monopolists and would-be monopolists. At least until *Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, LLP*, 540 U.S. 398 (2004), and *Pacific Bell Telephone Co. v. linkLine Communications, Inc.*, 129 S. Ct. 1109 (2009), which appear, when taken together, to have seriously placed those limits in doubt.

This paper will compare the pre-*Trinko* law of refusals to deal with the current state of the law under *Trinko* and *linkLine* (hereafter collectively "*TrinkLine*"). It will then ask what if anything is left on the limits of a monopolist's refusal to deal. Finally, this paper will ask whether the economic theory underlying *TrinkLine* has been called in question, in whole or in part, by our current economic difficulties, and, if so, what should be left of the law on refusals to deal.

II. THE LAW BEFORE *TRINKLINE*

Any discussion of the law of refusals to deal must start with two caveats. First, the discussion relates to unilateral refusals to deal, not concerted refusals to deal, which fall generally within the purview of Section 1 and the group boycott cases, such as *Klor's, Inc. v. Broadway-Hale Stores, Inc.*, 359 U.S. 207 (1959); *Northwest Wholesale Stationers, Inc. v. Pacific Stationery & Printing Co.*, 472 U.S. 284 (1985); and *Fashion Originators' Guild of America v. Federal Trade Commission*, 312 U.S. 457 (1941). Although the law of concerted refusals to deal has gradually shifted in this area from per se treatment to rule of reason treatment (compare *Klor's* with *Northwest Wholesale Stationers*), concerted refusals to deal are generally a concern of Section 1, and not Section 2.

Second, for any unilateral refusal to deal to be actionable under Section 2, the party refusing to deal must have sufficient market power either to amount to monopoly power, or to create a dangerous probability of success in achieving monopoly power. In the absence of actual monopoly power, or a dangerous likelihood of success, there can be neither monopolization nor an attempt to monopolize in violation of Section 2. This means, in practical terms, that before a court even reaches its analysis of whether a refusal to deal violates Section 2, the court will have to go through the exercise of finding a relevant product and geographic market, and determining the refuser's power, if any, in the relevant market. If monopoly or market power in a relevant market cannot be shown, the court never gets to the legality of a unilateral refusal to deal.

The refusal to deal law that is reviewed hereafter thus involves cases in which the refusal is unilateral, not joint, and the refuser has market or monopoly power.

Before *TrinkLine*, the law on refusals to deal was generally that a monopolist was free to refuse to deal, ***but***… The "but" was where the rubber met the road. In *Monsanto Company v. Spray-Rite Service Corporation*, 465 U.S. 752, 760-61 (1984), the Supreme Court articulated the right to refuse to deal:

> A manufacturer of course generally has a right to deal, or refuse to deal, with whomever it likes, as long as it does so independently. *United States v. Colgate & Co.*, 250 U.S. 300, 307, 63 L. Ed. 992, 39 S. Ct. 465, 7 ALR 443 (1919); *cf. United States v. Parke, Davis & Co.*, 362 U.S. 29, 4 L. Ed. 2d 505, 80 S. Ct. 503 (1960). Under *Colgate*, the manufacturer can announce its resale prices in advance and refuse to deal with those who fail to comply. And a distributor is free to acquiesce in the manufacturer's demand in order to avoid termination.

The "but" element was often expressed by courts as a requirement that the refusal to deal not be in furtherance of a plan to coerce the refused to party to agree to an anticompetitive scheme or otherwise promote a monopoly through anticompetitive means. *Tidmore Oil v. BP Oil Co./Gulf Products Div.*, 932 F.2d 1384, 1389 (11th Cir. 1991):

> This is not to say that all refusals to deal are beyond the reach of the antitrust laws. Courts have found the 'agreement' element to be satisfied where a manufacturer uses the threat of termination (1) to coerce retailers to adhere to stated resale prices, *United States v. Parke, Davis & Co.*, 362 U.S. 29, 80 S. Ct. 503, 4 L.Ed.2d 505 (1960), (2) to coerce compliance with a scheme of exclusive territories, *Graphic Products Distributors, Inc. v. ITEK Corp.*, 717 F.2d 1560, 1573 (115h Cir. 1983), (3) to coerce the retailer to refrain from selling competitor's products and to comply with certain tying arrangements, *Perma Life Mufflers*, 392 U.S. 134, 88 S. Ct. 1981, and (4) to obtain the buyer's acquiescence in some other anticompetitive provision, *Eiberger v. Sony Corp. of America*, 622 F.2d 1068 (2d Cir. 1980) (creating territorial restrictions by enforcing compliance with warranty fees on products sold outside dealer's territory).

Another formulation of the rule, as articulated by the Supreme Court, allows a monopolist to refuse to deal only where the monopolist has a legitimate business justification. *Eastman Kodak Co. v. Image Technical Servs.*, 504 U.S. 451, 483 n.32 (1992) ("It is true that, as a general matter, a firm can refuse to deal with its competitors. But such a right is not absolute; it exists only if there are legitimate competitive reasons for the refusal. See *Aspen Skiing Co. v. Aspen Highlands Skiing Co.*, 472 U.S., at 602-605, 105 S.Ct., at 2857-2859.") The citation to *Aspen Skiing Co. v. Aspen Highlands Skiing Co.*, 472 U.S. 585 (1985), is instructive as it reflects the Court's view of that case in 1992, a view that was to change radically with *Trinko* in 2004.

In *Aspen Skiing*, the Supreme Court affirmed a judgment under Section 2 when a large ski company discontinued a long-term relationship with a smaller competitor, which had been permitted to sell lift tickets as part of a four-site package offered by the large company. After first acknowledging the right of a monopolist to refuse to deal, the Court observed that the right was not unqualified:

> The absence of an unqualified duty to cooperate does not mean that every time a firm declines to participate in a particular cooperative venture, that decision may not have evidentiary significance, or that it may not give rise to liability in certain circumstances. The absence of a duty to transact business with another firm is, in some respects, merely the counterpart of the independent businessman's cherished right to select his customers and his associates. The high value that we have placed on the right to refuse to deal with other firms does not mean that the right is unqualified. [472 U.S. at 600-01.]

The limits on the right to refuse to deal, as described by the Supreme Court, citing *Lorain Journal Co. v. United States*, 342 U.S. 143 (1951), is that the right must be exercised "in the absence of any purpose to create or maintain a monopoly." 472 U.S. at 601-02. The Court found the refusal to deal in *Aspen Skiing* actionable because "the monopolist did not merely reject a novel offer to participate in a cooperative venture that had been proposed by a competitor. Rather, the monopolist elected to make an important change in a pattern of distribution that had originated in a competitive market and had persisted for several years." 472 U.S. at 603. The Court observed, "By disturbing optimal distribution patterns one rival can impose costs upon another, that is, force the other to accept higher costs." 472 U.S. at 604 n.31. The Court concluded,

> If a firm has been "attempting to exclude rivals on some basis other than efficiency," it is fair to characterize its behavior as predatory. It is, accordingly, appropriate to examine the effect of the challenged pattern of conduct on consumers, on Ski Co.'s smaller rival, and on Ski Co. itself. [472 U.S. at 605.]

"Attempting to exclude rivals on some basis other than efficiency" would seem on its face to offer an intelligible test for evaluating not only refusals to deal, but all other alleged forms of predatory or anticompetitive conduct. Perhaps the formulation may be deceptively simple, but no one can claim it to be outside the comprehension of the ordinary juror. A number of courts have employed it. *Trans Sport, Inc. v. Starter Sportswear, Inc.*, 964 F.2d 186, 189, 191 (2d Cir. 1992); *Minnesota Mining and Mfg. v. Appleton Papers*, 35 F.Supp.2d 1138, 1145-46 (D. Minn. 1999); *Conwood Co., L.P. v. U.S. Tobacco Co.*, 290 F.3d 768, 783 (6th Cir. 2002); *Universal Analytics, Inc. v. MacNeal-Schwendler Corp.*, 914 F.2d 1256, 1258 (9th Cir. 1990); *Pacific Exp., Inc. v. United Airlines, Inc.*, 959 F.2d 814, 818 (9th Cir. 1992); *Drinkwine v. Federated Publications, Inc.*, 780 F.2d 735 (9th Cir. 1986); *Willamette Dental Group, P.C. v. Oregon Dental Service Corp.*, 882 P.2d 637 (Or. Ct. App. 1995); *Catch Curve, Inc. v. Venali, Inc.*, 519 F.Supp.2d 1028, 1036 (C.D. Cal. 2007); *Compuware v. International Business Machines*, 366 F.Supp.2d 475, 486 (E.D. Mich. 2005); *M & M Medical Supplies and Service, Inc. v. Pleasant Valley Hosp., Inc.*, 981 F.2d 160, 166 (4th Cir. 1993); *Sunshine Cellular v. Vanguard Cellular Systems*, 810 F. Supp. 486, 497 (S.D.N.Y. 1992); *General Industries Corp. v. Hartz Mountain Corp.*, 810 F.2d 795, 803 (8th Cir. 1987); *Blue Cross v. Marshfield Clinic*, 883 F. Supp. 1247, 1257 (W.D. Wis., 1995), aff'd in part, rev'd in part on other grounds, 65 F.3d 1406 (7th Cir. 1995) ; *Tasty Baking Co. v. Ralston Purina, Inc.*, 653 F. Supp. 1250, 1273-74 (E.D. Pa. 1987); *Lepage's Inc. v. 3M*, 324 F.3d 141, 146-47 (3rd Cir. 2003); *Reazin v. Blue Cross & Blue Shield of Kansas, Inc.*, 635 F. Supp. 1287, 1332-33 (D. Kan. 1986), *aff'd*, 899 F.2d 951 (10th Cir. 1990). A search of case law has revealed 36 decisions employing the language since *Aspen Skiing*, and only a few in the past five years - not a particularly large number in light of the number of Section 2 cases going through the courts. As discussed in the next section of this paper, the survival of this test for evaluation of a refusal to deal by a monopolist is probably doubtful after *TrinkLine.*

Several other aspects of the law of refusals to deal pre-*Trinko* deserve mention. First is *Lorain Journal Co. v. United States*, 342 U.S. 143 (1951), cited in *Aspen Skiing*, which appears to be good law, certainly in the view of the current Antitrust Division. In *Lorain Journal*, the Supreme Court found a violation of Section 2 based on a monopoly newspaper's refusal to deal with advertisers who also placed ads with a radio station viewed as a competitor by the newspaper. The case appears to stand squarely for the proposition that a monopolist's refusal to deal violates Section 2 if it results from the other party's unwillingness to enter into an anticompetitive agreement intended to injure competition without any substantial business justification.

Second is the essential facilities or bottleneck doctrine, which imposes a duty to deal on a monopolist in control of a so-called essential facility under the following conditions: "(1) control of the essential facility by a monopolist; (2) a competitor's inability, practically or reasonably, to duplicate the essential facility; (3) the denial of the use of the facility to a competitor; and (4) the feasibility of providing the facility to competitors." *Advanced Health-Care Serv. v. Radford Com. Hosp.*, 910 F.2d 139, 151 (4th Cir. 1990); *MCI Communications v. AT&T*, 708 F.2d 1081, 1132-33 (7th Cir. 1983) (citing *Otter Tail Power Co. v. United States*, 410 U.S. 355 (1973); *Twin Laboratories v. Weider Health & Fitness*, 900 F.2d 566 (2d Cir. 1990); *Delaware & Hudson Ry. v. Consolidated Rail Corp.*, 902 F.2d 174, 178-79 (2d Cir. 1990); *City of Anaheim v. Southern California Edison Co.*, 955 F.2d 1373, 1379, 1380 (9th Cir. 1992) (" A company which has monopoly power over an essential facility may not refuse to make the facility available to others where there is no legitimate business reason for the refusal.") The doctrine finds its origins in *United States v. Terminal R.R. Ass'n*, 224 U.S. 383 (1912), in which the Court found an antitrust violation where an association formed by a group of competitors controlled all terminal facilities through which rail entry could be gained to the city of St. Louis.

Finally, an extraordinarily good overview of the law of refusals to deal pre-*TrinkLine* appears in *Byars v. Bluff City News Co., Inc.*, 609 F.2d 843 (6th Cir. 1979). The plaintiff, Byars, distributed magazines for Bluff City until he was terminated, and then brought a Section 2 claim. The Sixth Circuit observed that in addition to the essential facilities doctrine, a monopolist could be liable for a refusal to deal intended unfairly to preserve its monopoly. The Court's decision is here quoted at length because of the summary it provides of the various categories of unlawful refusals to deal by a monopolist, with a discussion of the leading cases:

> There exist two conceptually similar lines of cases which impose a duty to deal upon a monopolist. The first is a straightforward "intent" test which originated from dicta in *United States v. Colgate & Co.*, *supra*, 250 U.S. at 307, 39 S.Ct. at 468 where the Court stated that a business is free to deal with whomever it pleases so long as it has no "purpose to create or maintain a monopoly." *Poller v. Columbia Broadcasting System, Inc.*, 368 U.S. 464, 468-69 (1962). [The second line of cases is the essential facilities or bottleneck line.]
>
> The intent test was applied in *Eastman Kodak Co. v. Southern Photo Materials Co.*, 273 U.S. 359, 47 S.Ct. 400, 71 L.Ed. 684 (1927). There, the Court affirmed a finding of illegal monopolization where a monopolist refused to deal and the jury inferred monopolistic intent because the defendant's policies were "in pursuance of a purpose to monopolize." *Id.* at 375, 47 S.Ct. at 404. Kodak had monopoly power over the national wholesale market for certain photographic supplies. It desired to vertically integrate and take over retail distribution of these

supplies as well. To that end, it began buying out distributors. One such distributor, however, refused to be bought out. Kodak responded by refusing to sell photographic supplies to the distributor at wholesale prices. From these facts, the jury was permitted to infer an illegal intent to monopolize. *Kodak* comes perilously close to establishing an absolute duty to deal since it permitted a finding of illegal intent where the only evidence of monopolistic purpose was Kodak's desire to buy out retail distributors and its inability to provide an independent business reason for its refusal to deal.

Two other Supreme Court decisions have found refusal to deal by monopolists to be illegal. *Lorain Journal v. United States*, 342 U.S. 143, 72 S.Ct. 181, 96 L.Ed. 162 (1951), a newspaper which was indispensable to local businesses refused to sell advertising space to customers who bought advertising on a local radio station. The Court found that this conduct was designed to destroy the competitor and enjoined it as an illegal attempt to monopolize.

Otter Tail Power Co. v. United States, 410 U.S. 366, 93 S.Ct. 1022, 35 L.Ed.2d 359 (1973), the Court found that a private utility company with monopoly power in a wholesale power market had illegally "sought to substitute for competition anticompetitive uses of its dominant economic power." Id. at 380, 93 S.Ct. at 1031. In addition to supplying wholesale power, Otter Tail also was responsible for providing power at retail to the inhabitants of certain small towns. Otter Tail's wrongful conduct was its refusal to sell power at wholesale prices or to "wheel" power to municipalities who proposed to replace Otter Tail in the local retail market with their own power companies.

Two lower court cases have found a monopolist's refusal to deal unlawful because it was done with intent to preserve a monopoly. *See Poster Exchange, Inc. v. Nat'l Screen Service Corp.*, 431 F.2d 334 (5th Cir. 1970), Cert. denied, 401 U.S. 912, 91 S.Ct. 880, 27 L.Ed.2d 811 (1971) (". . . National Screen intentionally used the monopoly power it had at the manufacturing level to eliminate Poster as a competitor at the distributor-jobber level."); *United States v. Klearflax Linen Looms*, 63 F.Supp. 32, 39 (D. Minn. 1945) ("(monopolist) cannot refuse to sell if its design and purpose is to establish a wrongful monopoly.").

* * *

Whether the analysis of refusals-to-deal by monopolists is premised on the "intent" test or the "bottleneck" test, there is a discernible uniformity of holdings of illegal refusals-to-deal in various factual contexts.

First, there are situations where a monopolist uses its monopoly power in one market to distort competition in another market by refusing to deal. This is forbidden, at least absent a valid business justification for the refusal to deal. *Six Twenty-Nine Prods, Inc. v. Rollins Telecasting, Inc.*, 365 F.2d 478 (5th Cir. 1966); *Packaged Programs, Inc. v. Westinghouse Broadcasting Co.*, 255 F.2d 708 (3d Cir. 1958); *United States v. Griffith*, 334 U.S. 100, 68 S.Ct. 941, 92 L.Ed. 1236 (1948)

> (Invalidating use of monopoly power in one geographic market to acquire monopoly power in another).
>
> Second, there is the context in which a monopolist refuses to deal with customers who deal with its rivals. This behavior is inherently anti-competitive; *Lorain Journal, supra*, makes it clear that this is illegal, either as monopolization or attempt to monopolize. See also *North Texas Producers Ass'n v. Metzger Dairies, Inc.*, 348 F.2d 189 (5th Cir. 1965), Cert. denied, 382 U.S. 977, 86 S.Ct. 545, 15 L.Ed.2d 468 (1966); *Kansas City Star Co. v. United States*, 240 F.2d 643 (8th Cir.), Cert. denied, 354 U.S. 923, 77 S.Ct. 1381, 1 L.Ed.2d 1438 (1957).
>
> Third, there is the context in which a group of competitors control an indispensable facility which cannot be easily duplicated. This is the classic case where the "bottleneck theory" applies. Absent valid business reasons, equal access is required for all. See *United States v. Terminal R.R. Ass'n, Supra; Gamco, Inc. v. Providence Fruit & Produce Bldg.*, 194 F.2d 484 (1st Cir.), Cert. denied, 344 U.S. 817, 73 S.Ct. 11, 97 L.Ed. 636 (1952).
>
> Fourth, there is the most conceptually difficult context of all that in which a monopolist seeks to vertically integrate. These were the circumstances in *Eastman Kodak Co. v. Southern Photo Materials Co., supra,* where Kodak cut off one of its retail distributors as part of its efforts to vertically integrate into the retail distribution of photography supplies. Likewise, in *Otter Tail Power Co. v. United States*, 410 U.S. 366, 93 S.Ct. 1022, 35 L.Ed.2d 359 (1973), Otter Tail, among other things, refused to deal when small towns proposed to replace it with their own retail power systems. In that case, Otter Tail used its monopoly power in the wholesale power market to prevent the displacement of its (natural) monopoly in the local retail power market.

Of particular significance is that a number of the cases described in *Byars* are "price squeeze" cases, in which the monopolist vertically integrates while charging prices to its wholesale customers that are sufficiently high so that the wholesale customers cannot compete with the monopolist's prices at retail. These include *Southern Photo Materials* and *Otter Tail Power*. The same price squeeze claim also appears in *United States v. Aluminum Co. of America*, 148 F.2d 416, 436-38 (2d Cir. 1945) (L. Hand, J.), perhaps the leading example of a price squeeze in violation of Section 2. As will be seen, these cases may well not survive *linkLine*.

III. ALONG CAME *TRINKLINE*

A. *Trinko*

The first blow to refusal to deal cases comes in *Trinko*, where the issue was addressed largely in dictum in Justice Scalia's majority opinion. In *Trinko*, the issue before the Supreme Court was the intersection between the Section 2 of the Sherman Act and the Telecommunications Act of 1996, which required local telephone companies like the defendant, Verizon, to provide access to their networks to competitors wishing to sell various unbundled services. The plaintiffs, customers of Verizon's competitors, claimed that Verizon had not complied with its obligations under the Telecommunications Act to provide access to its network, and that Verizon's violation of the Act constituted anticompetitive conduct in violation of Section 2. The narrow issue for the Supreme Court was whether conduct violative of the Telecommunications Act could form the predicate for a violation of Section 2 where the defendant satisfied Section 2's other requirement, the possession of monopoly power in a relevant market.

The Court could easily have answered this question in the negative, and eventually did, but not before Justice Scalia authored what was essentially a paean to monopoly. In fulsome language, which can appropriately be characterized as Chicago-School-and-Bork-rampant, Scalia enthused:

> The mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free-market system. The opportunity to charge monopoly prices—at least for a short period— is what attracts "business acumen" in the first place; it induces risk taking that produces innovation and economic growth. To safeguard the incentive to innovate, the possession of monopoly power will not be found unlawful unless it is accompanied by an element of anticompetitive *conduct.*
>
> Firms may acquire monopoly power by establishing an infrastructure that renders them uniquely suited to serve their customers. Compelling such firms to share the source of their advantage is in some tension with the underlying purpose of antitrust law, since it may lessen the incentive for the monopolist, the rival, or both to invest in those economically beneficial facilities. Enforced sharing also requires antitrust courts to act as central planners, identifying the proper price, quantity, and other terms of dealing—a role for which they are ill-suited. Moreover, compelling negotiation between competitors may facilitate the supreme evil of antitrust: collusion. Thus, as a general matter, the Sherman Act "does not restrict the long recognized right of [a] trader or manufacturer engaged in an entirely private business, freely to exercise his own independent discretion as to parties with whom he will deal." *United States* v. *Colgate & Co.,* 250 U. S. 300, 307 (1919). [540 U.S. at 407-08.]

After paying lip service to *Aspen Skiing's* admonition that "[t]he high value that we have placed on the right to refuse to deal with other firms does not mean that the right is unqualified," Scalia went on to emasculate not only *Aspen Skiing*, but also *United States v. Terminal R.R. Ass'n* and every other prior decision holding a monopolist's refusal to deal unlawful, in language that is clearly dictum:

> Under certain circumstances, a refusal to cooperate with rivals can constitute anticompetitive conduct and violate Section 2. We have been very cautious in recognizing such exceptions, because of the uncertain virtue of forced sharing and the difficulty of identifying and remedying anticompetitive conduct by a single firm. [*Id.* at 408.]

As to *Aspen Skiing*, according to Scalia, that case "is at or near the outer boundary of Section 2 liability." *Id.* at 409.

> The Court there found significance in the defendant's decision to cease participation in a cooperative venture. See *id.,* at 608, 610-611. The

> unilateral termination of a voluntary *(and thus presumably profitable)* course of dealing suggested a willingness to forsake short-term profits to achieve an anticompetitive end. *Ibid.* Similarly, the defendant's unwillingness to renew the ticket *even if compensated at retail price* revealed a distinctly anticompetitive bent. [*Id.*; emphasis in original.]

Scalia found the case before the Court distinguishable from *Aspen Skiing* and *Otter Tail* based on the absence of a prior course of dealing between the parties, which was terminated by the defendant, and based on the absence of evidence of anticompetitive intent in Verizon's refusal to deal. *Id.* at 410.

He also took the somewhat gratuitous step of trying to squash any future Section 2 claims premised on the essential facilities doctrine:

> We conclude that Verizon's alleged insufficient assistance in the provision of service to rivals is not a recognized antitrust claim under this Court's existing refusal-to-deal precedents. This conclusion would be unchanged even if we considered to be established law the "essential facilities" doctrine crafted by some lower courts, under which the Court of Appeals concluded respondent's allegations might state a claim. *See generally* Areeda, Essential Facilities: An Epithet in Need of Limiting Principles, 58 Antitrust L. J. 841 (1989). We have never recognized such a doctrine, see *Aspen Skiing Co.,* 472 U. S., at 611, n. 44; *AT&T Corp.* v. *Iowa Utilities Bd.,* 525 U. S., at 428 (opinion of BREYER, J.), and we find no need either to recognize it or to repudiate it here. It suffices for present purposes to note that the indispensable requirement for invoking the doctrine is the unavailability of access to the "essential facilities"; where access exists, the doctrine serves no purpose. Thus, it is said that "essential facility claims should . . . be denied where a state or federal agency has effective power to compel sharing and to regulate its scope and terms." P. Areeda & H. Hovenkamp, Antitrust Law, p. 150, Paragraph 773e (2003 Supp.). Respondent believes that the existence of sharing duties under the 1996 Act supports its case. We think the opposite: The 1996 Act's extensive provision for access makes it unnecessary to impose a judicial doctrine of forced access. To the extent respondent's "essential facilities" argument is distinct from its general Section 2 argument, we reject it. [*Id.* at 410-11.]

In a footnote, Scalia distinguished the *Terminal Railroad Association* and *Associated Press* cases as being concerted action cases, and not refusals to deal by monopolists: "Respondent also relies upon *United States* v. *Terminal Railroad Assn. of St. Louis,* 224 U. S. 383 (1912), and *Associated Press* v. *United States,* 326 U. S. 1 (1945). These cases involved *concerted* action, which presents greater anticompetitive concerns and is amenable to a remedy that does not require judicial estimation of free-market forces: simply requiring that the outsider be granted nondiscriminatory admission to the club." *Id.* at n.3.

Finally, after holding that Verizon's violations of the Telecommunications Act could not be a predicate for a Section 2 violation, Scalia held that Verizon's refusal to share its network did not fit within any exceptions to a monopolist's right to refuse to deal, for the very reason that the Telecommunications Act provided an extensive regulatory framework and an administrative remedy for those denied access. *Id.* at 411-15.

In other words, according to Scalia, monopoly is good; monopolists should be free to do what they want because the market will fix everything; and don't complain if a monopolist violates a regulatory scheme by refusing to deal because those excluded can invoke the regulatory scheme the monopolist has violated.

Underlying all this are two possibilities repugnant to Scalia: first, that a monopolist might be found to have done something anticompetitive which really may not be anticompetitive; and, second, that judges and juries, who are not really qualified to evaluate the conduct of monopolists, may be asked to do so.

> Against the slight benefits of antitrust intervention here, we must weigh a realistic assessment of its costs. Under the best of circumstances, applying the requirements of Section 2 "can be difficult" because "the means of illicit exclusion, like the means of legitimate competition, are myriad." *United States* v. *Microsoft Corp.,* 253 F.3d 34, 58 (C.A.D.C. 2001) (en banc) *(per curiam).* Mistaken inferences and the resulting false condemnations "are especially costly, because they chill the very conduct the antitrust laws are designed to protect." *Matsushita Elec. Industrial Co.* v. *Zenith Radio Corp.,* 475 U. S. 574, 594 (1986). The cost of false positives counsels against an undue expansion of Section 2 liability.
>
> * * *
>
> Even if the problem of false positives did not exist, conduct consisting of anticompetitive violations of Section 251 may be, as we have concluded with respect to above-cost predatory pricing schemes, "beyond the practical ability of a judicial tribunal to control." *Brooke Group Ltd.* v. *Brown & Williamson Tobacco Corp.,* 509 U. S. 209, 223 (1993). Effective remediation of violations of regulatory sharing requirements will ordinarily require continuing supervision of a highly detailed decree. We think that Professor Areeda got it exactly right: "No court should impose a duty to deal that it cannot explain or adequately and reasonably supervise. The problem should be deemed irremedia[ble] by antitrust law when compulsory access requires the court to assume the day-to-day controls characteristic of a regulatory agency." Areeda, 58 Antitrust L. J., at 853. In this case, respondent has requested an equitable decree to "[p]reliminarily and permanently enjoi[n] [Verizon] from providing access to the local loop market . . . to [rivals] on terms and conditions that are not as favorable" as those that Verizon enjoys. App. 49-50. An antitrust court is unlikely to be an effective day-to-day enforcer of these detailed sharing obligations. [*Id.* at 414-15.]

The fear of forcing a court to act as a regulatory agency in prescribing terms of enforced dealing is, however, somewhat of a red herring. If a monopolist has refused to deal for anticompetitive reasons without a valid business justification, a finder of fact is surely competent to make such a determination. If it does, and it further finds that the refusal to deal has caused injury, then damages can be awarded. Further, if the refusal is continuing and threatens future harm, the court can enter an injunction prohibiting the refusal to deal unless supported by a valid business justification. Neither remedy places the court in the position of operating as a regulatory agency, or requires the court to engage in any type of fact-finding, application of law, or administration of justice for which it is unsuited or with which it is not familiar.

B. *linkLine*

Before one reads *linkLine*, one should carefully scour the opinion in *Trinko*, as well as every other antitrust decision prior to *linkLine* since the passage of the Sherman Act, for the term "antitrust duty to deal." The term is nowhere to be found in *Trinko*. It appears in only two other decisions, *Z-Tel Communications v. Sbc Communications*, 331 F.Supp.2d 513, 540 (E.D. Tex. 2004), and *Virginia Vermiculite v. W.R. Grace & Co.-Conn.*, 108 F.Supp.2d 549, 597, 600 (W.D. Va. 2000). In both cases, the context is a discussion of the essential facilities doctrine and the duty of a monopolist to share such facilities under certain conditions. The point is that prior to *linkLine*, the term either did not exist, or had no meaning outside the essential facilities doctrine. To say that a firm had no antitrust duty to deal was to use a phrase without content, beg the question, and engage in circular reasoning. A firm has an antitrust duty to deal when the courts say that a refusal to deal will violate Section 2. The relevant inquiry is under what circumstances courts will so hold. To be informative and intelligible, a decision should say that under the particular or generalized facts of record, a monopolist may or may not refuse to deal, and why. To refer in the abstract to the presence or absence of an antitrust duty to deal is really to state nothing except a bald conclusion without explanation or rationale, and to be intellectually evasive and unforthcoming.

The foregoing disquisition is prompted by the decision in *linkLine*, where, like Athena springing full-grown from the brow of Zeus, the term "antitrust duty to deal" suddenly appears full-grown and full-blown, without reasoned development, in Supreme Court antitrust jurisprudence.

In *linkLine*, the plaintiff, a provider of DSL services, claimed that the defendant, Pacific Bell, was subjecting the plaintiff to a price squeeze by charging the plaintiff a price for wholesale access to Pacific Bell's network that was too high to permit the plaintiff to compete with Pacific Bell in the retail sale of DSL services—a classic price squeeze in the mold of *Alcoa*. At the time the case went to the Supreme Court, there was no evidence that Pacific Bell was selling its DSL services at retail below any measure of cost. The first paragraph of the Court's opinion, written by Chief Justice Roberts, summarizes the issue, as seen by the Court, and its disposition:

> The plaintiffs in this case, respondents here, allege that a competitor subjected them to a "price squeeze" in violation of Section 2 of the Sherman Act. They assert that such a claim can arise when a vertically integrated firm sells inputs at wholesale and also sells finished goods or services at retail. If that firm has power in the wholesale market, it can simultaneously raise the wholesale price of inputs and cut the retail price of the finished good. This will have the effect of "squeezing" the profit margins of any competitors in the retail market. Those firms will have to pay more for the inputs they need; at the same time, they will have to cut their retail prices to match the other firm's prices. The question before us is whether such a price-squeeze claim may be brought under Section 2 of the Sherman Act when the defendant is under no antitrust obligation to sell the inputs to the plaintiff in the first place. We hold that no such claim may be brought.

Five short paragraphs later, "antitrust duty to deal" is born, as the Court states, "In *Verizon Communications Inc.* v. *Law Offices of Curtis V. Trinko, LLP,* 540 U. S. 398, 410 (2004), we held that a firm with no *antitrust duty to deal* with its rivals at all is under no obligation to provide those rivals with a 'sufficient' level of service." (Emphasis added.) Four paragraphs later: "We granted certiorari, 554 U. S. ___ (2008), to resolve a conflict over whether a plaintiff can bring price-squeeze claims under Section 2 of the Sherman Act when the defendant has no antitrust duty to deal with the plaintiff."

In analyzing Pacific Bell's pricing at the wholesale level, the Court observes that apart from predatory pricing,

> There are also limited circumstances in which a firm's unilateral refusal to deal with its rivals can give rise to antitrust liability. *See Aspen Skiing Co.* v. *Aspen Highlands Skiing Corp.,* 472 U. S. 585, 608-611 (1985). Here, however, the District Court held that AT&T had no such antitrust duty to deal with its competitors, App. to Pet. for Cert. 84a-85a, and this holding was not challenged on appeal.[2]
>
> 2. The Court of Appeals assumed that any duty to deal arose only from FCC regulations, 503 F. 3d, at 878-879, n. 6, and the question on which we granted certiorari made the same assumption. Even aside from the District Court's reasoning, App. to Pet. for Cert. 77a-85a, it seems quite unlikely that AT&T would have an antitrust duty to deal with the plaintiffs. Such a duty requires a showing of monopoly power, but—as the FCC has recognized, 20 FCC Rcd., at 14879-14887-the market for high-speed Internet service is now quite competitive; DSL providers face stiff competition from cable companies and wireless and satellite providers.

The footnote is both troubling and puzzling. The Supreme Court says that it is proceeding under the assumption that any requirement imposed on Pacific Bell to deal with the plaintiff arises only under FCC regulations. If so, the case is indistinguishable from *Trinko*, and could have been summarily reversed on that basis. But the Court goes on to observe that "the market for high-speed Internet service is now quite competitive; DSL providers face stiff competition from cable companies and wireless and satellite providers." If this is a finding, then the Court is saying that Pacific Bell cannot have monopoly power in the relevant market, which is another basis for summary reversal. If Pacific Bell cannot possibly be a monopolist, then there is no reason to discuss refusals to deal, price squeezes, or predatory pricing. The entire opinion begins to look more and more like dictum. This is particularly so, in light of the plaintiff's having conceded prior to oral argument that it was no longer asserting a Section 2 price squeeze, and instead was asking only for remand to the district court to prosecute a retail predatory pricing claim.

Nonetheless, the Court ploughs ahead:

> The challenge here focuses on retail prices—where there is no predatory pricing—and the terms of dealing—where there is no duty to deal. Plaintiffs' price-squeeze claims challenge a different type of unilateral conduct in which a firm "squeezes" the profit margins of its competitors. This requires the defendant to be operating in two markets, a wholesale ("upstream") market and a retail ("downstream") market. A firm with market power in the upstream market can squeeze

> its downstream competitors by raising the wholesale price of inputs while cutting its own retail prices. This will raise competitors' costs (because they will have to pay more for their inputs) and lower their revenues (because they will have to match the dominant firm's low retail price). Price-squeeze plaintiffs assert that defendants must leave them a "fair" or "adequate" margin between the wholesale price and the retail price. In this case, we consider whether a plaintiff can state a price-squeeze claim when the defendant has no obligation under the antitrust laws to deal with the plaintiff at wholesale.

In answering, the Court begins, "A straightforward application of our recent decision in *Trinko* forecloses any challenge to AT&T's *wholesale* prices." It certainly does. The Court could have stopped right there. Instead, it goes on to coat its decision with the oobleck of "antitrust duty to deal."

> Given that Verizon had no antitrust duty to deal with its rivals at all, we concluded that "Verizon's alleged insufficient assistance in the provision of service to rivals" did not violate the Sherman Act. *Id.,* at 410. *Trinko* thus makes clear that if a firm has no antitrust duty to deal with its competitors at wholesale, it certainly has no duty to deal under terms and conditions that the rivals find commercially advantageous.

With all due respect, this proposition is by no means self-evident and is rather self-serving question begging. Even if Pacific Bell had no obligation to sell to the plaintiff at wholesale, the real question presented in *linkLine*, assuming that Pacific Bell was a monopolist, was what obligations did Pacific Bell have once it chose or started to deal with the plaintiff. For instance, does Section 2 require that a monopolist that chooses to do business at wholesale not intentionally disadvantage customers that also compete with the monopolist at retail? Or that if the monopolist does so, the monopolist have a valid business reason or efficiency justification for doing so? These are not unreasonable or novel requirements for a monopolist that enters into a course of dealing with customers, and find support in the jurisprudence of Section 2 in such cases as *Aspen Skiing*, *Southern Photo Materials*, *Byars*, and *Alcoa*. Because it has already decided the issue, however, and masked its reasoning with the conclusory invocation of "no antitrust duty to deal," the Court never confronts the real issue raised in *linkLine*.

Instead, relying on "no antitrust duty to deal," the Court concludes:

> But a firm with no duty to deal in the wholesale market has no obligation to deal under terms and conditions favorable to its competitors. If AT&T had simply stopped providing DSL transport service to the plaintiffs, it would not have run afoul of the Sherman Act. Under these circumstances, AT&T was not required to offer this service at the wholesale prices the plaintiffs would have preferred.

Thus, the Court sweeps away almost a century of monopoly price squeeze cases, from *Southern Photo Materials*, to *Alcoa*, to *Otter Tail*. The closest the Court comes to any rationale for why it is overruling these cases *sub silentio* is a reference to the opinion of Justice, then Judge, Breyer in *Town of Concord* v. *Boston Edison Co.,* 915 F.2d 17, 25 (1st Cir. 1990):

> One of our colleagues has highlighted the flaws of this test in Socratic fashion:
>
> "[H]ow is a judge or jury to determine a 'fair price?' Is it the price charged by other suppliers of the primary product? None exist. Is it the price that competition 'would have set' were the primary level not monopolized? How can the court determine this price without examining costs and demands, indeed without acting like a rate-setting regulatory agency, the rate-setting proceedings of which often last for several years? Further, how is the court to decide the proper size of the price 'gap?' Must it be large enough for all independent competing firms to make a 'living profit,' no matter how inefficient they may be? . . . And how should the court respond when costs or demands change over time, as they inevitably will?" *Town of Concord, supra,* at 25.

But, as shown, this is not really an issue in a price squeeze case. No court has to decide whether a price is fair, or set any price whatsoever in a price squeeze case, if the inquiry is whether the monopolist is intentionally trying to disadvantage rivals, or lacks any valid business or efficiency justification for doing so. Is the monopolist trying to exclude rivals on a basis other than efficiency? That begins and ends the inquiry. The remedy need be nothing more than damages or an injunction to cease the practice. The court need set no price.

It is somewhat disconcerting when *Alcoa* and *Otter Tail* are consigned to dust without analysis and express overruling, and John Roberts is dismissive of Learned Hand.

A final word on *linkLine* is that Justice Breyer, concurring in the judgment with Justices Souter, Stevens, and Ginsburg, does not rise to the majority bait to declare *Alcoa* and the other price squeeze cases a dead letter. Instead, he points out that the wholesale pricing of Pacific Bell arose from a regulatory regime, and under those circumstances, as in *Trinko*, the plaintiff's resort should be to the regulators and not Section 2 of the Sherman Act. Breyer aptly observes that price squeeze claims should still retain their viability under traditional and appropriate circumstances:

> A "price squeeze" claim finds its natural home in a Sherman Act Section 2 monopolization case where the Government as plaintiff seeks to show that a defendant's monopoly power rests, not upon "skill, foresight and industry," *United States* v. *Aluminum Co. of America,* 148 F. 2d 416, 430 (CA2 1945) *(Alcoa),* but upon exclusionary conduct, *United States* v. *Grinnell Corp.,* 384 U. S. 563, 576 (1966). As this Court pointed out in *Verizon Communications Inc.* v. *Law Offices of Curtis V. Trinko, LLP,* 540 U. S. 398 (2004), the "'means of illicit exclusion, like the means of legitimate competition, are myriad.'" *Id.,* at 414 (quoting *United States* v. *Microsoft Corp.,* 253 F. 3d 34, 58 (C.A.D.C. 2001) (en banc) *(per curiam)*). They may involve a "course of dealing" that, even if profitable, indicates a "willingness to forsake short-term profits to achieve an anticompetitive end." *Trinko, supra,* at 409. See, *e.g., Aspen Skiing Co.* v. *Aspen Highlands Skiing Corp.,* 472 U. S. 585, 610-611 (1985); Complaint in *United States* v. *International Business Machines Corp.,* Civil Action No. 69 Civ. 200 (S.D.N.Y., filed Jan. 17, 1969), Paragraph 20(c), reprinted in F. Fisher, J. McGowan, & J. Greenwood, Folded, Spindled, and Mutilated: Economic Analysis and *U. S.* v. *IBM* 357

> (1983). And, as Judge Hand wrote many years ago, a "price squeeze" may fall within that latter category. *Alcoa, supra,* at 437-438. As a matter of logic, it may be that a particular price squeeze can only be exclusionary if a refusal by the monopolist to sell to the "squeezed customer" would also be exclusionary. But a court, faced with a price squeeze rather than a refusal to deal, is unlikely to find the latter (hypothetical) question any easier to answer than the former.

IV. IS ANYTHING LEFT OF REFUSALS TO DEAL?

After due consideration, the answer is yes. *Aspen Skiing* still stands, even with the shackles hung on it in *Trinko*. Just as Justice Scalia did to *Aspen Skiing* in *Trinko*, future courts may confine *Trinko* and *linkLine* to their facts, specifically where the defendant's duty to deal and pricing are subject to a regulatory regime, as the four concurring justices did in *linkLine*. Despite the broad language and Chicago School rhetoric of *TrinkLine*, those decisions do not expressly overrule any of the prior refusal to deal or price squeeze decisions. Even the cabining of *Terminal Railroad Ass'n* to a concerted refusal to deal context in *Trinko* is accompanied by a disclaimer that the Court has never been asked to rule on the essential facilities doctrine, and is not doing so in *Trinko*. *Aspen Skiing* even retains validity as proscribing an unjustified, anticompetitive refusal to deal by a monopolist ending a prolonged course of dealing. Surprisingly, in *Trinko*, Scalia also reads *Aspen Skiing* as authorizing an examination of a monopolist's state of mind and motive in refusing to deal. If done for an anticompetitive purpose, without a valid business justification, ending a course of dealing may run afoul of Section 2. Thus, there is nothing in *TrinkLine* to preclude reliance on pre-*TrinkLine* jurisprudence, particularly in the event of a future change of direction in the Court's philosophy and outlook on antitrust.

Moreover, bestriding the law of refusals to deal like a colossus is *Lorain Journal*. One would think it near impossible for Scalia, Roberts, or the most extreme Chicago School zealot to defend the conduct condemned in *Lorain Journal*, the refusal of a monopoly newspaper to do business with advertisers placing ads on a competing radio station. Certainly, the current Antitrust Division is strongly on record as not tolerating conduct of the type proscribed in *Lorain Journal*. "VIGOROUS ANTITRUST ENFORCEMENT IN THIS CHALLENGING ERA," Christine A. Varney, Assistant Attorney General, Antitrust Division, U.S. Department of Justice, Remarks as Prepared for the United States Chamber of Commerce, May 12, 2009.

> While the Department is not proposing any one specific test to govern all Section 2 matters at this time, I believe the balanced analyses reflected in the leading cases interpreting the reaches of the Sherman Act provide important guidance in this regard. In particular, leading Section 2 cases – from *Lorain Journal v. United States* to *Aspen Skiing Co. v. Aspen Highlands Skiing Corp*. to *United States v. Microsoft* highlight a common concern regarding the harmful effects of a monopolist's exclusionary or predatory conduct on competition and, ultimately, consumers. Reinvigorated Section 2 enforcement will thus require the Division to go "back to the basics" and evaluate single-firm conduct against these tried and true standards that set forth clear limitations on how monopoly firms are permitted to behave. There can be no better charter for our return to fundamental principles of antitrust enforcement.

> In 1951, the Supreme Court laid down a marker for Section 2 enforcement in its decision in *Lorain Journal*. In that case, the Court made a clear step forward in identifying single firm conduct that crossed the line separating lawful, fair competition from exclusionary, anticompetitive acts.

So long as *Lorain Journal* stands as good law, a formidable dike exists to prevent the seepage of *TrinkLine* to where the Court's conservative bloc would apparently like it to go. Roberts' remarks at the end of *linkLine* put the issue rather bluntly. If the monopolist has no duty to deal in the first place, then when the monopolist does choose to deal, there can be no antitrust consequences for the terms on which it deals: "But a firm with no duty to deal in the wholesale market has no obligation to deal under terms and conditions favorable to its competitors." If this reasoning were followed to its logical consequences, it would potentially wash away *Lorain Journal*, *Alcoa*, *United States v. Microsoft Corporation*, 253 F.3d 34 (D.C. Cir. 2001), and much other Section 2 jurisprudence. Because of *Lorain Journal*, this is unlikely.

As *TrinkLine* recedes in time, notwithstanding the extensive dictum purporting to create a monopolist's bill of rights, future courts are likely to look back on *TrinkLine* as standing for the rather limited and unexceptionable proposition that where a regulatory scheme establishes an obligation for a monopolist to deal, complaints are properly addressed to the regulators and not to courts in the form of antitrust suits.

V. SHOULD ANYTHING BE LEFT OF REFUSALS TO DEAL?

The answer is again yes. Obviously, one of the easiest ways for a monopolist to use its power to disadvantage rivals is to refuse to do business altogether, or to refuse to do business unless its customers accept terms that will exclude the monopolist's rivals, a la Microsoft. Behind *TrinkLine*, there is a century of antitrust jurisprudence condemning such uses of monopoly power. The current administration does not appear likely to accept the invitation of *TrinkLine* to turn its back on history and go soft on monopoly.

Beyond this, however, there is another compelling reason that *TrinkLine* will be limited to the regulatory context, and the law of refusal to deal should enjoy continued viability. Specifically, the economic world view of Scalia and Roberts has been shown by events to rest on false premises, or at least is subject to serious questions concerning its validity. Certainly many, if not most, economists today would not blindly accept Scalia's pronouncement in Trinko that "The mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free-market system."

FTC Commissioner Thomas Rosch has succinctly and brilliantly summarized the current questioning of Chicago School orthodoxy in his article, "The Redemption of a Republican," FTC: Watch No. 743, p. 4 (June 1, 2009). Beyond questioning "whether economics concepts expressed in complex formulae featuring the Greek alphabet are useful," Commissioner Rosch says,

> I have questioned the basic tenets of orthodox Chicago School law and economics as those tenets were set forth by Judge Robert Bork in the Antitrust Paradox—that antitrust law is concerned with maximizing societal welfare; that markets are generally perfect; that, if imperfect, they can and will correct themselves; that, accordingly, rational businesspeople will not engage in predatory conduct (because it is not profit-maximizing since markets will correct themselves). [*Id.* at 5.]

Commissioner Rosch cites, with approval, the views of George Osborne, the "shadow" Chancellor of the Exchequer for the Conservative Party in Britain, rejecting the "efficient markets hypothesis' on the grounds that prices 'do not [and] probably never will' 'accurately represent [] all relevant information,' and that even if they did, 'people would still not respond rationally.'" *Id.* at 6. He adds,

> Additionally, Mr. Osborne is reported to have said that since markets do not operate on the basis of complete and accurate information, they are "prone to speculative bubbles," which justify and indeed require control by "effective regulation."…Finally Mr. Osborne is said to have asserted that banks that are "too big to fail" should be "broken up," rejecting the "Chicago School-inspired dictum that market-generated monopolies are the most efficient distributor of resources and price utility." [*Id.*]

Commissioner Rosch concludes,

> Frankly, (except for "breaking up" banks that are "too big to fail" instead of just not letting them merge in the first place), I intended to communicate all of these things in my January New York remarks: that the "ideology of the free market fundamentalists" is arguably "bankrupt"; that markets cannot be as efficient and self-correcting as orthodox Chicago School economists would have it because information is imperfect and human beings do not always act rationally; that there is a need for government intervention to control speculative bubbles; and that monopolies are not the most efficient distributor of resources. I added that vigorous antitrust enforcement could and should play a substantial role in whatever government intervention is appropriate. [*Id.* at 6-7.]

These remarks have been quoted at length because they are right, and because they summarize as succinctly and accurately as possible the reexamination that is taking place of the thinking that has produced not just the country's greatest economic crisis since the Great Depression, but the antitrust jurisprudence of the past thirty years, and especially the last five, in which the Supreme Court has granted certiorari in nine straight antitrust cases in which the plaintiffs have prevailed, and reversed every one, with a consequent and drastic narrowing of both public and private antitrust enforcement. *Trinko*; *Volvo Trucks No. Am., Inc. v. Reeder-Simco GMC, Inc.*, 546 U.S. 164 (2006); *Illinois Tool Works, Inc. v. Independent Ink, Inc.*, 547 U.S. 28 (2006); *Texaco, Inc. v. Dagher*, 547 U.S. 1 (2006); *Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber Co., Inc.*, 549 U.S. 312, 127 S. Ct. 1069 (2007); *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544, 127 S. Ct. 1955 (2007); *PSKS, Inc. v. Leegin Creative Leather Prods.*, 551 U.S. 877, 127 S. Ct. 2705 (2007); *Billing v. Credit Suisse First Boston, Ltd.*, 551 U.S. 264, 127 S. Ct. 2383 (2007); *linkLine.*

One can justifiably question whether the same thinking that has brought the world economy to its knees, and is now largely if not completely discredited as a general matter, can nevertheless validly constitute a set of guiding principles for the interpretation, application, and enforcement of the antitrust laws. The emperor has been shown to have no clothes outside the court system. Are judges to be taken seriously when they try to dress the antitrust laws in the same invisible motley, telling us that markets are perfect and self-correcting, monopoly is both beneficial and inherently ephemeral, and regulation is not only dangerous to the health of the free market economy, but beyond the ken and abilities of judges and juries? Is *Trinko* to have credibility when it begins with the Court proclaiming that monopoly "is an important element of the free-market system"? These rhetorical questions answer themselves.

The point is that antitrust law has followed its recent trajectory propelled by a set of beliefs and ideology that are now in tatters in the world at large, as governments across the globe rush to rescue financial institutions too big to fail, institutions that have reached their present plight through their own greed, improvidence, and chicanery, as regulators sat on their hands in the belief that, as Commissioner Rosch summarizes, "markets are generally perfect; that, if imperfect, they can and will correct themselves; that, accordingly, rational businesspeople will not engage in predatory conduct (because it is not profit-maximizing since markets will correct themselves)." The world now knows that this one-time orthodoxy is pernicious nonsense.

In his column in The New York Times of August 24, 2009, Nobel laureate economist Paul Krugman wrote:

> Washington, it seems, is still ruled by Reaganism — by an ideology that says government intervention is always bad, and leaving the private sector to its own devices is always good.
>
> Call me naïve, but I actually hoped that the failure of Reaganism in practice would kill it. It turns out, however, to be a zombie doctrine: even though it should be dead, it keeps on coming.
>
> Let's talk for a moment about why the age of Reagan should be over.
>
> First of all, even before the current crisis Reaganomics had failed to deliver what it promised. Remember how lower taxes on high incomes and deregulation that unleashed the "magic of the marketplace" were supposed to lead to dramatically better outcomes for everyone? Well, it didn't happen.

One wonders whether the Supreme Court and the rest of the federal judiciary will be the last group in this country to see what everyone else in the world knows, if they ever do, particularly as they interpret and apply the antitrust laws. One hopes not. Refusals to deal may be a good place to start.

PERSONAL REFLECTIONS ON THE ECONOMIC FOUNDATIONS OF VERTICAL RESTRAINTS

Howard J. Bergman
3M
St. Paul, MN

I am very pleased that The Sedona Conference® will devote some time to discuss the validity of the economic foundations of antitrust.

I have spent the past 24 years working on channel management issues at 3M in the U.S. and Europe. I have had the opportunity to see from the inside the challenges businesses are trying to manage and how the antitrust/competition laws affect their decisions. I would like to share some of my personal observations.

As you will quickly observe, this is a literally a "thought" piece - my perceptions - not a researched review. I apologize for any misperceptions, but I hope that this adds something to the usually lively discussion at Sedona.[1]

My experience

I studied economics in the graduate program of the University of Minnesota, 1974 - 1976. My professors included Leo Hurwicz (welfare economics), Thomas Sargent (macroeconomics), Neil Wallace (monetary theory), James Henderson (microeconomics), and Chris Sims (econometrics). I was a teaching assistant to Walter Heller, and I taught Economics 101. (Hurwicz was awarded the Nobel prize in economics in 2006; Sargent and Wallace were influential in the early development of the "rational expectations" macroeconomic models.)

I decided not to continue in economics and eventually went to law school at the University of Minnesota, where I studied antitrust with Dan Gifford.

I joined the Office of General Counsel at 3M in 1985 after four years in the litigation department of the law firm of Faegre & Benson in Minneapolis, Minnesota. At 3M, I worked closely with various 3M businesses in the consumer and health care markets. I have been able to actively participate in the development of marketing, sales, and channel strategies, including, for example, bundled rebate programs. It has given me an inside view of the complexity of the issues businesses must manage when delivering products and services to customers.

1 *See generally, The Economics of Governance* (Jan. 2005), http://www.acaweb.org/annual_mtg-papers/2005/0107_1645_0101.pdf; *Transaction Cost Economics: An Introduction* (Mar. 1, 2007), http://www.economics-ejournal.org/economics/discussionpapers/2007-3.

In 2004, I moved to Brussels where I have been General Counsel for 3M in Europe, the Middle East, and Africa. I have had the opportunity to work with attorneys throughout the region on a variety of issues involving marketing, sales, and channel management. As a result, I have had exposure to the interesting developments in the competition law of the European Union.

Summary

Antitrust analysts fall into various schools that interpret the antitrust laws somewhat differently (Harvard, Chicago, post-Chicago). Regardless of their differences, they all use a mix of standard price theory, an outdated model of industrial organization, incomplete models of games, and "self-evident" principles to justify their proposed rules for regulating certain economic activities.

Unfortunately, this mix gives an inaccurate understanding of what is really happening among participants in an economy. Standard price theory explains the behavior of spot markets, where customers bring their demand curves, suppliers bring their supply curves, and equilibrium prices clear the spot market. It sheds no light on the complex interactions among suppliers, channel partners, and their customers. The old industrial organization model of vertical relationships is similarly inaccurate. The game theory models provide little understanding of the actual interactions, usually leaving customers out of the picture, and require price theory to explain the welfare benefits/losses of the game. And the "self-evident" principles are mostly tautologies.

Most economic relationships exist outside of spot markets where "price" does not give the economic agents the information they need to enable transactions to occur. Most economic activity looks beyond the instantaneous moment. Businesses make decisions based on their expectations of the future – success of new products, profitability of new plants, and training of employees. But the future is filled with uncertain events, and the partners to business transactions have interests that are not fully aligned.

In addition, people cannot adequately describe everything that needs to be accomplished because of uncertainty or inadequacy of language; they cannot collect, absorb, and use all of the information they need about the quality of their transactions; they do not know what information they can trust; and their counterparts have incentives not to disclose all of the needed information accurately. As a result, their behavior is best described as "bounded rationality". That is, people try to behave rationally, but there are limits to their ability to do so.[2]

Where there is a well-defined set of contract laws, impartial legal systems, and effective enforcement, some of the issues related to "bounded rationality" can be resolved using contracts. But contracts require substantial resources to enter and enforce, and they often cannot adequately protect the parties. As a result, people use a variety of informal methods to verify the information they receive from another party. For example, consumers look to "brands" to help them efficiently evaluate the quality of products. Or they entrust their savings to an investment advisor with an outstanding reputation (and hope he will not spend 150 years in prison).

2 *See* OLIVER WILLIAMSON, MARKETS AND HIERARCHIES (1975).

The issues caused by "bounded rationality" form the basis for much of the economic activity that we see, from the formation of firms, to the contracts defining the relationship among parties, to advertising and branding. Unfortunately, the standard economic analysis used in antitrust law basically assumes away the issues created by bounded rationality. As a result, we have a difficult time understanding why the parties develop certain relationships and therefore our rationale of why we are regulating those relationships is strained.

"Bounded rationality" has been part of the analysis of industrial organization since the 1970s. Consumer economics is currently using the methodology to explore the limitations of consumer rationality. However, a useful model of the market should analyze how all of the economic agents (including suppliers, channel partners, and consumers) organize to compete for resources, produce wealth, and allocate the value created, while confronted with "bounded rationality". (To be complete, we would also analyze the bounded rationality of economists and government agents, that is, the naturally limited ability of economists and regulators to understand what they are regulating and the impact their regulations will have on the economy. It seems illogical to suggest that everyone is limited in their ability to collect, absorb, and use the necessary data, except policy makers.)

I believe that a full model of economic behavior based on "bounded rationality" provides support for a new set of "self-evident" principles that define fairness in the economy. The antitrust conclusions that we would reach from these "self-evident" principles would be the same as those reached by the Chicago School, but with a different rationale. In particular, vertical restrictions and resale price maintenance would be presumed to be "fair", that is "procompetitive" regardless of the size of the supplier, channel partner, or customer, since they are needed to bring alignment among the parties.

A New Set of "Self-Evident" Principles

- Firms compete, not products.
- Firms use all of their resources to get the business of their customers.
- The potential for profits drives innovation and efficiency, not competition.
- Social welfare improves over time due to the "surplus" value created by producers and channel partners.
- Consumers do not "create" value; they consume it.
- The economic agents (suppliers, channel partners, customers, consumers) struggle in the economy to allocate the surplus. There is no reason to measure social welfare based only on the portion of that surplus that goes to any one of those groups (or their subgroups); for example, consumers.
- Consumers play an active role in determining how suppliers and channel partners are organized as well as what goods and services are offered.
- Consumer preferences should be respected. Market share based on consumer preference (for example, brand loyalty or network effects) should not be a basis for describing a supplier as "dominant".
- Aligning the interests of suppliers, channel partners, and customers is essential for a well-functioning economy. Efficient mechanisms for creating alignment (loyalty rebates, exclusive territories, marketing restrictions, maximum prices, etc.) should be presumed to be "pro-competitive".

Observations

1. Antitrust is a set of equitable principles used to justify regulation of certain economic activity.

We like to think of antitrust law as a set of legal doctrines, based on the antitrust laws. However, we know that the antitrust laws only vaguely define the activities we are regulating, certainly as it concerns single firm behavior and "vertical" restraints. We are familiar with the development of the antitrust law over the years and the significant changes that have occurred. Underlying those changes has been an evolution in what we consider "fair".

At its heart, antitrust is a set of equitable principles that reflect our view of the "fairness" of how economic agents attempt to collect the value created by economic activity.

During the "golden age" of antitrust law ("golden" for the attorneys, in any case), antitrust law had a strong populist orientation. Big companies were assumed to be too powerful, and they used their power "unfairly" to harm smaller rivals or channel partners. Restrictions imposed by big companies represented an "abuse" of their power, and the antitrust law was applied to prevent that "abuse".

Price theory gave an "economic" rationale for regulating big firm behavior. In a perfectly competitive market, firms need to be too small to influence demand or strategically eliminate competitors. Since big firms violated that principle, they must be "anti-competitive". Price theory was not only a tool used to understand the market. It was also the gold standard by which markets were measured.

Over time, the populist approach gave way to an "economic" approach. Antitrust commentators noted that protecting the smaller rivals and channel partners was not always in the best interests of consumers. They identified a number of benefits that consumers should realize if big firms were permitted to impose restrictions that harmed their channel partners or smaller rivals.

The Chicago School developed an interpretation of antitrust law based on a new notion of what was "fair". Instead of evaluating fairness on from the perspective of channel partners and smaller rivals, this school looked from the perspective of consumers. Big companies could take actions that harmed smaller rivals or channel partners, if the result ultimately provided a benefit to consumers, in which case, damaging competitors was "pro-competitive".

One element of the Chicago School analysis is the "self-evident" principle that "efficient" firms should survive. The market works well when big firms force smaller firms to demonstrate their efficiency. Inefficient firms deserve to be eliminated from the market, and it is "fair" for more efficient firms to eliminate less efficient firms.

Another element of the Chicago School analysis is the "rule of reason", which requires a close look at the possible benefits and harms that certain behavior by the big firms could bring to consumers. Firms that pass the rule of reason are considered to be acting fairly; their restrictions are "pro-competitive".

Weighing the pro-competitive benefits of a transaction against the anti-competitive harm ought to be a simple mathematical exercise, and judgment should go to the party with the greater weight. Unfortunately, competitive benefits and harms cannot be measured or compared, which means that the concept is not operational. We are left with a tautology: we deem actions "pro-competitive" if they are within our notion of how the market ought to work; if they are not, the actions must by "anti-competitive."

While the concept of the "rule of reason" is technically unworkable, it has helped change the entire notion of what is "fair" in economic relationships. As a result of the concept, antitrust has effectively rejected the populist notion that big companies are bad and that the tools that they use are "abuses". We now look more carefully at the actual nature of economic transactions. This approach makes more "sense" than the populist approach – but only if we accept that the conclusions of the doctrines are "fair".

The economic basis for the Chicago School remains standard price theory. In a "competitive market", firms compete and drive down prices. Lower prices increase consumer welfare. Therefore, restrictions that potentially lower prices to consumers must be a "pro-competitive", even if they eliminate the competitors. In addition, as the "competitive market" drives prices down, inefficient firms are driven from the market, leading to an overall increase in efficiency.

The competition in the "competitive market", as imagined by the Chicago School, may lead to the destruction of firms, but this competition is good: consumers pay lower prices and efficiency is improved. Moreover, using antitrust laws to protect inefficient suppliers or free-riding distributors would prevent the "competitive market" from providing the expected consumer welfare.

A useful development in the analysis of the Chicago School, compared to the previous "golden age" analysis, is the recognition that price theory does not fully describe the behavior of big firms or the interaction among channel partners in the production and distribution of goods and services. There has been a recognition that competition implies that there will be winners and losers and that the market will not deliver "efficiency" or "consumer welfare" if the antitrust laws protect the losers. There has also been recognition that some restrictions might be necessary to give channel partners the incentive to perform certain services without a "free rider" taking advantage of their work.

Not all antitrust commentators, however, have agreed that "fairness" extends to every action taken by large firms that potentially benefit customers. The Chicago School might accept as "self-evident" that the antitrust law is intended to "protect competition, not competitors", but the post-Chicago School commentators wonder how competition can exist without competitors. In addition, the consumer welfare justifications used by the Chicago School to defend certain restrictions seem strained. The Chicago School justifies "exclusive territories" as needed to protect distributors who provide services from "free-riding" distributors, but we also see them used where the service justifying the restriction seems hypothetical.

The post-Chicago School commentators demonstrate with simple game theory models that large firms are able to seize an "unfair" advantage over smaller rivals and channel partners, which could result in higher prices to consumers. For example, it could be "unfair" for a large, multiproduct firm to offer a bundled rebate that cannot be matched by the smaller, equally efficient, one-product rival. If the smaller rival is eliminated, the larger firm can raise prices for at least that product. Since consumer welfare is based on the low price of an individual product, the resolution of the game would harm consumers and must be considered "anti-competitive".

The post-Chicago School is also based on standard price theory. It returns to the idea that firms in a "competitive market" must have enough competition to force prices down. If large firms make it difficult for smaller firms to be a competitive threat, the "competitive market" cannot be "competitive", and we would not expect to obtain the benefits of consumer welfare or efficiency predicted by the Chicago School. In a "competitive market", we would expect to see prices fall to a point where "excess" profits are eliminated. But we see that large firms maintain large profits. We have to assume then that the large firms are manipulating the market to the detriment of consumers.

The debate between the Chicago School and the post-Chicago School will continue as each school works to define what behavior is "fair" in the marketplace. Their will debate will be based on a common platform. Each will use standard price theory to demonstrate whether the actions of large firms comform to the requirements of a "competitive market" and will therefore result in low prices and efficiency, promoting "consumer welfare".

Unfortunately, the "competitive market" defined by standard price theory is inadequate for understanding economic activity. As a result, we do not understand what we are regulating and why the regulation is overall beneficial. The debate will go on, but neither school will have a basis for its conclusions.

2. Price Theory does not describe the economic transactions we are regulating.

Price theory remains the core of our understanding of economic markets. In fact, the "competitive market" that it describes is not only our explanation of <u>how</u> markets work, it is our golden rule of how markets <u>should</u> work.

The basic assumptions of the price theory are that

- All parties have complete information, which is costless to obtain and process.
- Transactions are costless.
- There are only sellers and buyers, none of whom are big enough to affect overall supply or demand.
- The aggregate of the demand of all buyers can be represented by a downward sloping demand curve (demand falls as prices rise).
- The aggregate of the supply of all suppliers can be represented by an upward sloping supply curve (supply increases as prices rise).
- Sellers can be represented by a production function, in which marginal costs first decline then increase as more units are produced.
- The price where the supply curve and demand curve meet is an "equilibrium" price. It clears the market; that is, suppliers do not want to sell additional units and buyers do not want to buy additional units at that price. And it occurs instantaneously.
- At the equilibrium price, the price will equal the average cost of an efficient firm. That is, there will be no firms that operate with higher costs, and there will be no "profits".
- The equilibrium of the model is instantaneous, static, and partial. That is, the model only looks at the supply and demand at an instant in time. It does not examine how supply and demand evolve over time or how firms plan investments for the future. And it only looks at one "market" at a time; it does not look at the effect the change in one market has on another.

This model explains the behavior of spot markets, where customers bring their demand curves, suppliers bring their supply curves, and equilibrium prices clear the market. In the model, there is no need for firms to innovate, no need for firms to invest, no need for firms to advertise, no reason for firms to expand, and no opportunity for firms to generate profits. Since all information is known by all parties and all activity is simultaneous, there are no future contingencies and no need for contracts.

In short, the model does not explain most of the economic activity that we see.

Most economic transactions take place outside of spot markets. Price does not provide all of the information needed for a transaction to occur. Contracts are an obvious example of relationships that exist outside of the "competitive market". They define aspects of a relationship that the market cannot define. And often contracts cannot provide all of the direction needed to maintain an on-going relationship among economic agents. Suppliers form relationships with the customers and their distributors who have a role in the transactions. Consumers develop relationships with the brands they prefer and with the retailers who sell the products with those brands.

In the world of antitrust, we have turned the relationship between theory and reality on its head. We have assumed that standard price theory appropriately describes economic relationships. Instead of using reality to question the validity of the model, we use the model to question the legitimacy of reality; instead of asking why the model does not predict what we see, we question why reality does not match the model.

We will examine some aspects of price theory that play a significant role in the debate among the antitrust schools.

3. The Competitive Market

The gold standard of antitrust economic analysis is the "competitive market" described by standard price theory. We assume that "competitive markets" will behave as the model describes and that markets that do not behave this way are not competitive.

When we examine the market described by the model, however, we discover that it describes a market in which there is **no** competition. "Price" is all that firms and their customers need to know in order to transact their business. Firms can sell their entire output without marketers, sales reps, advertising, promotions, or incentives. Buyers can purchase the entire quantity of the product that they want to purchase at that instant (at the equilibrium price).

In this world of "perfect competition", there are no differentiated products, so there is no need to innovate. The quality of the products in the market is identical, and consumers never have to consider the source of their products.

Since the products are identical, and there is no question about quality or delivery, there is no need to educate buyers about the quality or performance of the products. There is no need to promote the products. Since firms can sell all of their output at the market price, there is no need to advertise. Since there is no future, there is no need to invest. Since prices will be driven to a level where there are no profits, it is not clear why investors would invest in an economic enterprise. There is no need to form relationships. There is no need for contracts. There is no need for all of the economic activity that we take for granted in the economy.

The question, then, is why this market held as the gold standard for economic activity? Why do we measure real competition with this model of "perfect competition"?

It is easy to see why price theory does not predict most of the economic activity that we see.

- Accurate information is often impossible to obtain. Collecting information, digesting it, and using are very costly. Complete knowledge assumes no uncertainty. Price does not give adequate information about the products (for example, quality, availability) needed by the parties in the transaction. As a result, they develop relationships, for example, with contracts, branding, or advertising.
- Negotiating and enforcing transactions are costly.
- There are multiple levels of economic agents, including distributors and retailers whose activities do not fit within the model.
- Where channel partners share customers with suppliers, it is not clear whose demand is represented. If the channel partners establish the price of the goods sold to the customers, it is not clear what demand curve is seen by the suppliers. Similarly, when channel partners set prices, they can alter the supply curve in ways that are against the interests of the suppliers.
- Describing an organization as a production function misses the reason why organizations exist, which is to solve problems created by "bounded rationality".
- Most economic transactions are not about instantaneous transactions, but relationships that develop over time. These relationships include loyalty to brands (including the brands of channel partners).

4. Firms Compete, Not Products

The firms in the perfectly competitive market are narrowly conceived. They are not organizations; they do not have issues with organizational alignment; they do not have issues arising from the difficulty in communicating across functions; there is no hierarchy; there is no intra-organizational competition.

Firms are represented simply as "production functions", usually of a single product. They do not easily provide services with their products – it would be difficult to include the resources needed to provide service in the production function.

Since the production function is two dimensional (price and quantity of a single product), antitrust analysts assume that competition among firms is two dimensional. Firms compete by offering a product at a price. This quickly reduces to single products competing on the basis of price. Antitrust analysis is full of references to the "most efficient" producer of a product or the producer of a more innovative product.

Since the firm is represented only in two dimensions, it is not possible to analyze other sources of efficiency. Organizations are difficult to manage and the larger they become, the more difficult they are to manage. The nature of a firm has been described as a loose aggregation of hostile tribes. Aligning those "tribes" takes skill and resources. When they are aligned, they are capable of acting as a powerful organization, able to develop and deliver a large bundle of products and services globally. They are able to develop a reputation for consistent quality that reduces the resources that customers need to search for appropriate products.

When we imagine the firm in more than two dimensions, we understand that firms compete, not products. And that the firms will use all of the resources at their disposal to increase their value to their customers.

Let's consider the retail market to illustrate these points. Suppliers that understand the basics of the retailer's business model can design programs to increase the total value of what they offer the retailer. The first step is to understand the relationship between the retailers and their customers. Retailers provide value to consumers in a number of ways:

- Provide buying opportunities to consumers
- Identify products to present to consumers
- Market the products
    - Select the products
    - Promote the products
    - Place the products
    - Price the products

While providing these services, retailers try to maximize their economic profits using various tools:

- Increase margins
  The importance of margins to the retailer is obvious.
- Increase turns
  The importance of turns is less obvious to people outside of the retail market. Increasing turns allows the retailer to generate greater returns on a fixed amount of capital. For example, imagine that a retailer purchases products for $100 at the beginning of the month and sells them at the end of the month for $101. Its margin is about one percent. However, the retailer can take the $101 it generated in the first month, save $1, buy products for $100 at the beginning of the second month, and sell all of the goods for $101 at the end of the second month. If the retailer is able to do this throughout the year, it will have generated $12 of profit from its initial investment of $100, an annual return of 12 percent.
- Increase other sources of income
  Retailers recognize the value that they bring to the real estate they occupy, in particular because of the flow of people they can generate. They are able to charge others, either the suppliers or other third parties (banks, florists, coffee shops), for using that space.
- Decrease cost of capital employed
  As with turns, the economic profit of the retailer increases if it can generate the same returns with less capital.
- Decrease cost of operations
  Strong retailers demand support from their suppliers in reducing the cost of operations by funding advertising, by enforcing delivery schedules, and by requiring suppliers to integrate with the retailer's electronic ordering and systems.

By understanding how retailers generate economic profit, suppliers are able to make themselves more valuable to the retailers:

- Increase margins
  - Lower costs. Simply lowering costs will not necessarily increase margins, since the retailer controls the resale price.
  - Increase recommended resale prices. Similarly, increasing recommended resale prices will not necessarily increase retailer's margins, unless the supplier is able to enforce a minimum resale price.
- Increase turns
  - Lower prices. As with lower costs, lower prices from the supplier to the retailer will not necessarily increase turns, since the retailer sets the price of the supplier's products and all other products it sells. Moreover, simply lowering the price of one supplier's products does not necessarily result in increased sales for the retailer. It might only shift sales from one supplier to another. And if customers associate high price with high quality, raising prices might increase turns.
  - Better products
  - Advertising/promotions
  - Customer appeal
  - Category management
- Other sources of income
  - Slotting fees

    Slotting fees are essentially a rental fee that retailers charge suppliers for the privilege of appearing on the retailer's shelf. The retailer can impose them when the power of its brand in attracting customers is more valuable to the supplier than the power of the supplier's brand is to the retailer. It is an example of the paradox of the beekeepers and orchard owners. Simply, beekeepers need to place their hives in an orchard to collect honey; therefore beekeepers will pay orchard owners to place their hives in the orchards. On the other hand, orchard owners need bees to pollinate their orchards; therefore orchard owners will pay beekeepers to place the hives in the orchards. Whether beekeepers pay orchard owners or orchard owners pay beekeepers will depend on the relative bargaining power each has.
  - Rebates
- Decrease cost of capital
  - Payment terms

    We can see the power of payment terms if we modify the story we told for increasing turns. Imagine that the retailer receives the products at the beginning of the month, but does not pay for the products until the end of the month. The retailer would not have to finance the purchase of the products – it can generate profit without capital costs. Moreover, if the payment is due at the end of the second month, the retailer could put the $101 it received for the product in the first month into a money market fund for a month and earn interest on it. By the end of the year, it would collect 11 months of interest on the $101 that it was able to invest each month before paying for the goods, plus the 12 percent.
  - Inventory Management

- Decrease operational costs
  - Ordering
  - Delivery
  - Inventory management
  - Reduction of errors
  - Additional services, for example, preparing planograms

It should be clear that the retailer is not solely focused on product by product price competition. The retailer will prefer working with suppliers that provide it with the best overall economic profit, using all of the tools available.

Let's consider how this would apply to bundled rebates. Obviously, a firm that is efficient in organizing the development, manufacturing, and marketing of many products is able to use that ability to offer its entire bundle of products to get the business of their customers. The bundle might be so attractive to the retailer that the producer of only one of the products cannot compete, even if it is the most efficient producer of that product.

However, the larger firm, if successful, has proven that it is the more efficient supplier of a large bundle of products. Efficiency from the perspective of the retailer is not limited to a single production function, but to the total value of what the supplier offers.

The producer of a single product can compete in a number of ways. Since it is smaller and more focused, it should have lower organizational costs, its decision making time should be shorter, it should be more flexible.

However, the "competitive market" does not require that the highly innovative and efficient single product firm continue as a single product firm. Customers might prefer the total value of a large bundle of products to the benefits of the sole product the single product firm produces, which gives the large, multi-product firm a competitive advantage. If that is the case, the single-product firm may join with other firms, either through alliances or mergers, to create its own bundle. Of course, it might discover that it does not have organizational skill or acumen to successfully manage a large bundle of products, which is a competitive shortcoming when it is the measure used by the customers. It would then be fair to say that the larger company is the more efficient supplier of the bundle.

5. Economic Profits

Economists have observed that the predicted outcome of the standard price theory is absurd: equilibrium prices fall until price equals marginal cost equals average cost, that is, a point where the suppliers earn no profit. This, of course, is not what we see in "real life", nor is it a desirable outcome. No one would invest in a business that could not earn profits.

Economists solve this dilemma by defining "economic profits" (as opposed to accounting profits) to include "opportunity costs", that is, the returns that could be gained from the next best alternative, in the production function.

Unfortunately, there is no basis for the definition. It seems obvious that, over time, industries that do not generate adequate returns will lose their attractiveness. Either investment will decline or the industry will consolidate, permitting the remaining firms to charge higher prices and cover their opportunity costs. (Alternatively, the firms will collude to increase prices to achieve the same result.) However, this process can take years, even decades. It is not a cost that should be considered in an instantaneous, static, equilibrium model.

Moreover, opportunity costs are subjective. Different firms have different expectations for returns. But in that case, a more "efficient" producer can drop out of the market because it has higher expectations for returns. And, as a result, our model of "perfect competition" does not generate the most "efficient" producer. (We should add that investments in research and technology, software development, and clinical trials are also not reflected in the production function, nor are investments in quality and safety.)

In the spot market, firms can remain in business forever selling at prices that equal average cost. It is in the nature of a spot market that they do so. Moreover, since all markets should be "competitive", the profits of all firms – across all spot markets – should be zero, that is, there are no better alternatives.

From a practical perspective, it is not clear how economic profits would be measured. Typically, we ask whether prices are "competitive" by looking at prices when markets are "competitive". And we subtract manufacturing costs to determine the profits in the industry. We do not ask if those prices include "opportunity costs". And how could we? What would the appropriate opportunity cost be?

This is a particularly difficult issue for industries that require a significant investment in research or capital. Taking some examples:

- In the oil industry, once a well begins pumping oil, the cost of continuing to pump is low. However, the price of a barrel of oil should depend on its replacement cost, not the cost of pumping. But, the cost of exploration and discovery of replacement reserves and returns on investment in the installation of oil wells in difficult environments are uncertain. Large investments can be made in exploring regions that are dry. These investments, whether successful or not, have no impact on the spot price of oil. In the spot market, the price reacts as predicted by standard price theory: suppliers and customers reach a price that clears the market. But if that price does not cover the cost of exploration, today's low price will fail to provide an incentive to find replacement reserves.
- Similarly, in the pharmaceutical industry, the significant costs of the industry are in the development and testing of new molecules. The investment in research and development runs in the hundreds of millions of dollars, with little likelihood for success. Once a molecule is appropriately tested and determined to be safe and effective, the cost of actually producing the molecule is minimal. If left to the spot market, the price of pharmaceuticals would be driven below the level needed to cover the costs of the research and development.

The theoretical fallacy is that we have assumed that the production function describes a firm for situations outside of spot markets. As a result, we naturally conclude that, in a "competitive market", prices will equal average cost. If firms are able to generate "excess" profits, it is evidence of a market failure, and antitrust needs to remedy the problem.

However, all of our "real world" experience tells us that this cannot be true, nor do we want it to be true. Profits are the incentive for firms to innovate and become more efficient.

6. Low Prices

Standard price theory leads antitrust analysts from all schools to identify low prices with consumer welfare.

A fundamental problem rests in the assumption that our social goal should be the increased consumption of goods at low prices, and it is not clear why this should be the goal pursued by governmental agencies.

We have already considered the problem that capital costs and the costs of research and development will generally not be covered by a spot market price. As a result, if prices remain at the level of spot prices, industries that require significant long-term investments will not be adequately rewarded. While the standard price theory relies on firms instantaneously leaving the market without cost, in a world of bounded rationality, we would expect the process of firm elimination to be long and painful, and the more capital intensive the industry is, the longer the process will take. We could expect several possible outcomes, including the decline of the industry as a whole, as investors and bright employees look for more profitable industries, the consolidation of the industry and increase in prices, the development of cartels.

While we have considered capital-intensive industries to demonstrate that "spot prices" are not likely to cover the actual costs of operating a business in the long run, the same statement could be made for all industries. The costs of environmental protection, occupational safety, food safety, and safety of air travel, etc., would not be covered by prices in the spot market. Similarly, the spot market is unlikely to reward companies for providing consistently high quality products and services.

An area of investigation in consumer economics is the complex pricing structures that we observe in highly competitive consumer industries such as airlines and telephone service. In each, competitors tend to lead with very low advertised prices, covering very low service levels, to which additional services can be added by the consumer for additional charges, leading to a maze of offerings which are difficult for many consumers to understand. Standard economics would assume the multiple choices are good for consumers since they can choose what they want. Modern consumer welfare analysis uses the bounded rationality of consumers to demonstrate that on average consumers are worse off from having complex choices – they are unable to compare the complicated offerings and often choose a more expensive package with unneeded services.

But looking from the supplier side, it seems likely that competition has driven the advertised price – the spot price – to a level below what is required to cover the additional services. The low cost airlines now need to rebuild the cost of services back into the price of flying, without giving up the "spot" price; similarly with phone companies, health care insurance, mortgage financing, and hospital supplies distributors. (Anyone who worked in the hospital supplies industry in the 1990s will recall the impact of the hospital group purchasing organizations on distributor margins, to the point where their revenues only covered order fulfillment. All other services had to be provided at extra charge.)

In some of these cases, consumers have little opportunity or ability to collect the necessary information and compare the offerings. In areas such as mortgage financing and health care insurance, the risk of a consumer making the wrong choice has significant social implications, suggesting that these are areas where consumers need regulatory protection from the complex pricing systems generated in the market.

We should take the analysis a step further. As we mentioned, the model of "perfect competition" is static. It does not take any interest in the future; there are no incentives to invest or innovate. The primary incentive for those investments, of course, is the prospect of greater profits in the future. This is clearly recognized under patent law which provides a limited monopoly to patent holders in return for making innovations public. There is an interesting hostility between the antitrust community and the patent community. Antitrust law considers monopolies and monopoly profits as potential problems; patent law sees them as one of the driving forces creating wealth in the economy. While antitrust claims that its model of "competition" creates innovation, it is difficult to see how. Ultimately, the lure of profits is the incentive necessary for investment in innovation, and the antitrust model of competition provides no incentives for that investment to take place.

7. The traditional industrial organization diagram

Every antitrust attorney knows how to diagram a vertical market. It requires three boxes, aligned vertically. The top box is labeled "manufacturer"; the middle is labeled "distributor/retailer"; the bottom is labeled "customer/consumer". Manufacturers sell their products to the "distributor"; they have no contact with the customers. Between the manufacturers and distributors, manufacturers have all of the power; they dictate the terms and conditions of their relationship. They are best described by the populist terminology of the "golden-age": they attempt to use their power to limit competition. The distributors sell the products to the customers. Their business model is simple: prepare a catalog of products with prices and take orders from the customers. They do not have a strategy. Consumers have the least power. They are passive participants in the economy. They are offered products by distributors, and they buy them according to their demand schedule.

We have already mentioned that price theory does not shed light on the relationships among suppliers, channel partners, and customers. The vertical diagram traditionally used in antitrust has the same defect – it does not help us understand the relationships among the participants in an economy.

8. Consumers

While current antitrust analysis has consumer welfare at the heart of its analysis, it is not clear that it gives consumers a role in economic transactions, other than looking for low prices. In the traditional industrial organization diagram, the customers are passive – they do not have a strategy and do not influence the transactions. Similarly, the games used to demonstrate the anticompetitive nature of loyalty rebates typically include two manufacturers, but customers (whether the distributor/retailers or consumers) have no strategy other than obtaining low prices.

We could ask how our view of the economy might change if we were to draw the industrial organization diagram in reverse, with consumers at the top, dictating the shopping environment they wish to shop in, as well as the goods and services they wish to purchase, to the distributors/retailers who translate the consumer demand into orders to suppliers.

Suppliers and channel partners, of course, understand the important role that consumers play in dictating what products and services they offer and how they offer them. As Sam Walton noted, "There is only one boss. The customer. And he can fire everybody in the company from the chairman on down, simply by spending his money somewhere else."

Understanding the consumer market is difficult because it is a mass market. It operates through the "laws" of large numbers. Individuals acting in their own interest, when combined with similar individuals, drive the organizational structure of the market. And, while standard price theory assumes that consumers are homogeneous, they obviously are diverse. Different consumers want different bundles of goods and services; they also want different shopping experiences. The assumption has traditionally been that no individual consumer can affect market conditions and therefore they do not. This, of course, is the equivalent in political science of stating that no individual affects the outcome of an election. In a narrow sense, the statement is true. But it misses the bigger picture: mass choices, made by individuals for individual reasons, affect the broad outcome of elections and mass consumer markets.

Suppliers, of course, recognize that consumers can fall into groups and they spend a lot of resources attempting to segment the market by understanding those groups and how consumers associate with different groups over time. They invest heavily in bringing a specific message to the group they want to communicate with. They are only successful, however, when a "large enough" group of consumers accept the message.

Given the nature of mass markets, individual consumers cannot always get what they want. Their ability to demand a bundle of goods and services is limited by the choices made by all other consumers. This leads to the possibility of different groups of consumers dictating to other groups of consumers what bundles of goods and services will be available, and how and where those goods and services will be offered.

A group of consumers might prefer to pay low prices at a large discount store. Another group might prefer to pay higher prices at small, family-owned shops. However, the family owned shops can only survive if enough consumers purchase there, which might include a number of consumers who only want low prices. If the large discount store attracts enough of the price-sensitive consumers from the family owned store, the family owned store will fail. All consumers will be forced to purchase products at the discount store. The outcome will be a benefit to the consumers who want low prices, but a loss to consumers who wanted to purchase at a local family owned store. We cannot state that this is an unambiguous gain to consumers or society in general.

From the perspective of consumers who want to purchase at small local shops, suppliers who impose resale prices on their retailers are, in fact, acting on their behalf. With retail prices fixed, price does not differentiate the large chain retailers from small family owned shops. Rather, convenience (assortment, parking) or customer intimacy will play more important roles. While customers who only want low prices are harmed, customers who want to purchase at small local shops benefit (as do the shopkeepers who serve them). Again, we cannot state that there is an unambiguous gain or loss to the consumers.

We see a similar situation among passengers on airlines. One group of passengers simply want the cheapest flight from Paris to New York City. They do not need a lot of airport infrastructure to move their baggage, coordinate schedules, etc. Another group might want to travel from Marseilles to Las Vegas without having to pick up luggage or leave the airport while traveling. They too will have to fly through Paris and New York

City, but their needs are very different. This group needs the extra infrastructure to coordinate flights, handle baggage, etc., as well as the additional aircraft that enables them to travel to different locations.

A full-service airline might be able to provide the service demanded by the group of customers traveling from Marseilles to Las Vegas, but only if enough passengers fly its planes from Paris to New York City – the plane needs to be full to cover all of the costs of the aircraft, infrastructure, and people required to provide this level of service. A budget airline, of course, does not have the same requirements for aircraft, infrastructure, and people, and can offer low cost flights just between Paris and New York City.

If the low-cost airline is able to attract enough price-sensitive customers from the full-service airline's connection between Paris and New York City, the full service airline would be unable to provide the full service to passengers traveling between Marseilles and Las Vegas. This would benefit the price sensitive passengers who only want to travel from Paris to New York City, but would harm the passengers who want full service. The social gain is ambiguous. In this situation, we would expect the full service airline to use various loyalty schemes as well as its stock of aircraft to keep as many of its passengers as possible. In the end, the full-service airline would be acting on behalf of the passengers requiring full service. If the full- service airline has enough resources to keep enough of its price-sensitive passengers on its flights, the low-cost carrier would fail, possibly harming the passengers who are only looking for low prices, but benefiting the customers who want full service. Again, the benefit or harm to "consumers" is ambiguous.

9. Consumer welfare under the curve

In addition to the issues of consumer welfare caused by the diversity of consumers there is a difficulty with the measure of "consumer welfare" typically used in antitrust analysis.

Economists use the area under the demand curve as a measure of consumer welfare - the larger the area under the curve, the larger the benefit. Since lower prices increase quantities demanded, the area increases with lower prices. As a result, antitrust has developed a self-evident principle that the ultimate goal of social welfare is low prices.

A practical problem is that consumers do not equally share in the welfare gain depicted under the curve. Consumers who only purchase products at the lower price have the benefit of purchasing product at the price they are willing to pay. The consumers who experience the benefit are those who would have gladly paid more. And there is no mechanism for the latter group of consumers to share the gain with the former.

We seem to forget that every point on the demand curve represents a benefit to consumers. At every price point, the consumers who purchase the product are happy with their purchase. Consumers who do not purchase because prices are too high have made the calculation that the product is not worth the price that is charged – they have better things to do with their money.

In fact, if we believe the demand curve really represents "consumer welfare", the best outcome for all consumers would be a firm that perfectly discriminated among all consumers. Every consumer would be able to purchase the product at a price they considered appropriate. While no consumer enjoys a "surplus", every consumer is happy with the bargain.

Looking at this issue more broadly, we need to question why "consumer welfare" should be our measure of the social benefit that we use to judge economic activity. While consumers generate demand for the added value, they do not create it. It seems odd to measure the social value of economic activity by measuring the amount of value that consumers capture. Similarly, a welfare function that only measures the value that channel partners are able to capture at the expense of large firms, as antitrust did in the "golden-age", is equally flawed. Clearly, suppliers have a major role in creating value for customers. The incentive for them to continue creating value is the share of that value that they can capture. But the analysis antitrust currently uses does not seem to recognize this obvious point.

Economic activity not only allocates resources to produce goods and services in a way that leaves everyone better off over time, it also allocates the value created by economic activities among the various participants in the economy. Suppliers, channel partners, and customers are constantly struggling to redistribute the value. Any analysis of economic activity should recognize the legitimacy of each group and sub-group to take actions which allocate to itself a greater share of the value created in the economy and to establish rules of "fairness" that are neutral with respect to the outcome of the allocation.

10. An Alternative Model

If Standard Price Theory provides a useful model of a spot market, is there a model that describes markets where relationships matter?

I believe that elements of a model have been developed, but that they have not been brought together in a complete model yet. Oliver Williamson, in his book *Markets & Hierarchies* (1975), describes an alternate model of industrial organizations that is explicitly based on bounded rationality. Similarly, consumer economists are using bounded rationality to develop a deeper understanding of consumer behavior. (DG SANCO, the consumer protection agency of the European Commission, is looking at these models to identify areas where greater regulation is required.)

Williamson identifies a number reasons why rationality is limited, beginning with the limited ability of people to effectively collect, digest, and make use of accurate information. These activities have real costs associated with them, and are often simply impossible. The quality of the information we collect is often suspect. A related problem is the ability of people to communicate what they know, even when they want to. We only need to consider that there are some things that are easier to learn by doing.

Uncertainty is a key factor in the analysis. We live in a world of different types of risks. Some are measured as known probabilities and can be insured against. Others are random and uninsurable. And some are in between – they might be measurable, but they are not insurable. A function of economic activity is to shift the allocation of risks and uncertainty among the participants in the economy.

Bounded rationality, uncertainty, and the inability to have full confidence in the quality of information one party gives to another, have significant effects on why and how firms are organized. These factors explain why certain transactions take place in a "spot market", why others take place through contract among parties, and why others are simply taken out of the market and placed in a hierarchy. The fact that firms are in a continual process of integrating and outsourcing functions demonstrates the on-going management of the issues created by bounded rationality.

However, in *Markets & Hierarchies*, Williamson only examines the "supply" side of activity. For a complete model, we need to understand the behavior of subsets of suppliers, channel partners and consumers.

Recent research in consumer economics uses "bounded rationality" to describe the limited ability of consumers to make rational choices, and the ability of firms to take advantage of that limited ability. It is leading some regulators to question whether there are certain products/services where the complexity is too great for consumers to manage, with the conclusion that many consumers would in fact benefit from less choice, and with the next conclusion these consumers would be better off if regulators limited the choices available to consumers.

Of course, we already do this in many product areas: we do not permit hospitals to choose to use inexpensive, but unsafe, medical devices or pharmaceuticals; we do not permit workers to choose to work with ineffective safety equipment. It would not be difficult to imagine that the welfare of some consumers could be generally improved by limiting the types of credit available to them. For example, we might conclude that "too many" consumers do not make credit decisions often enough to understand the risks that they are taking and that the downside of those risks are too great to allow uninformed consumers to take. In which case, it might be appropriate for a consumer protection agency to limit the choices available to consumers.

11. Brands

One topic of consumer economics that deserves special consideration is the existence of "brands". As we mentioned, standard price theory has no explanation for the existence or value of brands.

To appreciate the value of brands, we only need to consider the difficulties we would confront in a world without brands when selecting which products to buy. We would go to a shop where we would be confronted by a shelf lined with products with generic names and no indication of the source of the product or its quality. We might try a product and decide we like it. We would return to the store to purchase the product again, but there is no assurance that the product we purchase will be the same as the product we purchased previously. We might want to search for the original product in other stores, but we have no way to identify the product we are looking for. We would have no certainty that the products we purchased would ever be what we really wanted. We would become totally passive purchasers who have no influence on the source or "quality" of the products we buy.

We quickly understand that "brands" are a tool used by consumers to overcome the limitations of their "bounded rationality". Product quality varies, not only between well-made and poorly-made, but also performance, look, feel, smell, taste, etc. Consumers want to know if the product they are purchasing meets the requirements that they expect each time they purchase the product. Obviously, consumers cannot test each product for quality before using it. Instead, they use "brands" to help them identify products that meet their requirements for "quality". By developing a "brand", the brand owner delivers valuable information to the consumer, value that the consumer is willing to pay for.

We should note that "private label" brands have the same origin. "Private label" brands are simply the brands of well-known companies who outsource the production of products, but that customers trust for an expected level of quality. Where the private label brand is the brand of a retailer, the retailer has taken the responsibility to ensure that its "branded" products meet the quality expectations of its customers.

> In addition to identifying the source of a product, brands serve another function, which is to allow consumers to self-select a group (image) they want to be members of. Desire to be a member of a group is instinctive among people and consumption of certain products is one way that people express their membership. (It has been observed that the Goth tribe is a group of young people who demonstrate their distaste for conformity by dressing alike). The statement that a brand allows people to make about themselves has value above and beyond the product that is branded. A sweatshirt might be a sweatshirt, but a sweatshirt with a favorite sports team name is a statement. The decision of a consumer to identify with a particular brand should be respected. And the creator of the brand should be entitled to collect the full value that consumers place in that brand.

12. Consumers and Market Power

We previously discussed the role of consumers in determining the bundles of goods and services offered by retailers, and how and where they are offered. Ultimately, a "brand" is successful because consumers have determined that the "brand promise" of the brand has value. Suppliers and their channel partners invest significant resources to influence consumers in evaluating their brands, but consumers have the final "vote". A supplier that acquires significant market share because its brand is highly valued by consumers does not have "market power" – its "power" exists only because a large enough group of consumers have selected the "brand". The supplier will quickly lose its market share if it fails to continually convince consumers of the value of the brand.

We have a similar situation when network effects make a product or service more valuable as more consumers use it. The supplier whose product benefits from the network effect has power only because the consumers value the benefits of the network effect. Attempts by regulatory agencies to inhibit the network effect reduce the value of the network to the consumers who benefit from it. As we previously discussed, there might be a group of consumers who would benefit from limiting the benefits of the network effect, but the effect on total welfare is ambiguous.

13. A hypothetical illuminating the need for vertical restrictions based on bounded rationality

Let's say we develop and manufacture high-performance supplies that are used in the manufacturing processes of other companies. These products might seem simple but in fact involve significant expertise in materials sciences to develop and produce them. Thanks to our innovative engineers, we are experts. We are able to charge a premium price for these products, and because of our production skills, we make significant profit margins.

Our employees are motivated by the opportunity to use their skills to help solve interesting problems for our customers. They also enjoy the higher than average compensation and benefits packages that can be offered by a highly profitable company. As a result, we are able to attract very bright new employees to bring new ideas to our labs and plants, and we feel confident that we will continue to develop useful new products and services that are highly valued by our customers.

Our shareholders, of course, are motivated by our high margins. We have no trouble attracting investors or creditors when we need to fund new projects. Since we care about our margins, we will not make products if we cannot obtain significant profit margins. We will invest our money elsewhere, **even if we are the most efficient producer of those products**.

We simply do not have expertise in the development and production of our products. We also have superior organizational skills. It is not easy coordinating the activities of thousands of people, working in different functions, in different countries, with different objectives. It has been observed that a corporation is a loose aggregation of hostile tribes. It is difficult to get all of its tribes aligned and working cooperatively. And as it becomes harder to manage, the larger and more diverse it becomes. Transmitting information and establishing direction requires significant effort and overhead. The organization necessarily becomes more bureaucratic than small organizations.

Though it takes a lot of skill and acumen to manage a large organization, when the organization is effective, it is capable of creating a diverse array of products and services that it is able to deliver to customers around the world. Let's say that we have learned to manage our organizations very effectively, and we work hard to transmit knowledge and skills among our employees. Since we have these organizational skills, we have an excellent record in working with customers, developing solutions for them, and delivering the right products and services at the right time. We have developed significant credibility with many customers who have come to trust our ability to deliver the products and services that we have promised and we help find solutions to their problems.

As a supplier of premium priced products, we need to convince our customers (the people who use our products) that the superior performance of our products and services outweighs the additional cost. This requires a relationship with the customer. The customer must have confidence that our performance story is accurate, especially if they need to change their process to take advantage of our superior performance. Once they are convinced that we are credible and that our products provide the expected superior performance, they are likely to pay the higher price that we would like to charge for our products. Over time, if we are successful, our "brand promise" will help maintain our credibility, and we can spend less time trying to convince our customers to believe us. In that case, the "brand" gives the customer significant efficiencies in searching for appropriate products and gives us a very valuable asset.

Of course, there are other companies who would also like to sell products and services to the same customers. Some of these companies are large, bureaucratic organizations like ours that produce a wide variety of products and services, though not necessarily equivalent with ours. Other businesses are small, nimble companies that supply a very limited number of products.

All of these businesses will attempt to persuade customers that their **total** offering of products and services would provide more value to the customers than ours. They will also attempt to demonstrate that their organizations have as much credibility in providing products and services as ours.

We would expect that larger businesses like ours, will have an advantage in organizational depth and breadth, in terms of products and services and geographic coverage, but will be slower and more costly, since they need a large bureaucracy to manage the organization. The smaller businesses often lack the bundle of products and services and the organizational depth and breadth of the larger businesses, but they should be quicker, more nimble, and less costly than the larger businesses.

Regardless of size or complexity, the firms will use their own advantages to become the preferred provider of the products and services they can offer. It takes skill to persuade a customer that the customer should prefer one business' products and services over another. The sellers need to understand what the customers are trying to accomplish, how they are trying to accomplish it, and how the different combinations of goods and services will be of value to them. And the value to customers might not be based just on the specific goods and services and the prices they are offered at. The value might be influenced by the supplier's ability to provide other financial incentives; it might also be influenced by the supplier's credibility, reputation, and ability to deliver.

An obvious question is who will convince the customer that our entire bundle of goods, services, and reputation are worth the price we are asking.

If we have a sales force, our sales reps can make the call, develop the relationship, and provide the information needed to convince our customer. If we do not have a sales force, we might work through a distributor that has its own sales force. And, of course, we could have various combinations of the two sales forces. We can send our sales force to convince the customers that they should purchase our products for a premium price, and the distributor can send its sales force to call on the same customers and promote any of the products that the distributor carries, including our competitors.

(We should note that advertising in the consumer market plays the same role as our sales force in the business-to-business market. The advertising provides the contact between the supplier and the customers, but since the market is a mass market, individual contact is not possible.)

We have some (but not complete) confidence that our sales rep will properly represent our company as credible and demonstrate that the superior quality of our products justifies our higher prices. We can provide them with training, guidance, corrective action, and financial incentives. If all else fails, we can replace them.

When we work with distributors, we will have a complex relationship with them. They are a part of our sales force and supply chain – we rely on them to promote our products and deliver them to our customers. However, they can also be our competitors, promoting a different bundle of goods and services than ours. In fact, they might even be promoting their own goods and services in competition with ours.

As a result, we cannot be as confident that our distributor's sales force will be aligned with our interests. We do not necessarily have the same goals in selling our products or the same incentives to sell them, and we are faced with fundamental questions about our relationship:

- Will our distributors make an effort to find new customers for our products and train their sales reps to promote the benefits of our products? Or will they simply look for customers who are already buying our products from other distributors? Or will they try to persuade our customers to use a lower-priced competitive product that generates higher margins for the distributor?
- How do our distributors intend to position our products among their product mix? Will they see the same value in our products that we see? It is, of course,

> easier for a distributor to sell a product based on price than performance. So will they make an effort to sell the benefits? Or will they use our products as loss-leaders? Or will they use our popular brands to draw in customers, and then persuade the customers to purchase a competitive product.

A large distributor can take advantage of its strong position and demand large discounts from its suppliers on large orders. However, the distributor does not have to pass its discount to its customers. In fact, it does not have to sell the products to its customers. Instead, it can become an unintended wholesaler of our products to other distributors.

Similarly, a distributor can tell us that it would like a special price to participate in a bid. We would prefer to give the special pricing only to our distributors that will use the special price to bid our products. Other distributors could use the knowledge of our special price in order to submit a bid for a competitive product. Even if the distributor that bids our product wins the bid, we do not have assurance that they will only use the discounted product in sales to the customer that was asking for bids.

In business terms, how do we maintain the "loyalty" of our distributors? (For those who have studied business ethics, loyalty is the ethical imperative that no one likes to talk about, but that drives a significant amount of behavior in business relationships. It is the unacknowledged twin of "trust".)

If we find distributors whose interests are relatively aligned with ours, how will we reward them, compared to distributors whose interests are at best neutral relative to ours? A simple example:

> Distributor A has worked 12 months with a customer to demonstrate the value of using our products in its operations. The customer has agreed to write operational protocols to use our products and is now issuing a tender for our products. We need to find a mechanism to reward Distributor A for its efforts. If we do not, our distributors will not take the initiative to persuade customers to use our products. The easiest method is to give Distributor A the exclusive right to submit a bid for providing our products and to charge a price to the customer that compensates Distributor A for the services they provided. A second best method is to give a special discount to Distributor A for the bid. But, in that case, we are paying for the service provided by the distributor to the customer that we think the customer should pay for.

Once we have the interests of our distributor aligned with ours, how do we keep it aligned? In order to improve the ability of our distributors to sell our products, we might give them training in technical aspects of our products, sales techniques, sales management, organizational skills, and customer service. But what prevents the distributors from using those skills to sell products that directly compete with ours?

As we have noted, our distributors are not only an extension of our sales force, representing and promoting our products, they are also our competitors. A special case where we compete directly with our own distributors is the sale of the distributor's "own-label" products. Let's consider this scenario:

> We have developed a strong position in a product category by sending our sales reps to manufacturing sites to demonstrate the value of our products. But our customers purchase our products (with many others) from its distributors. The

> distributors establish the prices at which they sell our products. They know which of their customers purchase our products and which purchase competitive products; they know the quantities that their customers purchase, the prices they pay; and, of course, the prices that we and our competitors charge for the products. In fact, they have more information about our customers purchasing decisions than we have.
>
> One of our distributors has begun to sell low-cost, own-brand products similar to ours. They might have one of these strategies: (1) they could direct their sales reps to promote our products and use their own-brand products for customers who are only interested in price; or (2) they could direct their sales reps to switch customers to their own-brand products, and use our products as a lead-in to the sale.
>
> If the distributor follows the second strategy, it is in a strong position to compete against us. It can set the price of our products and its own-brand products to increase the sales of its own-brand products.
>
> Naturally, we would like to encourage the distributor to follow the first strategy: we would like to protect our sales efforts from competition from distributors who are supposed to be representing our products, who have complete information about customer purchases of our products, and who set the prices of our products to the customer.

We can quickly identify a number of contractual tools to help align the interests of our distributors with our own. They are very familiar to the antitrust community: resale price maintenance and vertical restraints.

However, we would use a different analysis to justify the use of the restraints. We would presume that the supplier has a legitimate interest in restricting the activities of its channel partners in order to align their interests. Unlike the Chicago School, we would not have to imagine that a distributor is providing actual "services" that need protection from free-riders.

For example, if we look from the perspective of alignment, we would probably resolve the Dentsply case in favor of Dentsply.[3] Traditional antitrust analysis looked at the ability of Dentsply to "exclude" its competitors from the strongest dental supplies distributors, and this appears "unfair". From the perspective of alignment, however, we can see what Dentsply was trying to achieve. It was "unfair" for its distributors to use the substantial income they made selling Dentsply products to support the sales of Dentsply's competitors. Dentsply simply gave the distributors a choice: you can be loyal to Dentsply or to the competitors, but not both, so make a choice. While most Dentsply distributors chose Dentsply, it was because they did not want to give up the income generated by Dentsply products.

The Dentsply example leads to a separate discussion regarding the protection of competitors. It is possible that small, single-product competitors of a large diversified company have difficulty entering the channels, even though they are more efficient (in some sense) and more innovative (in some sense) than the large firm. However, we should not presume that the only way for the economy to enjoy the benefits of the efficiency and

3 *U.S. v. Dentsply Int'l, Inc.,* 399 F.3d 181 (3d Cir. 2005).

innovation are through the small, single-product firm. It is possible that customers in the industry more highly value the ability to purchase a broad range of products from a single source of supply. The solution, then, is not to restrict the larger firm from requiring loyalty from its distributors – the solution is to encourage the smaller partner to find larger partners through which it can compete as part of a broad bundle, either through an alliance or merger.

Conclusion

We began our discussion with the idea that antitrust is a set of equitable principles. The equitable principles have been based on populist ideals supporting small firms against big firms or the economic approach that is based on "consumer welfare".

We have looked at the standard economic models used to analyze economic activity and questioned their relevance to the market they are supposed to describe. We have also suggested that a model based on "bounded rationality" among all parties would give us greater insight into the reasons for certain organizational structures and behaviors.

I believe that such a model would support a new set of "self-evident" principles:

- Firms compete, not products.
- Firms use all of their resources to get the business of their customers.
- The potential for profits drives innovation and efficiency, not competition.
- Social welfare improves over time due to the "surplus" value created by producers and channel partners. Consumers do not "create" value; they consume it.
- The economic agents (suppliers, channel partners, customers, consumers) struggle in the economy to allocate the surplus. There is no reason to measure social welfare based only on the portion of that surplus that goes to any one of those groups (or their subgroups), for example, consumers.
- Consumers play an active role in determining how suppliers and channel partners are organized as well as what goods and services are offered.
- Consumer preferences should be respected. Market share based on consumer preference (for example, brand loyalty or network effects) should not be the basis for antitrust action.
- Aligning the interests of suppliers, channel partners and customers is essential for a well-functioning economy. Efficient mechanisms for creating alignment (loyalty rebates, exclusive territories, marketing restrictions, maximum prices, etc) should be presumed to be procompetitive.

The antitrust conclusions that we would reach from these "self-evident" principles would be the same as those reached by the Chicago School, but with a different rationale. In particular, vertical restrictions and resale price maintenance would be presumed to be "fair", that is "procompetitive", regardless of the size of the supplier, channel partner, or customer, since they are needed to bring alignment among parties that are involved in the economic activity.

I believe these principles add an interesting light to the debate among schools of antitrust. But my own rationality is bounded.

THE SUPREME COURT'S 21ST CENTURY SECTION 2 JURISPRUDENCE:[1] PENELOPE[2] OR THERMOPYLAE?[3]

John DeQ. Briggs & Daniel J. Matheson[4]
Axinn Veltrop Harkrider LLP
Washington, DC

In recent decades, the Supreme Court has been in an affectionate embrace with unilateral conduct by a dominant firm. The Court has lauded the stimulating effect of efforts to achieve monopoly and generally has been reluctant to declare unlawful conduct except where particular practices are overwhelmingly likely not to represent competition on the merits. But at the same time these themes have been played differently by a number of lower courts in significant cases. And the announced enforcement intentions of the new administration's competition agencies - the Federal Trade Commission ("FTC") and the Antitrust Division of the Department of Justice ("DOJ") - also point in a different direction, exhibiting skepticism about, if not hostility to, the Supreme Court's serenade to monopoly and its virtues.

There is thus something of a struggle shaping up for the heart and soul of antitrust. It remains to be seen whether the Supreme Court will continue to be dominated by the Chicago-informed antitrust economics and resulting law; whether the lower courts, being tugged in the other direction by the executive branch and the plaintiffs bar, will follow the lead of the Supreme Court, or have to be pulled along somewhat by the heels; or whether the pendulum will swing back past center at all. Much depends on the politics of the Supreme Court in the next few years, but there is also at least some sense in the Congress that U.S. antitrust is out of step with competition law and policy in the rest of the world.[5]

---

1 We leave it to each reader to decide his or her own answer posed in the title. We suspect that where readers stand on the questions depends in some measure on where they sit.

2 In Homer's ODYSSEY, Penelope waits 20 years for the final return of her husband, during which she has a hard time snubbing marriage proposals from 108 suitors, many odious. She is a symbol of fidelity in the face of temptation.

3 The Battle of Thermopylae took place over three days during the second Persian invasion of Greece in September 480 B.C. It was fought between an alliance of Greek city-states, led by Sparta, and the Persian Empire of Xerxes. Vastly outnumbered, the Greeks held up the Persians for seven days in total (including three of battle) at the pass of Thermopylae, before the rear-guard was annihilated in one of history's most famous last stands. The battle has become a symbol for courage against overwhelming odds.

4 Mr. Briggs is Co-Chairman of Axinn Veltrop & Harkrider LLP's Antitrust Group and Managing Partner of the firm's Washington, D.C., office. He is a former Chair of the American Bar Association's Section of Antitrust Law (1995-96). Prior to joining AV&H, he served for more than a decade as Chair or Co-Chair of Howrey LLP's Antitrust Practice Group and then as Managing Partner, Strategy & Planning. Mr. Matheson is an associate in the Washington, D.C., office of Axinn Veltrop & Harkrider LLP. Both Mr. Briggs and Mr. Matheson have been counsel to the prevailing parties in several of the cases discussed herein.

5 *See* Letter of from 22 Congressmen to Christine Varney and Jon Leibowitz (Sept. 18, 2009) (on file with author). They express their " ... increasing concern ... about developments in international competition policy, how the EC is shaping the global competitive environment, and the impact these developments are having on American companies." The signers of the letter pointed specifically to recent or ongoing proceedings involving Google, IBM, Intel, Microsoft and QUALCOMM, and then went on to say:

> "…with so many of the world's successful technology companies based in the United States and very few located in the European Union or elsewhere, the Administration should be an advocate of the "American Way" both at home and in foreign jurisdictions. Otherwise, DG Comp will become the de facto super regulator in competition markets, and its approach in managing competition will shape commerce worldwide. Indeed, DG Comp already is spending millions of Euros exporting its competition policy to emerging markets like China."

See John DeQ. Briggs, *The U.S. Competition Law Regime is Losing the Competition in the World Market for Competition Regimes,* EUROPEAN AFFAIRS (Fall 2009).

I.

A brief overview of historical enforcement may help put the current state of affairs into perspective. For much of the 20th century both the judiciary and the executive branch wielded the Sherman Act to combat practices by which dominant firms disadvantaged smaller rivals. Section 2, which governs the unilateral conduct of dominant firms,[6] was directed against trusts in the first three decades of the century.[7] Then, after the decade and a half of the Great Depression and World War II, during which antitrust went into the closet in favor of the sort of centralized planning marked by the Industrial Recovery Act, antitrust emerged invigorated and refreshed. For some time after the war, the government brought § 2 cases against firms whose strength for dominance was perceived to present structural barriers to competition.[8] During this post-World War II period of aggressive enforcement, the Supreme Court and lower courts suggested that the Sherman Act condemned the use of monopoly power "to gain a competitive advantage;"[9] even where the firm's power was primarily attributable to "superior skill, industry, and foresight,"[10] and the dominant firm neither sacrificed profits to gain its advantage nor intended to use the advantage to maintain or further increase its monopoly power.[11] During this phase of American antitrust, a monopolist defending a challenged practice (for instance, leasing to customers capital equipment rather than selling it) was required to demonstrate that the practice made no contribution to its market power, and that the monopolist's strength was attributable "solely to [its] ability, economies of scale, research, natural advantages, and adaptation to inevitable economic laws."[12]

Not only was this period a time of aggressive enforcement, it was a time during which antitrust law was held in the highest esteem and enjoyed a place in the American constellation of laws near to the Constitution itself. This is what the Supreme Court said about the importance of antitrust in 1972:"Antitrust laws in general, and the Sherman Act in particular, are the Magna Carta of free enterprise. They are as important to the preservation of economic freedom and our free-enterprise system as the Bill of Rights is to the protection of our fundamental personal freedoms."[13]

And in 1978, the importance of antitrust meant that

> ...[E]ven when Congress by subsequent legislation establishes a regulatory regime over an area of commercial activity, the antitrust laws will not be displaced unless it appears that the antitrust and regulatory provisions are plainly repugnant. The presumption against repeal by implication reflects the understanding that the antitrust laws establish overarching and fundamental policies, a principle which argues with equal force against implied exclusions.[14]

---

6 Exclusive dealing and tying practices are often challenged under other statutes requiring concerted action or agreement, but inasmuch as those practices are fundamentally unilateral, we treat them so for purposes of this discussion.

7 *United States v. United States Steel Corp.*, 251 U.S. 417 (1920); *United States v. American Can Co.*, 230 F. 859 (D. Md. 1916), appeal dismissed by 256 U.S. 706 (1921). The oil and tobacco trusts were found guilty of violating both § 1 of the Act, which prohibits unreasonable restraints of trade, and § 2. *See Standard Oil Co. v. United States*, 221 U.S. 1, 55, 61-62, 75 (1911) (finding the Standard Oil trust guilty of monopolization); *United States v. American Tobacco Co.*, 221 U.S. 106, 182-83 (1911) (finding monopolization). Other combinations, however, were found to violate only § 1. *Northern Sec. Co. v. United States*, 193 U.S. 197 (1904) (injunction under § 1 of the Sherman Act prohibiting the Harriman-Hill-Morgan railroad holding company from exercising control over competing railroads); *Swift & Co. v. United States*, 196 U.S. 375 (1905) (injunction under § 1 of the Sherman Act prohibiting the beef trust from collusive price fixing).

8 *United States v. Aluminum Co. of America*, 148 F.2d 416 (2d Cir. 1945) (Hand, J.); *American Tobacco Co. v. United States*, 328 U.S. 781 (1946) (endorsing *Aluminum Co. of America*); *United States v. Griffith*, 334 U.S. 100 (1948); *United States v. United Shoe Machinery Corp.*, 110 F. Supp. 295 (D. Mass. Feb. 18, 1953), *aff'd per curiam*, 347 U.S. 521 (1954); *United States v. Grinnell Corp.*, 384 U.S. 563 (1966).

9 *Griffith*, 334 U.S. at 107.

10 *United States v. Aluminum Co. of Am.*, 148 F.2d 416, 430 (2d Cir. 1945).

11 *Griffith*, 334 U.S. 100 (1948).

12 *United States v. United Shoe Machinery Corp.*, 110 F. Supp. 295, 343 (1953), *aff'd per curiam*, 347 U.S. 521 (1954).

13 *United States v. Topco Assocs.*, 405 U.S. 596, 610 (1972).

14 *Lafayette v. La. Power & Light Co.*, 435 U.S. 389, 398-99 (1978).

The Court's loss of esteem for antitrust over the next several decades is perhaps nowhere better captured than in the juxtaposition of these ringing endorsements of antitrust law's fundamental importance to economic regulation with the language from *Credit Suisse* reflecting a deep mistrust of antitrust and antitrust courts. See quoted *infra*, at § III.

II.

This period of heartfelt and aggressive § 2 jurisprudence came to an end in the early 1980s. Important cases in the late 1970s, while not involving § 2, represented a harbinger of change with respect to the entire canvass of antitrust, including § 2. *General Dynamics* in 1976 marked an important turning point in merger analysis – introducing a certain rigor into the analysis of market share.[15] Three years later in *GTE Sylvania,* the Supreme Court overruled the *per se* rule against vertical territorial restraints and established a return to a rule of reason analysis for evaluating non-price vertical restraints.[16] While these cases suggested the direction in which the Court was moving, the greatest changes in doctrine began with the appointment of William Baxter as the head of DOJ. He introduced to the broader antitrust community, and the bench: the Chicago School; efficiencies; empiricism; economics-based guidelines; amicus briefs to lower courts in an effort to shape the law at the bottom of the judicial pyramid, and more.

In the 1980s antitrust policy makers attempted to impose strict, relatively objective, principles designed not only to restrict the growth of antitrust as it was then known, but to attack many of its accepted features root and branch. For the most part the Supreme Court enthusiastically joined in, not only adopting limited antitrust doctrines but altering procedural approaches in ways that limited private enforcement.

The 1980s began with a bang. Four of the largest antitrust cases in history were concluded – three of them just abandoned: the *IBM* case challenging IBM's dominance in mainframe computers and peripherals; the FTC's so-called *cereal* case; and the FTC's *Exxon* case seeking to dismember the oil industry. The settlement of the fourth case, *AT&T*, resulted in the breakup of AT&T's monopoly on local telephone service, but hardly brought an end to antitrust issues in telecommunications. The end of these cases could be regarded as the end of the era of antitrust challenges to structural dominance. Henceforth, § 2 enforcement and doctrine would primarily focus on delineating the boundaries of specific competitive (or anticompetitive) practices, in particular below-cost pricing, bundled pricing, exclusive dealing, and the use of intellectual property. And the approach to these specific practices has been consistent with the Supreme Court's post-1980 distrust of antitrust law's role in governing aggressive competition by single firms.

In 1984, for example, the Supreme Court said that unilateral conduct, regardless of its effect on rivals, runs afoul of the antitrust laws ***only*** if it "threatens actual monopolization. It is not enough that a single firm appears to 'restrain trade' unreasonably, for even a vigorous competitor may leave that impression."[17] By 1993 the Court had largely repudiated the *Alcoa* line of cases (while not explicitly overruling any case) that had suggested a monopolist could violate § 2 by seizing business opportunities from its rivals, holding instead that § 2 of the Sherman Act

15 415 U.S. 486, 501-504 (1976) (In *General Dynamics* the Court held that coal company's huge past and present market share was unrevealing about the firm's future ability to compete where the company had few reserves in the ground, and thus could not effectively compete for future long term supply contracts).
16 *GTE Sylvania Inc. v. Continental T.V.*, 433 U.S. 36, 59 (1977).
17 *Copperweld Corp. v. Independence Tube Corp.,* 467 U.S. 752, 767 (1984) (citations omitted).

> directs itself not against conduct which is competitive, even severely so, but against conduct which unfairly tends to destroy competition itself. It does so not out of solicitude for private concerns but out of concern for the public interest. Thus, this Court and other courts have been careful to avoid constructions of § 2 which might chill competition, rather than foster it.[18]

In a separate case that same year the Court made clear that

> [e]ven an act of pure malice by one business competitor against another does not, without more, state a claim under the federal antitrust laws; those laws do not create a federal law of unfair competition or "purport to afford remedies for all torts committed by or against persons engaged in interstate commerce."[19]

III.

Now, fifteen years later and in the early 21st century, the Court has gone further. Antitrust is no longer seen by the Supreme Court as the Magna Carta of free enterprise; rather, it seems to be seen as something of a beast on the verge of out of control. In *Verizon Communications Inc. v. Law Offices of Curtis V. Trinko*, the Court indicated not only comfort with the existence of monopoly power, but fawning approval of the stimulating effects of monopoly profits:

> Firms may acquire monopoly power by establishing an infrastructure that renders them uniquely suited to serve their customers. Compelling such firms to share the source of their advantage is in some tension with the underlying purpose of antitrust law, since it may lessen the incentive for the monopolist, the rival, or both to invest in those economically beneficial facilities.[20]

Even more dramatic was the Court's decision and language in *Credit Suisse*, which expressed a deep suspicion that the vagaries of antitrust litigation could not be trusted to produce consistent results. What the Court said about antitrust, while in context limited to the regulated securities markets, doubtless resonates with all critics of antitrust, class actions, treble damages, the lack of contribution and the American antirust regime in general:

> [A]ntitrust plaintiffs may bring lawsuits throughout the Nation in dozens of different courts with different nonexpert judges and different nonexpert juries. In light of the nuanced nature of the evidentiary evaluations necessary to separate the permissible from the impermissible, it will prove difficult for those many different courts to reach consistent results. And, given the fact-related nature of many such evaluations, it will also prove difficult to assure that the different courts evaluate different fact patterns consistently. The result is an unusually high risk that different courts will evaluate similar fact patterns differently.

---

18 *Spectrum Sports Inc v. McQuillan,* 506 U.S. 447, 458 (1993) (citations omitted). To prevail on a claim of attempted monopolization, plaintiff must prove that the defendant has (1) "engaged in predatory or anticompetitive conduct with (2) a specific intent to monopolize and (3) a dangerous probability of achieving monopoly power." *Id.* at 456.

19 *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209, 225 (1993) (citation omitted).

20 *Verizon Communications Inc. v. Law Offices of Curtis V. Trinko*, 540 U.S. 398, 407-408 (2004).

> …[T]hese factors suggest that antitrust courts are likely to make unusually serious mistakes in this respect. And the threat of antitrust mistakes, i.e., results that stray outside the narrow bounds that plaintiffs seek to set, means that underwriters must act in ways that will avoid not simply conduct that the securities law forbids … but also a wide range of joint conduct that the securities law permits or encourages ….[21]

This remarkable distaste for antitrust is a very far cry from antitrust as the Magna Carta of free enterprise.

The Supreme Court's solicitude for defendants, including monopolists, is captured in a single statistic. Since its 1992 decision in *Eastman Kodak Co. v. Image Technical Services Inc.,*[22] the Court has issued 17 antitrust decisions. Many were and are important; some less so. But not a single one ruled in favor of the party that was the antitrust plaintiff.[23] All involved holdings that favored the defendant. But the Supreme Court's doctrine must be applied by the lower courts, thus it may not be the end of the story.

In contrast to the endorsement of aggressive competition by dominant firms in *Spectrum Sports, Brooke Group*, *Trinko*, and *linkLine,*[24] stand the applications of § 2 by the Supreme Court in *Kodak*[25] and by lower courts in *LePage's, Inc. v. 3M,*[26] *Conwood Co. v. United States Tobacco Co.*,[27] and to a lesser extent *United States v. Microsoft Corp.,*[28] *United States v. Dentsply Int'l, Inc.,*[29] and *Broadcom Corp. v. Qualcomm Inc.*[30] These cases have all articulated a somewhat more expansive role for antitrust and demonstrate that dominant firms engaging in exclusionary conduct sometimes do so at their peril. Perhaps most importantly, DOJ's recent § 2 guidance, and also the FTC's enforcement actions in *In the Matter of Rambus, Inc.*,[31] and *In the Matter of Negotiated Data Solutions*[32]seem also, and more recently, to exhibit resistance to the Supreme Court's broad exculpatory mood.

Below, we seek to categorize the areas where litigation seems to be being shut off, as well as those where it may not be.

A. *The Supreme Court's Safe Harbors for Aggressive Competition*

1. *Pricing*

Plainly persuaded by the procompetitive benefits of aggressive competition, the Court in *Brooke Group* and *Trinko* created safe, or at least snug, harbors for certain types of

---

21 *Credit Suisse Securities LLC v. Billing*, 551 U.S. 264, 282 (2007).
22 504 U.S. 451 (1992).
23 *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 113 S. Ct. 2578 (1993); *Spectrum Sports Inc v. McQuillan*, 506 U.S. 447, 458 (1993); *Brown v. Pro Football, Inc.*, 518 U.S. 231 (1996); *State Oil Co. v. Khan*, 522 U.S. 3 (1997); *NYNEX Corp. v. Discon, Inc.*, 525 U.S. 128 (1998); *California Dental Ass'n v. F.T.C.*, 526 U.S. 756 (1999); *Verizon Communications v. Law Offices of Curtis V. Trinko*, 540 U.S. 398 (2004); *United States Postal Service v. Flamingo Indus. (U.S.A.) Ltd.*, 540 U.S. 736 (2004); *F. Hoffman-la Roche Ltd. V. Empagran S.A.*, 542 U.S. 155 (2004); *Volvo Trucks North America, Inc. v. Reeder Simco GMC, Inc.*, 546 U.S. 164 (2006); *Texaco, Inc. v. Dagher*, 547 U.S. 1 (2006); *Illinois Tool Works Inc. v. Independent Ink, Inc.*, 547 U.S. 28 (2006); *Leegin Creative Leather Products v. PSKS, Inc.*, 126 S. Ct. 2705 (2007); *Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber*, 127 S. Ct. 1069 (2007*); Bell Atlantic v. Twombly*, 127 S. Ct. 1955 (2007); *Credit Suisse Securities v. Billing*, 127 S. Ct. 2383 (2007); *Pacific Bell Telephone Co. v. linkLine Communications, Inc.*, 129 S. Ct. 1109 (2009).
24 *Pacific Bell Telephone Co. v. linkLine Communications, Inc.*, 129 S. Ct. 1109 (2009).
25 504 U.S. 451 (1992).
26 324 F.3d 141 (3d Cir. 2003), *cert. denied,* 542 U.S. 953 (2004).
27 290 F.3d 768 (6th Cir. 2002), *cert. denied,* 537 U.S. 1148 (2003).
28 253 F.3d 34 (D.C. Cir.), *cert. denied,* 534 U.S. 952 (2001).
29 399 F.3d 181 (3d Cir. 2005), *cert. denied,* 126 S. Ct. 1023 (2006).
30 501 F.3d 297 (3d Cir. 2007).
31 Docket No. 9302 (Aug. 2, 2006), *available at* http://www.ftc.gov/os/adjpro/d9302/index.shtm.
32 *In the Matter of Negotiated Data Solutions LLC,* File No. 0510094 (2008), *available at* http://www.ftc.gov/os/caselist/0510094/index.shtm.

conduct. Specifically, after *Brooke Group,* dominant firms may aggressively discount individual products against small rivals without fear of antitrust liability, provided those prices remain above some measure of cost. In *Brooke Group,* plaintiff-Liggett claimed that Brown & Williamson "cut prices on generic cigarettes below cost and offered discriminatory volume rebates to wholesalers to force Liggett to raise its own generic cigarette prices and introduce oligopoly pricing in the economy [cigarette] segment."[33] Characterizing Liggett's claim as one of predatory pricing, the Court said that, to prevail, plaintiff must prove that defendant's prices were below an "appropriate level" of cost, and that defendant "had a reasonable prospect, or, under § 2 of the Sherman Act, a dangerous probability, of recouping its investment in below-cost prices."[34] Key to that holding was the principle that:

> "[L]ow prices benefit consumers regardless of how those prices are set, and so long as they are above predatory levels, they do not threaten competition . . . ." As a general rule, the exclusionary effect of prices above a relevant measure of cost either reflects the lower cost structure of the alleged predator, and so represents competition on the merits, or is beyond the practical ability of a judicial tribunal to control without courting intolerable risks of chilling legitimate price cutting.[35]

The Court subsequently expanded this analysis to predatory bidding situations in *Weyerhauser*,[36] rejecting a competitor's complaint that a dominant firm's unreasonably high bids for alder logs excluded competitors from the market.

2. *Refusals to Deal*

Trinko and *linkLine*, the Supreme Court's most recent decisions focused directly on single firm conduct, follow in the steps of *Brooke Group* and increase the confidence with which a dominant firm may refuse to aid a rival.[37] *Trinko* narrowly defined the circumstances under which a refusal to deal may be characterized as anticompetitive conduct and expressly limited the Court's earlier decision in *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.*[38] to its facts. *Aspen* appeared to raise the stakes for dominant firms that refused to deal with a rival, where the refusal represented an "important change in a pattern of distribution that had originated in a competitive market," and facts suggested that the decision to cut off the rival amounted to a sacrifice of short-run profits in order to reduce competition in the long run.[39] However, the impact of *Aspen* extended beyond its facts. It renewed, at least for a time, the importance of intent in discerning whether conduct was "fairly characterized as 'exclusionary' or 'anticompetitive'" defined "exclusionary"

---

33 *Brooke Group, Ltd.,* 509 U.S. at 212.

34 *Id.* at 222, 224. Liggett's predatory pricing claim arose under § 2(a) of the Robinson-Patman Act. The Court held, however, that the same standard applied whether the claim arose under the Robinson-Patman Act or the Sherman Act: "the essence of the claim under either statute is the same: A business rival has priced its products in an unfair manner with an object to eliminate or retard competition and thereby gain and exercise control over prices in the relevant market." *Id.* at 222.

35 *Id.* at 223 (citations omitted).

36 *Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber*, 127 S. Ct. 1069 (2007).

37 While not focused tightly on the standards for liability in a single firm conduct setting, the Court's three decisions from the 2005 Term also give defendants a degree of aid and comfort. First, in *Volvo Trucks N. Am. Inc. v. Reeder-Simco GMC Inc.,* 126 S. Ct. 860 (2006), the Court reversed an Eighth Circuit decision that had allowed a manufacturer offering its dealers different wholesale prices to be held liable for price discrimination proscribed by the Robinson-Patman Act in the absence of a showing that the manufacturer discriminated between dealers contemporaneously competing to resell to the same retail customer. Second, in *Texaco Inc. v. Dagher,* 126 S. Ct. 1276 (2006), the Court reversed a Ninth Circuit decision applying the per se rule to claims of price fixing among parties to a joint venture who effectively operated as a single entity competing with other sellers in the market. Third, in *Independent Ink Inc. v. Illinois Tool Works Inc.,* 126 S. Ct. 1281 (2006), the Court vacated a judgment against a patent holder for tying and monopolization that had been affirmed by the Federal Circuit, holding that ownership of a patent does not presumptively confer market power in tying cases.

38 472 U.S. 585 (1985).

39 *Id.* at 603.

conduct as any conduct that "attempt[s] to exclude rivals on some basis other than efficiency," and suggested that a court should consider the impact of the challenged conduct on consumers and "whether it has impaired competition in an unnecessarily restrictive way."[40] The latter tinged with antitrust risk competitive strategies by a dominant firm (whether a refusal to deal or other type of conduct) that were not "efficient" in the sense of lowering cost or improving quality and not the least restrictive alternative to achieve the firm's objectives.

Trinko ended lingering ambiguity about the duty to deal with rivals, declaring *Aspen* "at or near the outer boundary of § 2 liability,"[41] and drove the result in *linkLine*. In *linkLine*, independent Internet service providers that competed with AT&T in the retail DSL market, and also leased DSL transport service from AT&T at the wholesale level, argued that AT&T subjected them to a price squeeze in violation of § 2.[42] The Supreme Court rejected the possibility of a price squeeze as a cognizable antitrust offense, at least in the absence of an "antitrust duty to deal." "*Trinko* ... makes clear that if a firm has no antitrust duty to deal with its competitors at wholesale, it certainly has no duty to deal under terms and conditions that the rivals find commercially advantageous."

It is rarely the case, as in *Trinko* and *linkLine,* that an alleged monopolist can claim that its products or services have never been willingly sold to third parties, and rarer still that a monopolist is required by regulation to sell at cost to its downstream rivals. Nevertheless, and doubtless going farther than the facts before it required, the Court in both cases left no doubt that the duty to deal is no broader than that arising on the facts of *Aspen* where "[t]he unilateral termination of a voluntary (and thus presumably profitable) course of dealing suggested a willingness to forsake short-term profits to achieve an anticompetitive end."[43] After *Trinko* and *linkLine,* unless facts surrounding a unilateral refusal to deal can be squeezed into *Aspen,* the antitrust analysis may well end.[44]

3. *Regulated Industries*

Another situation in which the Supreme Court has limited the application of the antitrust laws involves immunizing certain conduct in regulated industries from antitrust scrutiny. In 2007, the Supreme Court held in *Credit Suisse*, mentioned above, that the antitrust laws could not be applied to a conspiracy among securities underwriters to inflate the commissions on initial public offerings, due to a "plain repugnancy" between the antitrust claims and the federal securities laws.[45] The Court held that although the Securities and Exchange Commission ("SEC") condemned the practices in question, application of the antitrust laws to the unlawful practices would "threaten[] serious securities-related harm" due to the likelihood that "dozens of different courts with different nonexpert judges and different nonexpert juries" would have difficulty reaching consistent results.[46]

---

40 *Id.* at 602, 605.
41 124 S. Ct. at 879 (citation omitted).
42 The plaintiff ultimately argued that it should be given the opportunity to prove a predatory pricing claim in the retail market, but the Court rejected this effort as well.
43 *Id.* at 879.
44 A patentee or copyright owner generally has an absolute right to refuse unilaterally to license patents or copyrights (or refuse to sell patented or copyrighted products) for any reason. *See, e.g., In re Indep. Serv. Orgs. Antitrust Litig.,* 203 F.3d 1322, 1329 (Fed. Cir. 2000), *cert. denied,* 531 U.S. 1143 (2001). In the First and Ninth Circuits, however, the refusal to license (or sell) only gives rise to a rebuttable presumption that the refusal is supported by a legitimate business reason and, in the Ninth Circuit, that presumption can be rebutted by *subjective intent* evidence. *See Image Tech. Servs., Inc. V. Eastman Kodak Co.,* 125 F. 3d 1195, 1218 (9th Cir. 1997); *Data Gen. Corp. v. Grumman Sys. Support Corp.,* 36 F.3d 1147, 1187 (1st Cir. 1994); *see also Microsoft,* 253 F.3d at 63 ("frivolous" to argue "absolute and unfettered right" to use one's own intellectual property as one wishes). It remains to be seen whether the use of subjective intent evidence, without satisfying the factual predicate of *Aspen,* can stand after *Trinko.*
45 *Credit Suisse Securities LLC v. Billing*, 551 U.S. 264 (2007).
46 *Id.* at 280.

In the Court's view, while the risk of inconsistent results is present in all antitrust lawsuits, the difficulty in "separating the permissible from the impermissible" in the securities context meant that "there is no practical way to confine antitrust suits so that they challenge only activity ... unlawful under the securities law."[47] Thus significant risk existed that underwriters would be forced to "act in ways that will avoid not simply conduct that the securities law forbids ... but also a wide range of [efficiency-enhancing] joint conduct that the securities law permits or encourages."[48] Further, because the SEC "actively enforces the rules and regulations that forbid the conduct in question," the "enforcement-related need for ... antitrust ... " was held to be "unusually small."[49]

Credit Suisse went significantly further than any previous case in holding that the antitrust laws could not be applied to a conduct deemed illegal by a regulator due to the potential that hypothetical "nonexpert judges" and "nonexpert juries" would reach erroneous conclusions in future cases. It remains to be seen whether dominant firms will be able to avail themselves of this principle in other regulated industries (power generation and airlines spring immediately to mind), but given the tenor of the Court's decision, one might well wonder whether the *Credit Suisse* holding will remain limited to the particularly intricate securities industry and the SEC.

4. *Pleading Rules*

Perhaps even more important than immunizing from antitrust scrutiny specific practices or conduct in certain regulated industries, the Supreme Court has also, in the last three terms, erected more onerous pleading requirements for plaintiffs that may render it difficult for many § 2 plaintiffs to survive motions to dismiss. In *Bell Atlantic Corp. v. Twombly*,[50] the Supreme Court abandoned its 50-year-old precedent governing when a complaint states a claim under the Federal Rules of Civil Procedure. *Conley v. Gibson*[51] had for five decades instructed courts that "a complaint should not be dismissed for failure to state a claim unless it appears beyond doubt that the plaintiff can prove no set of facts in support of his claim which would entitle him to relief."[52] *Twombly* held that this standard had "earned its retirement," and that henceforth in antitrust cases alleging conspiracy, in order to survive a motion to dismiss a claim "requires a complaint with enough factual matter (taken as true) to suggest that an agreement was made."[53]

Some lower courts resisted expanding *Twombly* beyond the antitrust conspiracy context, leading the Supreme Court, in *Ashcroft v. Iqbal*, to expand *Twombly* and make clear that a complaint "does not unlock the doors of discovery for a plaintiff armed with nothing more than conclusions," and that all civil complaints "must contain sufficient factual matter ... to 'state a claim to relief that is plausible on its face.'"[54] In identifying the "factual matter," *Iqbal* instructs courts considering motions to dismiss "to begin by identifying pleadings that, because they are no more than conclusions, are not entitled to the assumption of truth."[55] In *Twombly*, the "legal conclusions" not entitled to the assumption of truth included the allegation that the defendants formed a conspiracy, while in *Iqbal*, they included the allegation that the defendants "knew of, condoned, and willfully and

47 *Id.* at 282.
48 *Id.*
49 *Id.*
50 550 U.S. 544 (2007).
51 355 U.S. 41 (1957).
52 *Id.* at 45.
53 550 U.S. at 557.
54 *Ashcroft v. Iqbal*, 129 S. Ct. 1937, 1950 (2009).
55 *Id.* at 1950.

maliciously agreed to subject [him]" to harsh conditions of confinement "as a matter of policy, solely on account of [his] religion, race, and/or national origin and for no legitimate penological interest."[56]

Under the *Twombly/Iqbal* approach, § 2 plaintiffs will face particular difficulty alleging the specific intent required to support a claim of attempt to monopolize, as well alleging that a practice such as a monopolist's refusal to deal are not justified or undertaken for the purpose of excluding a competitor. The assertion that a firm possesses a particular intent, and other fundamentally factual assertions may come to be treated as a "legal conclusion"[57] – indeed, it is difficult to picture circumstances in which a dominant firm would be so careless as to allow a potential plaintiff to acquire facts that would allow specific intent to be pled "plausibly."

B. *Lower Courts Accept Broader Theories of Liability*

The Supreme Court's last opinion opening the door to more, not less, risk for dominant firms was its 1992 decision in *Kodak,* which created antitrust risks for aggressive aftermarket competitors even if they lacked market power in the foremarket. *Kodak* precipitated a rush of claims challenging the aftermarket practices of manufacturers and franchisors,[58] but for years *Kodak's* seemingly expansive theory failed to take root, as lower courts' strict application of the conditions under which a small rival could become an aftermarket monopolist led to the rejection of the vast majority of post-*Kodak* claims.[59] Yet the Ninth Circuit relatively recently clarified the circumstances in which such a claim may lie, consistent with the willingness of plaintiffs to pursue and lower courts to allow claims not entirely foreclosed by the Supreme Court. Recent district courts have been more receptive to such aftermarket § 2 claims than was initially the case.[60]

In *Newcal Industries, Inc. v. IKON Office Solutions*,[61] the Ninth Circuit held that lack of market power in a primary market does not preclude an antitrust claim in an aftermarket where consumers make separate decisions to purchase in the primary market and the aftermarket, as long as customers do not explicitly contract away their ability to take advantage of competition in the aftermarket.[62] *Newcal* represents an interesting elaboration of *Kodak* for several reasons. First, *Newcal* suggests that a § 2 claim can be based on conduct that does not violate the reasonable expectations of customers at the time the primary good was purchased, as long as the conduct denied the customers the benefits of

---

56 *Id.* at 1951.

57 *Ginsburg et al. v. InBev NV/SA et al.*, No. 08-cv-1375 (E.D. Mo. Aug. 3, 2009) at 5 (granting defendants' motion for judgment on the pleadings and dismissing complaint with prejudice because allegations that beer makers were influenced by the possibility that defendant might enter the United States market were "legal conclusions."); *see also In Re Travel Agent Commission Antitrust Litigation*, No. 07-4464 (6th Cir. Oct. 2, 2009).

58 *See generally* John DeQ. Briggs and James G. Kress, *Trends in Private Antitrust Litigation: The Monopolist Next Door,* The Antitrust Review of the Americas 2003, Global Competition Review.

59 *See, e.g., Queen City Pizza, Inc. v. Domino's Pizza,* 124 F.3d 430 (3d Cir. 1997), *cert. denied,* 523 US. 1059 (1998); *PSI Repair Services Inc. v. Honeywell, Inc.,* 104 F.3d 811 (6th Cir.) *cert. denied,* 520 U.S. 1265 (1997); *Digital Equipment Corp. v. Uniq Digital Technologies, Inc.,* 73 F.3d 756 (7th Cir. 1996); *10 Security Systems Canada, Inc. v. Checkpoint Systems, Inc.,* 249 F. Supp. 2d 622 (E.D. Pa. 2003); *Universal Avionics Sys. Corp. v. Rockwell Int'l Corp.,* 184 F. Supp. 2d 947 (D. Ariz. 2001).

60 *E.g., Alternative Electrodes LLC v. Empi, Inc.,* 597 F. Supp. 2d 322 (E.D.N.Y. 2009) (denying motion to dismiss antitrust claim alleging monopolization of market for replacement electrodes for use with medical device sold by defendants); *Helicopter Transport Services, Inc. v. Erickson Air-Crane Inc.*, 2008 U.S. Dist. LEXIS 3466 (D. Or. Jan. 14, 2008) (denying defendant's motion for summary judgment on antitrust claim alleging defendant monopolized the market for helicopter spare parts and leveraged its parts monopoly into the heavy helicopter services market); *Xerox Corp. v. Media Sciences Int'l,* 511 F. Supp. 2d 372 (S.D.N.Y. 2007) (denying motion to dismiss antitrust counterclaim alleging monopolization or attempt to monopolize the market for replacement solid ink sticks for use in Xerox's phase change color printers); *Fin. & Sec. Prods. Ass'n v. Diebold,* 2005 U.S. Dist. LEXIS 45409 (N.D. Cal. July 8, 2005) (denying preliminary injunction to plaintiff alleging monopolization by Diebold of a parts and service aftermarket relating to ATMs).

61 513 F.3d 1038 (9th Cir. 2008).

62 *Id.* at 1050 ("Just as the plaintiffs had in *Eastman Kodak,* Newcal offers factual allegations to rebut the economic presumption that IKON consumers make a knowing choice to restrict their aftermarket options when they decide in the initial (competitive) market to enter an IKON contract.").

competition. Unlike some *Kodak*-type cases, *Newcal* does not mention any requirement that a policy change violates the expectations of locked-in customers – the allegations in *Newcal* were all about the exclusionary conduct and its impact on competitors in the aftermarket.

Second, *Newcal* recognized that a number of post-*Kodak* cases held that contractual rights did not give rise to market power,[63] but distinguished them because

> [t]his case is not a case in which the alleged market power flows from contractual exclusivity. IKON is not simply enforcing a contractual provision that gives it the exclusive right to provide replacement equipment and lease-end services. Rather, it is leveraging a special relationship with its contracting partners to restrain trade in a wholly derivative aftermarket.[64]

Thus, under *Newcal*, a firm's contractual relationship with its customers can create a "special relationship" that can in turn be leveraged to the detriment of competitors, even absent a policy change violating the expectations of locked-in firms.

Similarly, in contrast to the pro-defendant outcomes in *Brooke Group* and *Trinko*, stand Circuit court decisions in *LePage's, Conwood* and *Microsoft,* affirming adverse verdicts against dominant firms because of aggressive competitive or distribution strategies, as well as *Dentsply,* which reversed the trial court's decision and found § 2 liability in a dominant firm's exclusive dealing policies. Of these, *Microsoft* and *Dentsply* are closest to the Supreme Court's monopolization law, but even those opinions reflect important shifts from center. *Microsoft* could be viewed, in part, as repackaging "mainstream" standards for market definition and suppression of potential competition. But its willingness to find a § 2 violation based on exclusive dealing agreements that did not violate § 1 of the Sherman Act is a dramatic departure from antitrust norms, foreshadowed in *Conwood* and followed in *LePage's.* The Third Circuit in *Dentsply* followed a similar path, finding a violation of § 2 even where there were unappealed findings of no violation of § 3 of the Clayton Act or § 1 of the Sherman Act. Likewise, *Microsoft's* treatment of legitimate business justifications as a balancing exercise consistent with rule of reason analysis under § 1- in which anticompetitive effects are weighed against any offsetting procompetitive benefits - further dilutes the role of such justifications in a § 2 case. In the past, there was at least the argument that the existence of a demonstrable procompetitive business justification served as a shield to § 2 liability.[65] *Microsoft* may have closed that door.

The most serious ambiguity in modern American law bearing upon unilateral conduct may well be the lack of a coherent standard governing "exclusionary" or "anticompetitive" conduct, as reflected in the controversy surrounding the outcomes in *LePage's* and *Conwood.* In particular, dominant firms are faced with uncertainty by the emergence of theories that permit the imposition of § 2 Sherman Act liability for allegedly exclusionary agreements that do not violate § 1. The Supreme Court had three chances in 2004 and 2005 to provide guidance regarding the appropriate standard for exclusionary conduct: first, in response to 3M's petition for *certiorari* in *LePage's;* second, in its decision in *Trinko;* and third, in response to Dentsply's petition for certiorari in *Dentsply.* The Court declined to consider *LePage's* and *Dentsply,* and did not purport to define a standard in *Trinko* that would apply generally to exclusionary conduct. As a consequence, counselors

63 *Id.* at 1048 ("the law prohibits an antitrust claimant from resting on market *power* that arises solely from contractual rights that consumers knowingly and voluntarily gave to the defendant (as in *Queen City Pizza* and *Forsyth*).").
64 *Id.* at 1050.
65 *See Kodak,* 504 U.S. at 483; *ACT, Inc. v. Sylvan Learning Sys. Inc.,* 296 F.3d 657, 669-70 (8th Cir. 2002).

continue to struggle with a patchwork of different standards in the circuit courts and the vestiges of *Aspen's* "efficiency" based definition of "exclusionary" conduct.[66]

1. *Conwood*

Depending on your point of view, *Conwood* is either a refreshing throwback to the days in which a monopolist was punished for engaging in dirty tricks, or a cautionary tale of unstructured § 2 analysis. No matter your view, the size of the *Conwood* judgment, amounting to trebled damages in excess of USD ***$l billion,*** signals the significant risk to defendants posed by § 2 challenges from rivals losing ground to aggressive merchandising strategies for more popular brands. In *Conwood,* a rival snuff manufacturer complained that United States Tobacco ("UST") "engaged in a concerted effort, directed from the highest levels of a national monopoly, to shut Conwood out from effective competition through the elimination of its racks and [point of sale] advertising, all in the unusual moist snuff market, where [point of sale] is the central marketplace battleground."[67] The jury awarded Conwood $350 million on its § 2 claim, trebled to $1.05 billion. On appeal, the Sixth Circuit affirmed, pointing to UST's misuse of its role as category captain, unauthorized destruction of Conwood racks, burying competitive products on the UST industry rack and misrepresenting its sales performance to increase facings of slower moving UST products.

The decision of the Sixth Circuit is remarkable because of its emphasis on intent, reliance on internal UST documents, dismissal of testimony by retailers that they (not UST) controlled in-store placements, and apparent refusal to assess the extent to which UST's conduct actually foreclosed Conwood and other rivals from reaching consumers. The appellate court was, instead, satisfied that the allegedly exclusionary conduct undertaken by a "conceded monopolist" was widespread, unjustified and driven by anticompetitive intent, and (based largely on testimony by plaintiff's expert) harmed consumers by raising prices, limiting choice, and slowing the growth of rivals.[68] It distinguished, and rejected as irrelevant, established foreclosure analysis of exclusive dealing under § 1.[69] However, the record suggested that less than 10 percent of stores used UST racks exclusively, and evidence of "widespread" destruction of racks was anecdotal.[70] *Conwood* supplies no rules of general applicability beyond the notion that a collection of torts can at times amount to a §2 case and thus tort lead in the hands of a skilled alchemist can be at times converted to antitrust gold.

2. *LePage's*, *Cascade Health Solutions v. PeaceHealth*,[71] and *Doe v. Abbott Labs*[72]

The Third Circuit's *en banc* decision in *LePage's, Inc. v. 3M*[73] illustrates the manner in which lower courts have expanded § 2 liability in a discernible counter-trend to the

66 *See, e.g., Microsoft,* 253 F.3d at 58-59 (setting forth framework to prove anticompetitive conduct that requires proof of anticompetitive effects and permits balancing of procompetitive benefits); *Taylor Publishing Co. v. fastens, Inc.,* 216 F.3d 465, 475 (5th Cir. 2000) ("'exclusionary conduct' is conduct, other than competition on the merits or restraints reasonably 'necessary' to competition on the merits, that reasonably appears capable of making a significant contribution to creating or maintaining monopoly power.") (citation omitted); *Gen. Indus. Corp. v. Hartz Mountain Corp.,* 810 F.2d 795, 804 (8th Cir. 1987) ("exclusionary" conduct for § 2 purposes is conduct "without legitimate business purpose that makes sense only because it eliminates competition"). *Compare also Concord Boat Corp. v. Brunswick Corp.,* 207 F.3d 1039, 1062-63 (8th Cir. 2000) (ordinary business practices cannot serve as anticompetitive conduct for § 2 purposes), *cert. denied,* 531 U.S. 979 (2000), *with LePage's,* 324 F.3d at 151-52 ("monopolist is not free to take certain actions that a company in a competitive (or even oligopolistic) market may take").

67 290 F.3d at 787.

68 *Id.* at 784-90.

69 *Id.* at 787 n.4.

70 *Id.* at 775, 784-85.

71 515 F.3d 883 (9th Cir. 2008).

72 571 F.3d 930 (9th Cir. 2009).

73 324 F.3d 141 (3d Cir. 2003), *cert. denied,* 124 S. Ct. 2932 (2004).

Supreme Court, although admittedly many of the Supreme Court's pronouncements postdate *LePage's*. However, the related decisions of the Ninth Circuit in *PeaceHealth* and *Abbott Laboratories* illustrate that even more limited interpretations of § 2 do not resolve the tension between *LePage's* and the Supreme Court's § 2 jurisprudence, and that lower court efforts to police monopolists' pricing may not survive the Supreme Court's retrenchments.

LePage's affirmed an approximately $68 million trebled damage award stemming from 3M's practice of offering certain bundled and incentive discounts. The holding that such arguably ordinary discounting practices amounted to illegal monopolization serves as a further example of juries and lower courts' willingness to sanction dominant firms despite the Supreme Court's more permissive rhetoric. *LePage's,* however, is even more notable because the conduct condemned in *LePage's* is strikingly similar to conduct the prospect of which led the EU Commission in the summer of 2001 to block the proposed merger of General Electric Company and Honeywell International, Inc.[74] - a transaction cleared by its U.S. counterpart subject to minor conditions.[75] That action precipitated an unusual and highly public outcry from U.S. regulators to the effect that the EU action was contrary to fundamental antitrust principles.[76] Then-Assistant Attorney General Charles A. James, in written remarks, described the divergent outcomes:

> We concluded that the merged firm would have offered improved products at more attractive prices than either firm could have offered on its own, and that the merged firm's competitors would then have had a great incentive to improve their own product offerings. This, to us, is the very essence of competition, and no principle is more central to U.S. law than that antitrust protects competition, not competitors.
>
> In stark contrast, the EC focused on how the merger would affect European and U.S. competitors, essentially concluding that the very efficiencies and lower prices the transaction would produce would be anticompetitive because they might ultimately drive some of those competitors from the market or reduce their market shares to a point where they could not [sic] longer compete effectively. In other words, the EC determined that the fact that customers would be "induced" to purchase more attractive and lower-priced GE/Honeywell products, rather than those of its competitors, was a bad thing of a sort that its antitrust law ought to prohibit.[77]

---

74 Commission Decision of July 3, 2001, Case No COMP/M.2220 - *General Electric/Honeywell.* Appeals are currently pending at the Court of First Instance in Luxembourg, Case T-209/01 *Honeywell v. Commission* and Case T-21O/01 *GE v. Commission.*

75 In the view of the European Commission, the proposed transaction would have combined GE's position in engines, described by the Commission as "dominant," and GE's influence as a purchaser and financier of aircraft through GE Capital Aviation Services and GE Credit, with Honeywell's leading position in avionics and other products. The Commission was concerned that the proposed merger would have permitted the merged firm to strengthen its position in engines and achieve dominance in avionics and other products through bundled package deals. In particular, the merged firm would allegedly have had an incentive to bundle engines with avionics (and other products such as auxiliary power units, environmental control systems, electric power, wheels and brakes, landing gear, and aircraft lighting) in sales to aircraft manufacturers and airlines, to gain an advantage over its competitors.

76 Charles A. James, Assistant Attorney General, Antitrust Division, USDOJ, International Antitrust in the 21st Century: Cooperation and Convergence, Address Before the OECD Global Forum on Competition, Paris, France (Oct. 17, 2001); *see* Charles A. James, Assistant Attorney General, Antitrust Division, USDOJ, Reconciling Divergent Enforcement Policies: Where Do We Go From Here? Address Before the Fordham Corporate Law Institute, New York, N.Y. (Oct. 25, 2001); William J. Kolasky, Deputy Assistant Attorney General, Antitrust Division, USDOJ, Conglomerate Mergers and Range Effects: It's a Long Way from Chicago to Brussels, Address Before the George Mason University Symposium, Washington, D.C. (Nov. 9, 2001); *see also* William J. Kolasky & Leon B. Greenfield, *A View to a Kill: The Lost GE/Honeywell Deal Reveals a Trans-Atlantic Clash of Essentials,* Legal Times, (July 30, 2001) at 28. *See generally* James F. Rill & John DeQ. Briggs, *GE-Honeywell: Chill or Challenge for Global Cooperation?* Antitrust Rep. 3 (Sept. 2001); John DeQ. Briggs & Howard Rosenblatt, *A Bundle of Trouble: The Aftermath of GE/Honeywell,* Antitrust Magazine (Fall 2001); John DeQ. Briggs & Howard Rosenblatt, *Live and Let Die,* 10 Geo. Mason L. Rev. 3 (Apr. 2002).

77 Charles A. James, *supra,* International Antitrust in the 21st Century: Cooperation and Convergence.

That characterization of the EU Commission's reasoning, whether or not correct in *GE,* lies at the very heart of the *LePage's* decision.

It is curious that the bundling theories pursued by the EC in *GE/Honeywell,* and so vigorously attacked by senior U.S. government officials, have found fertile soil in the Third Circuit. Even more surprising, in the circumstances, is the decision of the Antitrust Division of the United States Department of Justice not to support 3M's petition for review by the U.S. Supreme Court.[78]

Until *LePage's,* American rules governing discounting and bundling by single firms were believed to be relatively clear, and were fairly reflected in the U.S. reaction to *GE/Honeywell. First,* as a general matter, low prices were hailed as the essence of competition.[79] Whether offered by a dominant firm and "regardless of how ... set," they did not raise the specter of antitrust liability under § 2 unless the price fell below an "appropriate measure of cost," and defendant had a "dangerous probability of recouping its investment in below-cost prices."[80] That standard, set out in *Brooke Group,* ended - so many thought - the debate concerning the circumstances under which discount strategies, without more, could serve as a predicate for § 2 liability. Under that test, the distinction between volume discounts for single products or across product lines attracted little attention.[81] *Second,* rivals generally did not have antitrust standing to complain about increased competition from a competitor's low, above-cost prices.[82] *Next,* package discounts that simply offered two separately priced products at a discount were not suspect,[83] and volume discounts were viewed as procompetitive.[84] Indeed, the Robinson-Patman Act, prohibiting certain price discriminations, was typically the only legal constraint in framing the latter. Bundled discounts could raise issues in some courts if a seller offered two products at a discount, had market power in the market for one of those products and set the stand-alone price of the monopoly product at such a *high* level that, when added to the cost of purchasing the second product from an alternate supplier, the discounted bundle was the customer's only viable option. Such "offers" left the customer virtually no choice, prompting those courts to treat the practice, if proven, as a coercive *de facto* tie.[85] *Finally,* whether price incentives that encourage customers to shift purchases to the discounter might rise to the level of an exclusive dealing arrangement was an open question, but, as noted above, the prospect that a § 2 claim predicated on exclusive dealing could survive if the underlying conduct was lawful under § 1 was considered unlikely.

Not until *Concord Boat Corp. v. Brunswick Corp.*[86] – a case wending its way through the Eighth Circuit in the late 1990s – did counselors raise § 2 concern about non-coercive above-cost price reductions that encouraged customers to increase purchases at the expense of rivals. The district court in that case sustained a trebled damages award of approximately $133 million against an engine manufacturer for certain antitrust violations including, in particular, its use of above-cost volume and market share discounts to increase engine sales. On appeal, the Eighth Circuit rejected out of hand the notion that such

78 Although the brief of the government to the Supreme Court seemed clearly critical of the decision of the Third Circuit, that criticism was not for the reasons of policy more fervently enunciated in *GE/Honeywell.* Indeed, it seems entirely plausible that the decision of the government to oppose *certiorari* may have sprung from a fear, based on the record in *LePage's,* that the Court might well have affirmed the judgment.

79 *Barry Wright Corp. v. ITT Grinnell Corp.,* 724 F.2d 227 (1st Cir. 1983) (Breyer, J.).

80 *Brooke Group Ltd.,* 509 U.S. at 223-24.

81 *But cf Ortho Diagnostics Systems, Inc. v. Abbott Laboratories, Inc.,* 920 F. Supp. 455, 471 (S.D.N.Y. 1996) (§ 2 claims based on package discounts failed).

82 *Atlantic Richfield Co. v. USA Petroleum Co.,* 495 U.S. 328, 340-41 (1990).

83 *Jefferson Parish Hospital District No.2 v. Hyde,* 466 U.S. 2, 12 (1984).

84 *Fedway Associates, Inc. v. United States Treasury,* 976 F.2d 1416, 1422, 1423 (D.C. Cir. 1992) (Ginsburg, J.).

85 *See, e.g., Marts v. Xerox, Inc.,* 77 F.3d 1109, 1113 (8th Cir. 1996).

86 21 F. Supp. 2d 923 (E.D. Ark 1998), *rev'd,* 207 F.3d 1039 (8th Cir.), *cert. denied,* 531 U.S. 979 (2000).

discounts could violate § 2 of the Sherman Act.[87] The court distinguished, without much analysis, the bundled multi-product discounts then under attack in *LePage's.* However, early signs from other circuits suggested that the successful attack on such strategies, whether single-product or bundled discounts, would be short-lived.[88] That, of course, proved untrue, which brings us most directly to the facts of *LePage's* itself.

The *LePage's* facts are fairly straightforward. 3M manufactures Scotch® brand tape for home and office use and, in the early 1990s, had a share of about 90 percent in a market for transparent tape. LePage's competed against 3M with "second brand" and private label tape and, by 1992, accounted for some 88 percent of private label tape sales in the United States, although private label tape accounted for a relatively small percentage of overall transparent tape sales.

LePage's claimed that 3M improperly maintained its monopoly in transparent tape by offering higher rebates to customers for purchasing products across 3M's different product lines from home care and leisure products to audio/visual and stationery products (the bundled discounts), and by offering certain large customers lump-sum cash payments, promotional allowances and other cash incentives to encourage them to purchase 3M tape (allegedly *de facto* exclusive dealing arrangements). According to the court, the multi-product rebate program

> set customer-specific target growth rates in each product line. The size of the rebate was linked to the number of product lines in which targets were met, and the number of targets met by the buyer determined the rebate it would receive on all of its purchases. If a customer failed to meet the target for any one product, its failure would cause it to lose the rebate across the line. This created a substantial incentive for each customer to meet the targets across all product lines to maximize its rebates. . . . LePage's claim[ed] that customers could not meet these growth targets without eliminating it as a supplier of transparent tape.[89]

The Third Circuit agreed that 3M's bundled and incentive discounts - ordinary business practices in the hands of smaller rivals - were anticompetitive conduct in the hands of an alleged monopolist and, together, caused anticompetitive effects. Notably, however, its analysis was largely limited to the effects on LePage's and *assumed,* in particular, that LePage's would have to absorb the total bundled discount on its smaller volume of tape sales.[90] The court chose to ignore the possibility that LePage's was not as efficient a tape producer as 3M, the availability of other competitive responses such as joint marketing to spread the bundled discount over multiple firms, and the role of power buyers. It undertook scant analysis of the disputed effects of the conduct on tape prices and no analysis of output, which some evidence suggested had increased. It also rejected 3M's evidence of legitimate efficiencies, rejecting the arguments that the challenged discounts

87 207 F.3d 1039 (8th Cir.), *cert. denied,* 531 U.S. 979 (2000).

88 *See e.g., Virgin Atlantic Airways Ltd. v. British Airways Plc,* 257 F.3d 256,265-72 (2d Cir. 2001) (§ 2 claim based on bundled sales of tickets tested under *Brooke;* noting in connection with § 1 analysis that "[r]ewarding customer loyalty promotes competition on the merits"); *Microsoft,* 253 F.3d at 68 ("The rare case of price predation aside, the antitrust laws do not condemn even a monopolist for offering its product at an attractive price."); *Western Parcel Express v. UPS of America, Inc.,* 190 F.3d 974, 976 (9th Cir. 1999) (rejecting argument that volume discounts were tantamount to exclusive dealing agreements). *But see Avery Dennison Corp. v. ACCO Brands, Inc.,* 2000-1 Trade Cas. (CCH) 1[72,882, at 87,559-60 (CD. Cal. 2000) (permitting § 2 challenge to exclusivity payments, bundled rebates and other promotional payments); *In re Warfarin Sodium Antitrust Litigation,* 1999-1 Trade Cas. (CCH) 'If 72,457, at 84,219 (D. Del. 1998) (permitting § 2 challenge based on defendant's use of, among other things, "rebates and market retention agreements as part of its allegedly multifaceted effort to restrain trade in the oral anticoagulant market").

89 324 F.3d at 154, 170.

90 *ld.* at 159-63.

were consistent with legitimate economic interests to increase sales and further rejecting proffered benefits in the form of single invoices and consolidated shipments for lack of narrowly tailored cost-justification evidence.[91]

Relying on *Brooke Group,* 3M argued that its conduct was lawful because its tape prices were above cost. Declaring that "the most significant legal issue in this case," the Third Circuit dismissed *Brooke Group* as "[in]applicable to a monopolist with its unconstrained market power."[92] Looking to more general standards for exclusionary conduct reached in refusal to deal (not pricing) cases, the majority concluded that bundled rebates - even if above-cost - could be exclusionary for § 2 purposes.[93] In its view, the "principal anticompetitive effect of bundled rebates as offered by 3M is that when offered by a monopolist they may foreclose portions of the market to a potential competitor who does not manufacture an equally diverse group of products and who therefore cannot make a comparable offer," where the branded product is "indispensable to any retailer in the transparent tape market."[94]

At the urging of the Department of Justice, the Supreme Court denied 3M's petition for *certiorari,* leaving *LePage's* the rule at least in the Third Circuit. Companies are not free to ignore *LePage's*, because most large businesses in the United States are amenable to suit in the Third Circuit, which encompasses an economically significant region in the mid-Atlantic area of the United Sates' East Coast. However, the Ninth Circuit's subsequent decision in *Cascade Health Solutions v. PeaceHealth*[95] provides a rather different analysis of bundled discounts offered by dominant firms, and another recent case suggests, at least obliquely, that the Supreme Court's *linkLine* decision might ultimately eliminate all such challenges.

In *PeaceHealth*, the Ninth Circuit addressed a straightforward challenge to a bundled pricing offer. The defendant hospital operator offered both (1) tertiary care services[96] and (2) primary and secondary care services.[97] A competing hospital operator, offering only primary and secondary care services, charged that PeaceHealth monopolized and attempted to monopolize the relevant market for primary and secondary acute care hospital services by offering more favorable rates on tertiary services to purchasers (insurance companies) that made PeaceHealth their sole preferred provider for *all* services-primary, secondary, and tertiary. The district court, relying on *LePage's*, instructed the jury that "a defendant with monopoly power (or, in the case of an attempted monopolization claim, a dangerous probability of achieving monopoly power) engaged in exclusionary conduct by simply offering a bundled discount that its competitor could not match. The instruction did not require the jury to consider whether the defendant priced below cost."[98]

The Ninth Circuit reversed, based on *Brooke Group* and *Weyerhauser*, holding that a plaintiff must allege pricing below cost to allege exclusionary conduct.[99] To determine the appropriate measure of cost, the Ninth Circuit adopted the "discount attribution" standard:

91 *Id.* at 163-64.
92 *Id.* at 147, 151.
93 *Id.* at 146-52, 154-57.
94 *Id.* at 155, 156.
95 515 F.3d 883 (9th Cir. 2008).
96 Tertiary care includes more complex services such as invasive cardiovascular surgery and intensive neonatal care.
97 Primary and secondary acute care hospital services are common medical services such as setting a broken bone and performing a tonsillectomy.
98 *Id.* at 898-99.
99 *Id.* at 903.

> Under this standard, the full amount of the discounts given by the defendant on the bundle are allocated to the competitive product or products [primary and secondary care]. If the resulting price of the competitive product or products is below the defendant's incremental cost to produce them, the trier of fact may find that the bundled discount is exclusionary for the purpose of § 2.[100]

But while the discount attribution standard is far more rigorous than the Third Circuit's *LePage's* approach, it is obviously more likely to result in liability than a safe harbor based on *linkLine*. *PeaceHealth* and *linkLine* differ in that *PeaceHealth* did not involve the dominant firm providing an input directly to its competitors, while in *linkLine* AT&T squeezed its retail competitor on a good that AT&T provided at wholesale. But there is not a tremendous amount of analytical space between the Supreme Court's insistence that dominant firms be able to deal with whomever they choose and the *Brooke Group/Weyerhauser* doctrine that firms should generally be able to price however they choose. And as the Ninth Circuit demonstrated in *Abbott Laboratories*, to which we now turn, *linkLine* can be read entirely to insulate monopolists from liability based on the interplay between the doctrine of bundled discounts and the doctrine of refusal to deal.

In *Abbott Laboratories*, the defendant was allegedly a monopolist in a drug known as ritonavir, sold under a brand name as Norvir, that "boosts" the effectiveness of protease inhibitors used to fight HIV. Abbott originally sold Norvir as a standalone protease inhibitor, but later discovered it was more useful as a "booster" taken in low dosages along with other inhibitors. Abbott sold such a "boosted" protease inhibitor, Kaletra. Once Abbott's competitors received FDA approval to advertise that their protease inhibitors could be "boosted" by taking them in conjunction with Norvir, Abbott more than quadrupled the price of Norvir from $1.71 to $8.57 per 100 mg, but did not increase the price of Kaletra. According to the plaintiffs, the effect was to raise the total cost of boosted protease inhibitors provided by Abbott's competitors, and leverage its "Norvir monopoly to attempt to monopolize the boosted market for Kaletra."[101]

The district court denied Abbott's motions to dismiss, holding that *PeaceHealth* did not foreclose the plaintiffs' claims because the characteristics of the prescription drug market were not appropriate for the discount allocation approach. In particular, the insignificant marginal costs of manufacturing drugs compared to the tremendous research expenditures necessary to invent the drugs allows a discounter to offer a price that discourages competitive entry without ever slipping below marginal cost. The Ninth Circuit did not question the district court's reasoning on this point, but reversed in any event. According to the Ninth Circuit, "[t]ime, and the United States Supreme Court, have overtaken this case," and in light of *linkLine*, "allegations of monopoly leveraging through pricing conduct in two markets [do not] state a claim under § 2 of the Sherman Act, absent an antitrust refusal to deal (or some other exclusionary practice) in the monopoly market or below-cost pricing in the second market."[102]

The Ninth Circuit thus read *linkLine* to preclude antitrust challenges to ***all*** pricing – including bundled pricing – of a monopoly product. This holding risks some possibility of being adopted by other Circuits in other contexts. The Ninth Circuit could have distinguished *linkLine* on the facts, because Abbott did not sell directly to its

---

100 *Id.* at 906.
101 *Doe v. Abbott Labs.*, 571 F.3d 930, 932 (9th Cir. 2009).
102 *Id.* at 931.

competitors; it sold Norvir directly to consumers. While the Supreme Court may need a more explicit decision to bring all lower courts into line, *Abbott Laboratories* may foreshadow the proposition that, even where a refusal to deal is not directly at issue, some circuits may read the Court's 21st century jurisprudence as nearly eliminating lower courts' ability to police monopolists' pricing.

3. *Dentsply*

Joining the cluster of lower court cases condemning exclusionary distribution practices by dominant firms is *United States v. Dentsply International, Inc.,*[103] a case brought by the DOJ against Dentsply, the nation's largest manufacturer of dental equipment and supplies. DOJ alleged that Dentsply's policy, prohibiting dealers that carried Dentsply's artificial teeth from carrying competitive products, amounted to unlawful exclusive dealing and unlawful maintenance of a monopoly. After a five-week trial, the district court issued a 165-page opinion finding that, although Dentsply had a high market share, it was not able to exclude competition from a substantial share of the market for artificial teeth. The court also found that Dentsply's exclusive dealing did not violate § 2 of the Sherman Act because the government had failed to prove that Dentsply had monopoly power. Notably, even though distribution through Dentsply's dealers "may be easier" for rivals, the court held that "it is not the function of the antitrust laws to ease the burden of competing with an established and focused rival."[104]

The court also held that the absence of liability under § 3 of the Clayton Act, which prohibits exclusive dealing (as does § 1 of the Sherman Act), prevented the government from prevailing under § 2.

The Third Circuit reversed the District Court, finding liability for monopolization based mainly upon the propositions that: (1) Dentsply had a persistent share of some 75-80 percent of the U.S. tooth market and enjoyed monopoly power; (2) Dentsply's purpose in adopting the exclusive dealing policy was anticompetitive; (3) certain of Dentsply's proffered non-exclusionary business reasons were "pretextual"; and (4) the policy of exclusive dealing (and not the ineffectiveness of rivals and other factors that the District Court had found) in fact foreclosed rivals from (unquantified) access to "key dealers" that represented a "narrow, but heavily traveled channel" of distribution.

Dentsply asked for rehearing and rehearing *en banc,* on the grounds, among others, that the Court of Appeals simply found its own set of facts, and petitioned the Supreme Court for a writ of *certiorari*, but both courts denied review. The case is yet another instance in which the soaring language of the Supreme Court has a rather different sound down on the ground in the courts of appeal.

Probably the most interesting basis for the Third Circuit's *Denstply* opinion is the notion that exclusive dealing arrangements that raise no issue under § 1 of the Sherman Act can nonetheless provide the basis for § 2 liability. In the past, exclusive dealing arrangements that did not violate § 1 were generally not treated as anticompetitive for purposes of § 2. As noted above, that has now changed at least in some courts.[105] But the basis for that fresh § 2 liability is far from clear, although what is clear is that properly

103 399 F.3d 181 (3d Cir. 2005), *cert. denied,* 126 S. Ct. 1023 (2006).
104 277 F. Supp. 2d 387, 450 (D. Del. 2003).
105 *Microsoft,* 253 F.3d at 70 ("[m]onopolist's use of exclusive contracts, in certain circumstances, may give rise to a § 2 violation even though the contracts foreclose less than the roughly 40% or 50% share usually required in order to establish a § 1 violation.").

instructed juries have a great deal of latitude in finding liability under their own general notion of "fairness". The courts finding a defendant liable under § 2 have simply characterized the channels, foreclosure from which was held illegal, to be "key" or "efficient".[106] The failure of *Microsoft* and *LePage's* - and, now, *Dentsply* - to provide guidance concerning the degree of foreclosure or quality of channel that may lead a court to declare, on a different set of facts, that § 2 has been violated has left the law of exclusive dealing in shambles.

C. *Obama DOJ Breaks With Bush DOJ, and Possibly the Supreme Court*

The administration of the Antitrust Division, Department of Justice, during the years of the second President Bush was remarkably flaccid. The DOJ began its tenure by accepting a consent decree with Microsoft that would almost certainly never have been adequate for the Clinton administration, which began the action and sought a structural remedy. During these Bush years, the DOJ did not initiate a single § 2 case or, so far as the record discloses, commence any meaningful investigations with respect to single firm conduct. It did issue a highly controversial report articulating its enforcement priorities, *Competition and Monopoly: Single-Firm Conduct Under Section 2 of the Sherman Act*, (the "Report"). The Report, issued in September 2008, grew out of a joint project that began in 2006 between DOJ and the FTC, including a year-long series of joint hearings, with 29 separate panels and 119 witnesses covering a wide range of topics and perspectives. Despite the effort that went into it, the Report was not well-received. The FTC declined to endorse the Report, and the majority of the FTC Commissioners issued a strongly-worded critique. Commissioners Harbour, Leibowitz, and Rosch wrote that the Report, "if adopted by the courts, will be a blueprint for radically weakened enforcement of Section 2," and that the enforcement principles DOJ sets forth "would place a thumb on the scales in favor of firms with monopoly or near-monopoly power."[107] At its very first opportunity, the new administration signaled its agreement with the FTC, and perhaps also its disagreement with the approach to antitrust reflected in the most recent Supreme Court decisions.

In her first speech after her confirmation as Assistant Attorney General, Christine Varney expressly withdrew the Report. Announcing that the Report "no longer represents the policy of the Department of Justice with regard to antitrust enforcement under Section 2 of the Sherman Act [and that] the Report and its conclusions should not be used as guidance by the courts, antitrust practitioners, and the business community," AAG Varney criticized the Report for "rais[ing] many hurdles to Government antitrust enforcement."[108] Among other concerns, she criticized the report's skepticism about the "ability of antitrust enforcers—as well as antitrust courts—to distinguish between anticompetitive acts and lawful conduct," and argued the Report placed excessive emphasis on "a dominant firm's ability to act efficiently" and "understate[d] the importance of redressing exclusionary and predatory acts that result in harm to competition, distort markets, and increase barriers to entry."[109]

In any case, AAG Varney did not criticize, at least explicitly, the opinions on which the Report was based, or any Supreme Court holdings, but she cited as lodestars for

---

106 *Id.* (despite no unlawful exclusive dealing under § 1, district court found § 2 violation; "Microsoft had substantially excluded Netscape from 'the most efficient channels for Navigator to achieve browser usage share,' ... and had relegated it to more costly and less effective methods") (citation omitted); *LePage's,* 324 F.3d at 160 (loss of "key retail pipelines necessary to permit it to compete profitably" sufficient under § 2).

107 Statement of Commissioners Harbour, Liebowitz, and Rosch On the Issuance of the Section 2 Report By the Department of Justice (Sept. 8, 2008) at 1, *available at* www.ftc.gov/os/2008/09/080908section2stmt.pdf.

108 Christine A. Varney, Vigorous Antitrust Enforcement In This Challenging Era (May 12, 2009) at 7, 9, *available at* http://www.usdoj.gov/atr/public/speeches/245777.htm.

109 *Id.* at 6, 7.

the new administration's enforcement policy cases far removed from the Court's current jurisprudence: *Lorain Journal v. United States*,[110] *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.*,[111] *Microsoft*, *Dentsply*, and *Conwood*. Relying on these five cases and their undergirding principles, the DOJ has made clear that it will attack exclusionary or predatory conduct where it has an effect on competition, and ultimately consumers. The cases pointed to are well-known and to a helpful degree somewhat straightforward, at least in their articulation of key principles. The three lower court opinions are discussed above, while *Lorain Journal* and *Aspen* pre-date the current Supreme Court's rollback of § 2. But we summarize them all to simplify an elucidation of their guiding principles.

1. *Lorain Journal:* a newspaper publisher was the only business disseminating news and advertising in the Ohio town of Lorain until a small radio station began broadcasting in a neighboring community. The newspaper publisher sought to destroy the competitor by refusing to sell advertising space to anyone who also used the radio station for local advertising. The Supreme Court found the publisher to have violated § 2 of the Sherman Act by virtue of its exclusionary and predatory conduct.

2. *Aspen Skiing:* Ski Co. owned three of the four major downhill skiing facilities in Aspen, Colorado. Highlands owned the fourth. After many years of cooperating with Highlands to offer interchangeable ski passes that could be used at all four facilities, Ski Co. discontinued the practice and refused to sell lift tickets to Highlands even if Highlands was willing to pay full retail rates for them. Thus, Ski Co. was willing to sacrifice short-run benefits and consumer goodwill in exchange for a desired long-run impact on Highlands's business.

    a. The prior cooperation between the two companies was probably pivotal to the outcome. Absent the prior cooperation, the case likely never would have been brought, and if brought almost certainly would not have been decided as it was.

    b. The period of cooperation also began at a time when Aspen did not own all three of the mountains; prior to changing its policies, it obtained its dominant position by acquisition (a point rarely mentioned in writings about the case), making the change in policy more effective in excluding the smaller rival Highlands.

3. *Microsoft:* Microsoft was found to have violated § 2 of the Sherman Act by tying its browser (Internet Explorer) to its Windows operating system, thus effectively excluding Netscape from the market and thereby protecting its monopoly in operating systems.

4. *Dentsply:* Dentsply had a dominant share (74 to 80 percent) of the market for false teeth distributed to dentists in the U.S. The company had a policy of refusing to deal with distributors who handled the product of a rival and there was evidence that (a) Dentsply's purpose in adopting the exclusive dealing policy was anticompetitive and (b) the policy of exclusive dealing in

110 342 U.S. 143 (1951).
111 472 U.S. 585 (1985).

fact foreclosed rivals from access to "key dealers" that represented a "narrow, but heavily traveled channel" of distribution.[112]

5. *Conwood*: U.S. Tobacco ("UST") had a dominant share (more than 80 percent) of the U.S. market for moist snuff. In affirming the enormous judgment against UST, the Sixth Circuit referred to UST's misuse of its role as category Captain, unauthorized destruction of Conwood racks, burying competitive products on the UST industry rack and misrepresenting its sales performance to increase display space given to slower-moving UST snuff products. Based on internal UST documents that demonstrated an anticompetitive intent, the Sixth Circuit was satisfied that the exclusionary conduct was widespread, unjustified, driven by anticompetitive intent, and harmed consumers by raising prices, limiting choice, and slowing the growth of rivals.

The main takeaway from these cases, and from the DOJ's recent statements about its change of policy with respect to single firm conduct, is certain that such conduct will become a target of government investigation and/or litigation if the conduct:

a. Is exclusionary or predatory

b. Has no apparent legitimate business purpose (advancing one's own business interest solely by injuring a rival is not thought of as "legitimate")

c. Is engaged in by a dominant firm (any firm with 50 percent or more of a defined market), and

d. Injures competition (perhaps by slowing down or injuring a rival or raising barriers to entry).

Great change could occur if the courts permit cases to go forward that are based on these four principles. These principles seem to promote back to a position of primacy the importance of evidence going to the "legitimacy" of the business conduct at issue. Indeed, it is interesting and perhaps instructive to revisit what the DOJ said about this in its brief to the *en banc* D.C. circuit in *Microsoft* back on January 12, 2001, about a week before the Bush administration took over the case:

> The Supreme Court has described exclusionary conduct as conduct that "not only (1) tends to impair the opportunities of rivals, but also (2) either does not further competition on the merits or does so in an unnecessarily restrictive way". ... if "valid reasons" do not justify conduct that tends to impair the opportunities of a monopolist's rivals, that conduct is exclusionary.

Brief for Appellees United States and the State plaintiffs in *United States v. Microsoft,* at 47 (Jan. 12, 2001) (internal citations omitted). And further:

> Microsoft is mistaken if it means to suggest that a series of actions, which standing alone would not be unlawful, can never, in combination, resulting in a violation of

112 This is the last successful monopolization case brought by the Antitrust Division. It was brought in Wilmington, and ultimately decided by the Third Circuit in Philadelphia in 2005.

> the Sherman Act. ... an individual practice that serves no legitimate purpose and is intended to exclude a rival might nevertheless have so modest an effect on competition as not to violate the Sherman Act. But a coordinated campaign of such acts that in the aggregate has the requisite impact on the marketplace is unlawful. ... as a matter of both logic and sound antitrust law, the market effects of Microsoft's anticompetitive actions should be considered in their totality. It would be irrational to allow a monopolist to inflict a thousand anticompetitive cuts, many perhaps causing only small injury in isolation, that collectively extinguish or disable competition in the relevant market. The Sherman Act does not require courts to ignore the realities of an anticompetitive course of conduct.

Id. at 82. One can, for example, imagine many *Kodak*-type cases and other cases that would come out rather differently under this sort of standard, compared to what would happen under the screens articulated in the now withdrawn DOJ Guidelines.

Perhaps Varney's most significant statement was that, following the D.C. Circuit's opinion in *Microsoft*, the DOJ "will need to look closely at both the perceived procompetitive and anticompetitive aspects of a dominant firm's conduct, weigh those factors, and determine whether on balance the net effect of this conduct harms competition and consumers."[113] Her statement suggests that the new administration would not be comfortable with advocacy in favor of bright-line rules creating safe harbors for conduct by dominant firms – an advocacy that finds strong support in the Supreme Court's cases since *Brooke Group* and that of course found strong support in the most recent Bush administration.

It remains to be seen whether or to what extent DOJ will seek to expand the categories of conduct that are deemed "exclusionary" or "predatory" under § 2, but it is entirely possible that DOJ will assert that § 2 covers conduct that the courts have not yet recognized as exclusionary or predatory or that the courts have previously found to be within a safe harbor. If DOJ identifies conduct that "on balance" harms consumers and competition, they can be expected to both target the conduct directly and file *amicus* briefs in existing private litigation seeking to advance a more expansive view of the rights of plaintiffs under § 2 of the Sherman Act.

D. *The Federal Trade Commission*

The Federal Trade Commission was far out in front of DOJ for most of the Bush administration, and appears likely to remain so even in the new environment, in at least two areas. First, the FTC has claimed a lead role in challenging patent settlements between brand and generic prescription drug manufacturers when the settlements are accompanied by "reverse payments" from the brand manufacturer to the generic, challenging such alleged reverse payments agreements both under § 1 and under § 2 in *FTC v. Cephalon*. Second, the FTC has actively pursued companies perceived to exploit standard setting organizations by concealing technology or breaching apparent obligations to license on "reasonable and non-discriminatory" ("RAND") terms. The FTC has successfully challenged alleged abuses of standard-setting procedures under § 5 of the FTC Act, which allows the FTC to both enforce the antitrust laws to prohibit "unfair methods of competition," and to protect consumers by punishing "unfair or deceptive acts or practices."[114] But the FTC has been less successful when it bases liability theories explicitly on § 2 of the Sherman Act.[115]

113 Vigorous Antitrust Enforcement In This Challenging Era, *supra,* at 13.
114 15 U.S.C. § 45(a)(1).
115 After years of litigation, the FTC found a defendant guilty of violating § 2 in a lengthy opinion *In the Matter of Rambus, Inc*, Docket No. 9302 (Aug. 2, 2006), *available at* http://www.ftc.gov/os/adjpro/d9302/index.shtm, but ultimately suffered reversal on appeal. *Rambus, Inc. v. Federal Trade Commission*, 522 F.3d 456 (D.C. Cir. 2008).

The FTC has vigorously pursued defendants in the area of patent settlements, undeterred either by resistance from the Bush DOJ or consistent failure in court. For instance, in the *Schering Plough* case,[116] involving the controversial settlement of a patent dispute between a branded manufacturer and a generic manufacturer, Schering won the trial before the FTC's ALJ, lost the appeal to the Federal Trade Commission, and won in the Eleventh Circuit. The FTC then filed its own petition for *certiorari,* which was opposed by the Department of Justice and the Solicitor General,[117] and the Supreme Court denied the petition. The new administration's DOJ, by contrast, appears more willing to support the FTC's efforts. In *Arkansas Carpenters v. Bayer, AG* (the ciprofloxacin hydrochloride antitrust litigation), the FTC had already filed a brief in the litigation when the Second Circuit invited "the United States" to express its position on patent settlements accompanied by reverse payments. DOJ took the opportunity to support the FTC, and filed a brief arguing that reverse payments should be treated as "presumptively unlawful" under § 1 of the Sherman Act.[118] This signal from DOJ suggests that it will be similarly willing to support the FTC's position in § 2 litigations such as the *Cephalon* case. (DOJ has not yet had the opportunity to weigh in on the *Cephalon* case - the defendant's motion to dismiss has not yet been decided more than 18 months after the FTC's Complaint was filed.)

The FTC has also taken the lead in challenging "patent holdup" situations, in which one member of a standard setting organization ("SSO") exploits intellectual property rights covering a technology essential to practice the standard. Normally SSOs require all participants to disclose any intellectual property rights essential to practice the standard, and commit to license such essential intellectual property on RAND terms. Holdup situations arise when a member of the SSO either (a) makes a RAND commitment and fails to honor it, or (b) fails to disclose its intellectual property in order to avoid making a RAND commitment. The FTC has challenged both types of patent holdup under § 5 of the FTC Act.

The FTC's first action against holdup came in 1996, when the FTC alleged that Dell Computer Corporation violated § 5 by breaching its commitment to disclose patents to an SSO before the organization developed a standard relying on those patents.[119] The FTC alleged a violation of the antitrust laws and reached a consent decree with Dell, under which Dell agreed not to enforce its patent rights against computer manufacturers complying with the standard. In 2005, the FTC reached a consent decree with Union Oil Company of California ("Unocal") settling charges that Unocal violated § 5 by misrepresenting its intellectual property rights to a board promulgating standards governing low-emissions gasoline, and thus wrongfully obtained monopoly power after refiners became locked-in to regulations that required the use of defendant's proprietary technology.[120] Just as in *Dell*, the FTC alleged a violation of the antitrust laws, and Unocal agreed not to enforce the relevant patents.

Most recently, and more controversially, in 2008 the FTC extracted a consent decree from Negotiated Data Solutions LLC ("N-Data").[121] Unlike *Dell* and *Unocal*, which

116 *Schering-Plough v. FTC,* 402 F.3d 1056 (11th Cir. 2005), *cert. denied,* 126 S. Ct. 2929 (2006).
117 DOJ's brief opposing certiorari criticized some aspects of the 11th Circuit's opinion but argued that there was no circuit split, *available at* http://www.usdoj.gov/atr/cases/f216300/216358.htm.
118 http://www.usdoj.gov/atr/cases/f247700/247708.htm.
119 *In the Matter of Dell Computer Corp.*, 121 F.T.C. 616 (May 20, 1996).
120 In re *Union Oil Co. of California*, Docket No. 9312 (Aug. 2, 2005), *available at* http://www.ftc.gov/os/adjpro/d9305/index.shtm. The FTC was able to extract the consent decree in connection with Chevron's acquisition of Unocal.
121 *In the Matter of Negotiated Data Solutions LLC,* File No. 0510094 (2008), *available at* http://www.ftc.gov/os/caselist/0510094/index.shtm.

were based on failure to disclose technologies, the FTC alleged a violation in *N-Data* based on the respondent's change in the rates at which it would make available licenses to essential patents covering an industry standard. The controversy arose because of a significant distinction between the Commission's liability theory in *N-Data* and its liability theories in *Dell* and *Unocal*. In N-Data, the FTC alleged a violation of § 5 (just as in *Dell* and *Unocal*), but it did not allege a violation of the antitrust laws. Instead, in *N-Data* the FTC alleged that N-Data's conduct constituted a freestanding violation § 5 without violating § 2. Chairman Majoras and Commissioner Kovacic both dissented from the Commission's decision to lodge a complaint, arguing that the "unfair acts or practices" prong of § 5 should not be applied, and that the "the preconditions for use of stand-alone § 5 authority to find an unfair method of competition [absent an antitrust violation] are not present."[122] The FTC's approach in N-Data indicates that the Commission is willing to resort to other theories of liability when § 2 jurisprudence fails to offer a weapon to combat practices to which the FTC objects.

But the FTC's most significant recent action was probably its *Rambus* opinion, which was ultimately reversed by the D.C. Circuit. The FTC's theory of liability in *Rambus* was that the defendant should have revealed its technology to the SSO, which would have given the members of the organization the opportunity to either (a) seek a commitment that the technology would be licensed on RAND terms, or (b) develop an alternative to the technology. The D.C. Circuit found it a relatively easy case on appeal, holding that the FTC's theory of liability depended on the assertion that the standard-setting organization would have sought a RAND commitment if the technology had been disclosed. This, to the D.C. Circuit, distinguished the Third Circuit's conclusion that § 2 could be violated by deception of an SSO in *Broadcom v. Qualcomm*.[123] In that case, Broadcom alleged that Qualcomm violated § 2 "by falsely promising to license its patents on [RAND] terms, and then reneging on those promises after it succeeded in having its technology included in the standard."[124] *Rambus* and *Broadcom* differ because in *Broadcom* the plaintiff alleged that the SSO relied on the RAND commitment such that absent the deception Qualcomm's technology would not have been selected,[125] while in *Rambus*, "the Commission expressly left open the likelihood that [the SSO] would have standardized Rambus's technologies *even if Rambus had disclosed* its intellectual property."[126] Because all that Rambus accomplished through its alleged deception was to avoid making a RAND commitment, the outcome was controlled by the Supreme Court's decision in *Nynex Corp. v. Discon, Inc.*,[127] which held that a monopolist does not violate § 2 by avoiding a constraint on its pricing because avoiding a pricing constraint does not actually exclude competition.[128] The FTC sought *certiorari,* unsupported by the Bush administration's DOJ, and the Petition was denied.

While the FTC was unsuccessful in *Rambus*, its positions in both *Rambus* and *N-Data* suggest that the FTC is willing to push the envelope of § 2 liability in the standard-setting context. If DOJ joins the FTC in this area as well, the Supreme Court may face assaults on its jurisprudence from both agencies, heard by lower courts similarly willing to seek the cracks in the protection the Supreme Court has erected for dominant firms.

---

122 Dissenting Statement of Chairman Majoras *In the Matter of Negotiated Data Solutions LLC,* File No. 0510094 (2008), *available at* http://www.ftc.gov/os/caselist/0510094/index.shtm.
123 *Broadcom Corp. v. Qualcomm, Inc.*, 501 F.3d 297 (3d Cir. 2007).
124 *Id.* at 306.
125 *Id.* at 314.
126 522 F.3d at 466 (emphasis in original).
127 525 U.S. 128 (1998).
128 522 F.3d at 464-66.

E. *Final Thoughts*

The current tension between the Supreme Court's jurisprudence and the more activist inclinations that persist in some of the lower courts and at the agencies is neither a struggle between good and evil, nor an all-or-nothing proposition. Antitrust jurisprudence since the mid-1980s has added analytical rigor and depth to § 2 antitrust analysis, and the current Supreme Court's formidable antitrust thinkers – including both the Chicago-influenced Justice Scalia and the Harvard-school Justice Breyer – have made imprints on antitrust doctrine that may prove indelible. Section 2 enforcement will likely continue to proceed on the ground they have prepared, although perhaps with detours here and there and some road building into new areas. For instance, the divergence between *Rambus* and *Broadcom* was entirely about whether Justice Breyer's 1998 *Nynex v. Discon* opinion was applicable, not over whether it was correct. The fact that outcomes have been pro-defendant since 1993 may reflect as much about the enforcement excesses of previous decades as about the Supreme Court's rigid adherence to dogma, although all of this remains to be seen. Politics and judicial appointments will play a role in the future trajectory of U.S. antitrust. Indeed, for many years there was the notion, almost certainly a convenient fiction, that antitrust had reached a sort of lasting bipartisanship. But politics drives policy and the last several years have demonstrated that politics matters no less in antitrust than elsewhere.

It is worth pointing out, finally, that the substantial influence of the Chicago School in antitrust since the 1980s was a byproduct of a strong and relatively widespread belief in the operation, efficiency, and self-correcting nature of markets in general. The recent credit crisis – encompassing the spectacular collapse of Lehman Brothers and AIG, and the near collapse of the domestic financial system – has led to a broad and deep consensus among policymakers (at least in the Obama administration, and almost certainly much more broadly) that market-based economies are neither as efficient nor self-correcting as previously thought. This collapse in faith and confidence in markets strikes directly at core antitrust principles of the last two and a half decades and represents an independent basis for the current administration to become both more regulatory and more interventionist in a variety of markets, including especially those where there appears to be a "dominant" player.

What is not so clear is what sort of economic or regulatory model will replace the Chicago school's heady optimism that all would be well if we just trusted markets. But now that we know enough to distrust, indeed be fearful of, markets, what can or should we trust to provide a framework within which to construct an antitrust regime and a coherent set of rules? Congress? Sectoral regulators? Courts? These are the larger questions that the new administration needs to address and that conferences such as this can aid.

CARROTS & STICKS: IN DEFENSE OF A DIFFERENTIATED APPROACH TO BUNDLED DISCOUNTS & TYING

Jeane A. Thomas and Ryan C. Tisch[1]
Crowell & Moring LLP
Washington, DC

I. INTRODUCTION

Bundled discounts and tying are both practices used by sellers to increase sales, achieve transaction cost and other efficiencies, and encourage loyalty on the part of their customers. In their simplest forms, the two practices can be considered as bookends of a continuum of conduct – ranging from procompetitive arrangements that reduce prices by providing incentives for consumers to purchase the bundle ("carrots"), to anticompetitive arrangements that coerce buyers into purchasing products they otherwise would chose to obtain from competitive sources ("sticks").

Some commentators have advanced the view that the difference between bundling and tying is chiefly one of the degree of competitive foreclosure that they produce. As a result, some have advanced the position that the most appropriate mode of analysis for the two practices is a unified one, in which all forms of bundled sales are judged under a full-blown rule of reason approach, weighing the anticompetitive foreclosure produced by the practice against any procompetitive efficiency benefits that might accrue.

In this paper, we set forth the approaches taken by courts and enforcement authorities in the U.S. and Europe to various forms of tying and bundling arrangements. We then evaluate the arguments for and against a uniform approach to the analysis of such practices, and conclude that a differentiated approach based on the type of bundling/tying conduct at issue is more consistent with well-settled economic principles, legal and business practicalities, and the underlying goals of antitrust law and policy.

II. TYING: THE STICK

Tying is among the classic forms of conduct condemned by the antitrust laws. The analytical framework under which such conduct is assessed has changed markedly over time, as the "[Supreme] Court's strong disapproval of tying arrangements has substantially diminished,"[2] but it is possible to argue that the analysis remains relatively strict under U.S. law – and is far stricter as applied in Europe.

---

1 Jeane A. Thomas is a partner and Ryan C. Tisch is a counsel at Crowell & Moring LLP. The authors gratefully acknowledge the valuable contributions to this paper by Timothy C. Carson, as associate at Crowell & Moring LLP.
2 *Illinois Tool Works Inc. v. Independent Ink, Inc.*, 547 U.S. 28, 35 (2006).

Even as tying analysis has changed from a strict *per se* analysis with several threshold prerequisites to one that incorporates at least some elements of a rule of reason inquiry, the analysis clearly presumes to disfavor conduct deemed to constitute tying. This presumption has developed from a long history of observations by the courts that tying behavior more often involves forcing a buyer to do something against its interests ("coercion") rather than providing tangible benefits to consumers, such as lower prices in the case of bundled discounts.

A. The Classic *Per Se* Approach

In its classic form, tying is "an agreement by a party to sell one product only on the condition that the buyer also purchases a different (or tied) product, or at least agrees that he will not purchase that product from any other supplier."[3] In so doing, the seller harnesses market power he has over the "tying" product to prevent competition for sales of the "tied" product, and "[w]here such conditions are successfully exacted competition on the merits with respect to the tied product is inevitably curbed."[4]

Tying harms consumers by 1) inducing more sales at higher prices of the tied product than the seller would make in a fully competitive market, 2) "deny[ing] competitors free access to the market for the tied product," and 3) forcing buyers "to forego their free choice between competing products."[5]

Classic tying analysis is based on strong presumptions against the legality of the practice, which the Supreme Court articulated in its 1947 decision in *Standard Oil Co. v. United States*: "[T]ying agreements serve hardly any purpose beyond the suppression of competition."[6] Because of the strength of this presumption, "tying agreements fare harshly under the laws preventing restraints of trade."[7] In fact, conduct meeting all elements of the test was historically condemned *per se*.[8]

Tying arrangements have traditionally been evaluated under Section 1 of the Sherman Act as agreements in restraint of trade. The first element of the test requires that there be "two separate product markets [that] have been linked."[9] This element is designed to exclude the possibility that the two products are of a type that *should* be sold together because of efficiencies inherent in the tied offering.

Second, the tie must involve the element of "coercion," or "conditioning." As the Supreme Court explained in *Jefferson Parish*,

> . . . the essential characteristic of an invalid tying arrangement lies in the seller's exploitation of its control over the tying product to force the buyer into the purchase of a tied product that the buyer either did not want at all, or might have preferred to purchase elsewhere on different terms. When such "forcing" is present, competition on the merits in the market for the tied item is restrained and the Sherman Act is violated.[10]

3 *N. Pac. Ry. Co. v. United States*, 356 U.S. 1, 5-6 (1958).
4 *Id.* at 6.
5 *Id.*
6 *Standard Oil Co. v. United States*, 337 U.S. 293, 305 (1949).
7 *Times-Picayune Publ'g Co. v. United States*, 345 U.S. 594, 606 (1953).
8 *See, e.g.*, *Jefferson Parish Hosp. Dist. No. 2 v. Hyde*, 466 U.S. 2, 9-10 (1984).
9 *Id.* at 21.
10 *Id.* at 12.

Courts have considered a variety of different types of conduct to constitute coercion in alleged cases of tying. An "express refusal" to sell the tied products separately, often in the form of a clear contractual requirement that two products be purchased together, has consistently been accepted by courts as evidence of coercion.[11] Short of an explicit contract, some courts have accepted threats by sales personnel to withhold the tying product as evidence of coercion,[12] but generally have not accepted that mere sales pressure to accept both products involves a sufficient level of coercion.[13] More recently, courts have evaluated "technological" or "technical" tying, in which the "tied good [is] physically and technologically integrated with the tied good."[14] However, treatment of technological tying is particularly controversial. Some U.S. courts have applied a rule of reason analysis to technological ties for fear of chilling potentially cognizable procompetitive benefits of such ties.[15] By contrast, the European Commission's recent Guidance reflects a belief that technological ties pose greater risks to competition than do other forms of tying, because they are "costly to reverse" and "reduce[] the opportunities for resale of individual components."[16]

Some courts have explored more quantitative evidence about the purchasing behavior of buyers, refusing to find coercion where a significant number of buyers purchased the tied products from other sellers or separately.[17] And several courts have explicitly considered price manipulation (offering the elements of a tied offering singly only at higher prices) as a form of coercion, something that will be discussed further in this paper.[18]

Third, to constitute a tying violation, a seller must have "some special ability – usually called 'market power' – to force a purchaser to do something that he would not do in a competitive market."[19] The Court in *Jefferson Parish* clarified that this market power could be shown "when the seller's share of the market is high . . . or when the seller offers a unique product that competitors are not able to offer."[20]

Finally, tying behavior must affect "a 'not insubstantial' amount of interstate commerce,"[21] a *de minimis* test that requires only that foreclosure affect more than a tiny *absolute* amount of business.[22] The "not insubstantial" element does not incorporate any real or practical economic analysis of competitive effects.[23]

The traditional tying test thus operates as a short-form analysis that imposes *per se* liability categorically on all conduct that satisfies the requisite elements: separate products, coercion and market power. The elements of the test operate as "screens" to weed out cases in which the strong presumption against tying might not be warranted. For instance, the "separate products" test seeks to identify cases in which tying two products together could have efficiencies, a circumstance that has recently been attributed to tying in the software

---

11 *Amerinet, Inc. v. Xerox Corp.*, 972 F.2d 1483, 1500 (8th Cir. 1992); *see also N. Pac. Ry. Co.*, 356 U.S. 1 (1958).
12 *DataGate, Inc. v. Hewlett-Packard Co.*, 60 F.3d 1421, 1426-27 (9th Cir. 1995).
13 *Bob Maxfield, Inc. v. Am. Motors Corp.*, 637 F.2d 1033, 1037 (5th Cir. 1981) ("cajolery to the point of obnoxiousness" falls short of required coercion).
14 *United States v. Microsoft Corp.*, 253 F.3d 34, 90-92 (D.C. Cir. 2001).
15 *Id.* at 92-94.
16 *Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings*, ¶ 53, COM (2009) 864 final (Feb. 9, 2009).
17 *E.g., Paladin Assocs. Inc. v. Montana Power Co.*, 328 F.3d 1145, 1160 (9th Cir. 2003).
18 *Amerinet, Inc. v. Xerox Corp.*, 972 F.2d 1483, 1500 (8th Cir. 1992); *see also Moore v. Jas. H. Matthews & Co.*, 550 F.2d 1207, 1216-17 (9th Cir. 1977).
19 *Jefferson Parish Hosp. Dist. No. 2 v. Hyde*, 466 U.S. 2, 13-14 (1984) (internal citations omitted).
20 *Id.* at 17 (internal citations omitted).
21 *N. Pac. Ry. Co. v. United States*, 356 U.S. 1, 6 (1958).
22 *Telerate Sys., Inc. v. Caro*, 689 F. Supp. 221, 234 (S.D.N.Y. 1988).
23 *Id.* (citing cases where courts have held amounts as little as approximately $60,000 to be "not insubstantial").

industry.[24] Likewise, the market power screen excludes cases in which the seller likely would be unsuccessful in foreclosing competition because consumers would ignore its tied offering unless it were priced competitively.

Most significantly, the "coercion" test screens for cases in which anticompetitive foreclosure is unlikely because consumers really do have the ability to choose feasible alternatives. The test assumes that in cases in which the other prerequisites are met, the existence of a clear, hard "tie" in the form of an overt or contractual refusal to sell separately presumptively results in a "forcing" of consumer choices that excludes competition in the tied product with intolerable anticompetitive consequences.

B. The *Per Se* "Plus" Approach

The traditional test for tying amounts to "condemnation without inquiry into actual market conditions," omits any analysis of the actual foreclosure effects of the alleged tying conduct.[25] In its most recent decisions laying out the classical tying test, *Eastman Kodak Co. v. Image Technical Servs.* and *Jefferson Parish*, the Supreme Court did not prescribe any opportunity for defendants, under the *per se* rule, to argue that any procompetitive benefits their tying conduct might provide would outweigh its anticompetitive effects – where all of the elements of the tying standard have been met.[26]

However, the Supreme Court has articulated a role for competitive effects analysis that can be applied to cases of alleged tying that do not satisfy all of the "screens" for *per se* condemnation.[27] This mechanism allows plaintiffs to offer additional proof with respect to the actual competitive consequences of the complained-of conduct where they do not meet the higher burden required for *per se* treatment.

Lower courts have begun to incorporate some elements of competitive effects analysis into the tying standard, although the Supreme Court has not explicitly ratified such an approach. In *United States v. Microsoft*, the D.C. Circuit considered whether combining the Windows operating system with the Internet Explorer browser constituted an illegal tie.[28] The court found that although the conduct in question satisfied the elements of the traditional tying test, it would be inappropriate to find it *per se* illegal because there could be "a number of efficiencies that, although very real, have been ignored in the calculations underlying the adoption of a per se rule for tying."[29] The court found that while some of these concerns could be dealt with through the "separate products" element of the tying test (ostensibly, a tie manifesting significant efficiencies would be seen as a single product rather than two separate products), it decided that "the separate products test is a poor proxy for net efficiency from newly integrated products," and that a full rule of reason analysis would be a more appropriate analytical tool to explore such efficiencies.[30] While the *Microsoft* court went to great lengths to limit its concern to high-technology markets with which courts have little experience, it is easy to imagine such concerns regarding potential efficiencies in other tied markets.

---

24 *United States v. Microsoft*, 253 F.3d 34, 90-92 (D.C. Cir. 2001); "But not all ties are bad. Bundling obviously saves distribution and consumer transaction costs. PHILLIP E. AREEDA & HERBERT HOVENKAMP, *in* ANTITRUST LAW, ¶ 1703g2, at 51-52 (1991). This is likely to be true, to take some examples from the computer industry, with the integration of math co-processors and memory into microprocessor chips and the inclusion of spell checkers in word processors." *Id.* at 87. "Indeed, if there were no efficiencies from a tie (including economizing on consumer transaction costs such as the time and effort involved in choice), we would expect distinct consumer demand for each individual component of every good." *Id.*

25 *Jefferson Parish Hosp. Dist. No. 2 v. Hyde*, 466 U.S. 2, 15 (1984).

26 *Eastman Kodak Co. v. Image Technical Servs, Inc.*, 504 U.S. 451, 479 (1992); *Jefferson Parish*, 466 U.S. at 16-17.

27 *NCAA v. Bd. of Regents of the Univ. of Okla.*, 468 U.S. 85, 104 (1984).

28 253 F.3d 34 (D.C. Cir. 2001).

29 *Id.* at 94.

30 *Id.* at 92.

Other courts have acknowledged this potential, leading them to import effects analysis as a fairly standard component of the tying test. As the Eleventh Circuit found in *U.S. Philips Corp. v. International Trade Commission*, "[i]n order to show that a tying arrangement is per se unlawful, a complaining party must demonstrate that it links two separate products *and has an anticompetitive effect in the market for the second product*."[31]

C. The European Approach

Similar to the traditional U.S. approach, courts in the EU have been reflexively hostile to tying practices, generally condemning them *per se*.[32] In *Tetra Pak v. Commission*, for example, the European Court of First Instance ("CFI") found against a manufacturer of equipment for packaging food in cartons that required buyers purchasing its packaging machines also to purchase all packaging cartons used with those machines from Tetra Pak.[33] The Commission successfully met the threshold predicates for establishing an unlawful tie, including that Tetra Pak had market power in the markets for the tying products (in this case, separate markets for non-aseptic and aseptic packaging machines) and in the market for the tied products (packaging for those machines).[34] The CFI also accepted the Commission's allegation that the machines and packaging materials should be considered separate products, rejecting various technical and public safety arguments advanced by Tetra Pak that the two products are essentially separate.[35] That Tetra Pak's sales agreements involved a high degree of coercion was accepted as obvious from the face of its agreements, which specifically provided that all packaging had to be purchased from Tetra Pak.[36]

Having established the basic threshold elements for a tying case, the *Tetra Pak* court deemed the tying conduct to be "abusive" and therefore a violation of Article 82. The decision involves no real effects analysis at all: while it considers and rejects "objective justifications" advanced by Tetra Pak for its policies, these were found to amount to no more than public policy reasons (having to do with food safety) advanced by Tetra Pak for wanting to tie the two products together. The court did not consider harm to competition other than to say that other packaging manufacturers are necessarily excluded by Tetra Pak's conduct, and it did not consider any procompetitive effects the conduct might have.

In more recent cases, European courts have applied the thresholds tests for tying, but have been more receptive to a fuller effects analysis, at least in theory. In *Microsoft v. Commission*, in which the Commission alleged illegal tying of Microsoft's Windows Media Player (WMP) to the Windows operating system, the Commission established all of the elements of an illegal tie as set forth in *Tetra Pak.* However, the CFI noted that a presumption of competitive effects from the mere existence of the tying elements would not be appropriate:

> There are . . . circumstances relating to the tying of [Windows Media Player] which warrant a closer examination of the effects that tying has on competition in this case. While in classical tying cases, the Commission and the Courts considered the foreclosure effect for competing vendors to be demonstrated by the bundling of a separate

31 *U.S. Philips Corp. v. Int'l Trade Comm'n*, 424 F.3d 1179, 1193-94 (Fed. Cir. 2005) (emphasis added).
32 *See, e.g.*, Case T-30/89, *Hilti AG v. Comm'n*, [1991] ECR II-01439.
33 Case T-83/91, *Tetra Pak v. Commission*, [1994] ECR II-00755; *aff'd* Case C-333/94, *Tetra Pak v. Commission,* [1996] ECR I-5951.
34 *Tetra Pak*, [1994] ECR II-00755 at ¶ 42-78.
35 *Id.* at ¶ 79-81.
36 *Id.* at ¶ 12.

> product with the dominant product, in the case at issue, users can and do to a certain extent obtain third party media players through the Internet, sometimes [free of charge]. There are therefore indeed good reasons not to assume without further analysis that tying [Windows Media Player] constitutes conduct which by its very nature is liable to foreclose competition.[37]

However, in the final analysis, the CFI accepted the Commission's argument that Microsoft's conduct resulted in anticompetitive foreclosure, with a very limited analysis that focused more on Microsoft's distribution advantages and implied disadvantages for rivals and consumers than on direct analysis of the competitive effects these parties would suffer.[38] Nowhere does the Commission consider evidence of competitive effects such as prices to consumers, an analysis which might have led to a different result, given that media players were generally priced at zero at the time the action was brought.

The European Commission's recent Guidance on the application of Article 82 suggests a modern approach similar to that taken to tying cases in the U.S.:

> The Commission will normally take action under Article 82 where an undertaking is dominant in the tying market and where, in addition, the following conditions are fulfilled: (i) the tying and tied products are distinct products, and (ii) the tying practice is likely to lead to anticompetitive foreclosure.[39]

The Commission notes a number of circumstances in which it presumes that tying conduct is likely to be harmful to competition, including longer-term tying arrangements, such as those in which the two products are technically integrated into a whole, regulated markets, and markets in which the two tied products are both substitutable inputs into a manufacturing process.[40] Notably, the Commission's presumption that "technological" tying is likely to be more harmful than contractual tying or tying based on price incentives is directly at odds with the American presumption, which ascribes potential efficiencies to technological tying and treats this practice as worthy of more deference.[41] This can be interpreted as a judgment by the Commission that technological tying is more coercive than contractual tying, whereas U.S. courts tend to come to the opposite conclusion.

The Commission does account for efficiencies in its general formulation of the tying test, in accordance with the trend in modern U.S. case law. It will consider "claims by dominant undertakings that their tying and bundling practices may lead to savings in production or distribution that would benefit customers."[42]

It is also important to note that the threshold for dominance under EU law is significantly lower than it is in the United States, which may lead to the result that similar

---

37 Case T-201/04, *Microsoft Corp. v. Comm'n*, [2007] ECR II-3601, ¶ 977 (*quoting* Case COMP/C-3/37.792 Microsoft at ¶ 841).
38 *Id.* at ¶ 1054.
39 *Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings*, ¶ 50, COM (2009) 864 final (Feb. 9, 2009).
40 *Id.* at ¶¶ 53-57.
41 *Microsoft*, [2007] ECR II-3601, ¶ 293.
42 *Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings*, ¶ 62, COM (2009) 864 final (Feb. 9, 2009).

rules may have significantly broader application as applied by the European Commission and courts.[43]

* * * * *

The traditional tying analysis began as a relatively blunt instrument, efficient in its application but imprecise in its results, condemning absolutely a relatively narrow category of conduct. More recently, both in the U.S. and Europe, it has developed a more nuanced formulation, taking into account more sophisticated arguments regarding the separate products and coercion elements, and softening the poleax of *per se* treatment with some consideration of competitive effects. These developments retain the relative efficiency of the presumption against tying conduct, while preventing false positives at the margins where conduct is less coercive or results in more procompetitive efficiencies.

III. BUNDLED DISCOUNTING: THE CARROT

Analysis of bundled discounting is a relatively newer field than is analysis of tying, and as such is less settled. A "bundled discount" can come in any number of forms, but, like tying, it involves a seller offering two or more products for a single price. In one product ("Product A"), the seller has market power, and in another ("Product B"), it does not. The seller encourages buyers to purchase both Product A and Product B by offering a discount on the two products when they are purchased together.

Bundled discounting arrangements differ from tying because, in theory, customers receive something in return for their purchase of both products from the seller: lower prices in the form of a discount. By contrast, purchasers of tied products receive nothing – they simply do not have the option to buy the tying product without also purchasing the tied product.

U.S. antitrust laws favor lower prices,[44] and courts analyzing bundled discounting have generally shaped their analysis around the assumption that "[b]undled discounts generally benefit buyers because the discounts allow the buyer to get more for less."[45] As many commentators have noted, sellers may offer bundled discounts for a number of reasons, including "to save costs in distribution and packaging, to reduce transaction costs for themselves and their customers, and to increase reliability for customers."[46] Recognition of the potential efficiencies of bundled discounting, and of the potential benefits to consumers, are recognized, to a greater or lesser degree, in the two principal modes of analysis by U.S. courts.

On the one hand, some courts have applied an "anticompetitive foreclosure" test to bundled discounting.[47] This approach is exemplified by the Third Circuit's decision in *LePage's Inc. v. 3M*.[48] The foreclosure analysis stands in sharp contrast to the "discount

43 *See, e.g.*, Case T-30/89, *Hilti AG v. Comm'n*, [1991] ECR II-1439 (finding 55% market share as dominant); Case C-62/86, *AKZO v Comm'n*, [1991] ECR I-03359 (finding 50% market share as dominant); Case 85/76, *Hoffmann-La Roche & Co.* v *Comm'n*, [1979] ECR 461 (finding 47% market share as dominant). "The Commission's experience suggests that dominance is not likely if the undertaking's market share is below 40% in the relevant market." *Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings*, ¶ 14, COM (2009) 864 final (Feb. 9, 2009).

44 *See Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, 475 U.S. 574, 594 (1986).

45 *Cascade Health Solutions v. PeaceHealth*, 515 F.3d 883, 895 (9th Cir. 2008).

46 Antitrust Modernization Commission Report and Recommendations, at 95 (2007).

47 *See, e.g., SmithKline Corp. v. Eli Lilly & Co.,* 427 F. Supp. 1089 (E.D. Pa. 1976), *aff'd,* 575 F.2d 1056 (3d Cir. 1978), *cert. denied,* 439 U.S. 838 (1978).

48 324 F.3d 141 (2003).

attribution" test developed by the district court in *Ortho Diagnostic Systems, Inc. v. Abbott Laboratories, Inc.*, and refined and extended by the Ninth Circuit in *Cascade Health Solutions v. PeaceHealth*.[49]

A. The Foreclosure Standard

The anticompetitive foreclosure test evaluates whether the bundled discount arrangement impermissibly excludes competition in the market for the competitive product. In the classic situation, the seller has market power in Product A and faces competition in the market for Product B. The seller's bundled discount harms competition for sales of Product B by making it difficult, if not impossible, for competitors of Product B to successfully compete with the bundled package price. Employing a rule of reason-type approach, this anticompetitive effect must not be outweighed by the procompetitive business justifications offered by the seller.

The decision handed down by the Third Circuit in *LePage's* illustrates the application, and shortcomings, of the anticompetitive foreclosure approach. In *LePage's*, defendant 3M was deemed to have market power in the market for transparent tape (particularly Scotch brand tape), but faced competition from plaintiff LePage's in the market for "second brand" and "private label" tape. Among other things, 3M offered bundled discounts in the form of tiered rebates that increased for customers who bought multiple categories of 3M products. LePage's alleged that this discounting behavior foreclosed it from being able to sell second brand and private label tape to key retail accounts, thus protecting not only 3M's position in transparent tape but also in second brand and private label tape.[50]

The Third Circuit upheld a jury verdict in LePage's favor. It found that 3M "concedes it possesses monopoly power in the United States transparent tape market, with a 90% market share."[51] It also found that "the jury could have reasonably found that 3M's exclusionary conduct cut LePage's off from key retail pipelines necessary to permit it to compete profitably,"[52] specifically finding that "3M foreclosed LePage's from that critical bridge to consumers that [office] superstores provide, namely, cheap, high volume supply lines."[53]

The Third Circuit specifically credited evidence that LePage's retail customers would have suffered the loss of large rebates from 3M if they had purchased LePage's products in significant quantities, and evidence that LePage's profits and market share declined after 3M began to implement its bundled discounting program, as support for LePage's contention that 3M's program harmed competition. Further, the court pointed to evidence that 3M could later recoup some of the money spent on rebates if it was successful in driving LePage's from the market, citing evidence of barriers to entry to

49 *Ortho Diagnostic Sys., Inc. v. Abbott Labs., Inc.*, 920 F. Supp. 455 (S.D.N.Y. 1996); *Cascade Health Solutions v. PeaceHealth*, 515 F.3d 883 (9th Cir. 2008). Note that several other courts have considered bundled discounting and reached decisions that have formed the backdrop for *PeaceHealth*, but none have undertaken as full an analysis as did the Ninth Circuit. *See, e.g.*, *Masimo Corp. v. Tyco Health Care Group, L.P.*, No. CV 02-4770, 2006 WL 1236666 (C.D. Cal. Mar. 23, 2006) (finding that plaintiffs failed to allege sufficient facts to make out a case based on bundling conduct, and rejecting *LePage's*); *Invacare Corp. v. Respironics, Inc.*, No. 1:04 CV 1580, 2006 WL 3022968 (N.D. Ohio Oct. 23, 2006) (disposing of plaintiff's bundling claim because "unlike in *SmithKline* or *LePage's,* Plaintiff manufactures the same products and can match the product bundles offered by Defendant"); *J.B.D.L. Corp. v. Wyeth-Ayerst Labs., Inc.*, Nos. 1:01-CV-704 & 1:03-CV-781, 2005 WL 1396940 (S.D. Ohio Jun. 13, 2005), *reh'g and reh'g en banc denied*, 485 F.3d 880 (6th Cir. 2007) (granting summary judgment against a plaintiff alleging anticompetitive discounting and specifically declining to apply *LePage's*).

50 *LePage's Inc. v. 3M*, 324 F.3d 141, 160 (3d Cir. 2002).

51 *Id.* at 146.

52 *Id.* at 160.

53 *Id.* at 160 n.14.

show that no other competition likely would restrain 3M.[54] Finally, the Third Circuit dismissed the argument that 3M's discounting conduct was based on efficiencies that 3M would pass on to consumers, holding that 3M had introduced no record evidence to support such arguments.[55]

The *LePage's* decision has been widely criticized, however, for focusing on the impact of 3M's conduct on LePage's itself, rather than on the impact on competition generally or on consumer welfare.[56] The court did not address the possibility that LePage's was unable to compete with 3M because it is a less efficient or effective competitor than 3M. The court explicitly rejected 3M's contention that the antitrust laws do not condemn pricing practices unless they involve some form of below cost pricing, which 3M's discounts did not.[57] As the Antitrust Modernization Commission observed, "[l]ower prices may harm a rival but benefit consumers," and the court in *LePage's* did not apply an objective standard to segregate conduct that harms less efficient rivals from conduct that harms competition itself.[58] As such, the decision "is therefore likely to chill welfare-enhancing bundled discounts or rebates."[59]

B. The Price-Cost Test

The decision in *LePage's* marked a distinct departure from a nascent view of bundled discounting taken by the district court in *Ortho Diagnostic Systems, Inc. v. Abbott Laboratories, Inc*.[60] In that court's early articulation of a bundled discounting analysis, it applied a type of price-cost test known as a discount attribution standard to condemn pricing that fell below the seller's average cost of goods sold after the total amount of discounts for the entire bundle were attributed to the competitive good. However, the court left open the possibility that above-cost pricing could also be anticompetitive.[61] As the *Ortho* court laid out its test, pricing could be deemed anticompetitive where either "(a) the monopolist has priced below its average variable cost or (b) the plaintiff is at least as efficient a producer of the competitive product as the defendant, but . . . the defendant's pricing makes it unprofitable for the plaintiff to continue to produce."[62]

More than a decade later in the *PeaceHealth* decision the Ninth Circuit refined the *Ortho* approach in order to "discern where antitrust law draws the line between bundled discounts that are procompetitive and part of the normal rough-and-tumble of our competitive economy and bundled discounts, offered by firms holding or on the verge of gaining monopoly power in the relevant market, that harm competition and are thus proscribed by § 2 of the Sherman Act."[63]

In that case, plaintiff hospital owner Cascade alleged that defendant, rival hospital system PeaceHealth, offered bundled discounting involving several classes of hospital care.

---

54 *Id.* at 162-63.
55 *Id.* at 164.
56 *See, e.g., J.B.D.L. Corp. v. Wyeth-Ayerst Labs., Inc.*, Nos. 1:01-CV-704 & 1:03-CV-781, 2005 WL 1396940, at *14 (S.D. Ohio Jun. 13, 2005), *reh'g and reh'g en banc denied*, 485 F.3d 880 (6th Cir. 2007) (holding that the *LePage's* "verdict imposed a heavy penalty on 3M without producing consistent guidance for what is permissible price competition in the retail market for a simple item like transparent tape").
57 *LePage's*, 324 F.3d at 151-52 (holding that the predatory pricing standards of *Brooke Group Ltd. v. Brown & Williamson Tobacco Corp.*, 509 U.S. 209 (1993) should not be applied to 3M, which unlike *Brooke Group* defendant Brown & Williamson was a monopolist and therefore 1) could price its monopoly product without regard to competitive restraints and 2) should be subject to a higher level of scrutiny than a mere oligopolist).
58 Antitrust Modernization Commission Report and Recommendations, at 97 (2007).
59 *Id.*
60 *Ortho Diagnostic Sys., Inc. v. Abbott Labs., Inc.*, 920 F. Supp. 455 (S.D.N.Y. 1996).
61 *Id. at* 469.
62 *Id.*
63 *Cascade Health Solutions v. PeaceHealth*, 515 F.3d 883, 897 (9th Cir. 2008).

PeaceHealth was the only local hospital provider to offer the most complex type of hospital services, known as "tertiary care," while both PeaceHealth and Cascade offered the less-sophisticated primary and secondary classes of hospital care services. PeaceHealth sold all three classes of care together to insurers at a lower price than it offered for tertiary services and primary/secondary services as separate *a la carte* offerings. Cascade alleged that this pricing scheme resulted in its exclusion from being able to sell primary/secondary services to these insurers because the "bundled" offering from PeaceHealth was economically more attractive than purchasing tertiary services from PeaceHealth and primary/secondary services from Cascade.[64]

After explaining in detail the potentially procompetitive benefits of bundled discounts, the Ninth Circuit unanimously rejected the *LePage's* analysis because "it does not consider whether the bundled discounts constitute competition on the merits, but simply concludes that all bundled discounts offered by a monopolist are anticompetitive with respect to its competitors who do not manufacture an equally diverse product line."[65] The court went on to cite Judge Greenberg's dissent in *LePage's*, which argued that the anticompetitive foreclosure test "risks curtailing price competition and a method of pricing beneficial to customers because the bundled rebates effectively lowered [the seller's] costs."[66]

Instead, the Ninth Circuit applied a "discount attribution" test to determine whether PeaceHealth's discounts could meet a threshold showing demonstrating that they were unlikely to harm competition and therefore not worthy of further scrutiny.

> A plaintiff who challenges a package discount as anticompetitive must prove that, when the full amount of the discounts given by the defendant is allocated to the competitive product or products, the resulting price of the competitive product or products is below the defendant's incremental cost to produce them.[67]

Like the anticompetitive foreclosure test, the *PeaceHealth* standard applies only to a seller who is a monopolist in Product A, faces actual or potential competition for Product B, and sells both products as a bundle at a discounted price. However, the *PeaceHealth* discount attribution test attempts to determine whether the discount is anticompetitive based on whether an equally-efficient competitor could compete for sales of Product B if the defendant's discount on the entire bundle is attributed solely to Product B. If the net price of Product B after the bundled discount is fully attributed to it is below the seller's average variable cost for Product B, then the arrangement is likely to be anticompetitive because the discount would tend to exclude competition from a hypothetical, equally efficient producer of Product B.[68] The *PeaceHealth* court went on to overturn the jury verdict on the attempted monopolization claim because the jury instructions were based on the *LePage's* anticompetitive foreclosure standard.[69]

---

64 *Id.* at 891-892.

65 *Id.* at 899 (citing Antitrust Modernization Commission Report and Recommendations, at 97 (2007)).

66 *Id.* at 899 (citing *LePage's*, 324 F.3d at 179 (Greenberg, J., dissenting)).

67 *Id.* at 909.

68 *Id.* at 910. In adopting the discount attribution standard and comparing price to costs incurred by the seller, the *PeaceHealth* court explicitly rejected the standard set by the district court in *Ortho Diagnostic Sys., Inc. v. Abbott Labs., Inc.*, 920 F. Supp. 455 (S.D.N.Y. 1996). In *Ortho*, the court condemned bundled discounting where the pricing excluded an equally efficient competitor, requiring that plaintiffs prove their cost bases were equal to or below the defendants; the Ninth Circuit rejected this standard in part because it "does not provide adequate guidance to sellers . . . because the standard looks to the costs of the actual plaintiff." *PeaceHealth*, 515 F.3d at 905. *PeaceHealth* remedies this by comparing price to the costs of the defendant, analogizing these to those of a hypothetical, equally efficient competitor.

69 *PeaceHealth*, 515 F.3d at 910; *see also Safeway, Inc. v. Abbott Labs., Inc.*, 2010 WL 147988 (N.D. Cal. Jan. 12, 2010) (holding that plaintiff drug purchasers sufficiently alleged that bundled pricing of pharmaceuticals violated Section Two under the *PeaceHealth* discount attribution test).

The Ninth Circuit's decision does not automatically condemn all bundled discounting that results in below-cost pricing. The court noted in a footnote that plaintiffs would still need to prove "antitrust injury," meaning harm to competition (as opposed to harm to an individual competitor) of the type the antitrust laws were designed to prevent.[70]

However, after applying the discount attribution standard to plaintiff's monopolization claim, the Ninth Circuit went on to consider the district court's grant of summary judgment in favor of PeaceHealth on plaintiff's claim that "PeaceHealth illegally *tied* primary and secondary services to its provision of tertiary services."[71] The court held that the only issue in contention was the coercion element of the tying test; the district court had granted summary judgment based on evidence presented by PeaceHealth that buyers of its services did not consider themselves to be coerced, and that some insurers bought PeaceHealth services on a non-exclusive basis.[72] The appellate court disagreed, pointing to evidence presented by the plaintiff that (1) the loss of PeaceHealth's bundled discounts "would have had a 'large impact'" on certain insurers, and (2) just 4 of 28 (14 percent) of local insurers bought unbundled services from PeaceHealth.[73] Taken together with evidence of "substantial market power" wielded by PeaceHealth in the market for tertiary services, the court held this evidence was sufficient to establish "genuine factual disputes about whether PeaceHealth forced insurers, either as an implied condition of dealing or as a matter of economic imperative through its bundled discounting, to take its primary and secondary services if the insurers wanted tertiary services."[74]

C. The European Approach

Bundled discounting has generally been treated with greater suspicion by European courts, which have considered bundled discounts to be an abuse when practiced by a dominant firm. Take, for example, *Hilti v. Commission*, in which a dominant manufacturer of nail guns and the consumable nail cartridges and nails used in those guns was condemned by the European Commission and the Court of First Instance for offering larger discounts to purchasers who bought cartridge strips with nails than to those who bought only cartridge strips alone.[75] In that matter, the Commission based its decision on the fact that Hilti had market power and that, among other things, it had "attempted to block the sale of competitors' nails by a policy of reducing discounts for orders of cartridges without nails."[76] The Commission undertook no particular analysis of effects on competition, other than to say that Hilti had been able "to limit the market penetration of independent nail and cartridge strip producers . . . ."[77] The CFI's decision confirming the Commission's analysis reflects neither the consideration of empirical evidence demonstrating foreclosure, nor of evidence that the levels of discounting offered by Hilti prevented an equally-efficient competitor from being able to compete.

The recent Guidance issued by the European Commission indicates an evolution in its approach, suggesting that bundled discount arrangements should be evaluated using

70 *Id.* at 910 n.21 (specifically rejecting the Antitrust Modernization Committee's proposal that plaintiffs be required to show a third element in bundled discounting cases: that "the bundled discount or rebate program has had or is likely to have an adverse effect on competition," instead finding that this requirement is identical to "the general requirement of 'antitrust injury' that a plaintiff must prove in any private antitrust action") (citing Antitrust Modernization Commission Report and Recommendations, at 99 (2007); *Brunswick Corp. v. Pueblo Bowl-O-Mat, Inc.*, 429 U.S. 477, 489 (1977)).
71 *PeaceHealth*, 515 F.3d at 912 (emphasis added).
72 *Id.* at 914.
73 *Id.* at 914-15.
74 *Id.* at 914.
75 Case T-30/89, *Hilti AG v. Comm'n*, [1991] ECR II-1439.
76 *Id.* at ¶ 16.
77 *Id.* at ¶ 17.

the same anticompetitive foreclosure standard applied to tying arrangements.[78] However, the Commission's Guidance incorporates a price-cost test as a sort of "safe harbor" for multi-product discounts:

> If the incremental price that customers pay for each of the dominant undertaking's products in the bundle remains above the [long run average incremental cost ("LRAIC")] of the dominant undertaking from including that product in the bundle, the Commission will normally not intervene since an equally efficient competitor with only one product should in principle be able to compete profitably against the bundle.[79]

For bundled discounts that result in a price below the LRAIC of including that product in a bundle, the Guidance states that "enforcement action may . . . be warranted."[80] This standard roughly mirrors the *PeaceHealth* discount attribution test, although the lower threshold for "dominance" may result in broader application in the EU than in the U.S.

IV. THE SEARCH FOR APPROPRIATE ANALYTICAL STANDARDS

There has recently been a great deal of commentary and very healthy debate over the appropriate standards that courts and regulators should apply to various forms of bundling practices. While much is unsettled in this area of competition law, it would appear that there is consensus regarding the goals of avoiding overdeterrence and underdeterrence that would harm consumer welfare – although experts advocate a variety of approaches designed to achieve those goals.

A. Underlying Goals of Antitrust Standards

One of the most challenging issues in antitrust law is distinguishing, with respect to single-firm conduct, what constitutes competition on the merits that benefits consumers from what constitutes "exclusionary" conduct that ultimately harms competition itself and should be prohibited. As stated by the Antitrust Modernization Commission:

> How to evaluate single-firm conduct under Section 2 poses among the most difficult questions in antitrust law. Appropriate antitrust enforcement must distinguish aggressive competition that benefits consumers, such as most price discounting, from conduct that tends to destroy competition itself, and thus maintains, or facilitates acquiring, monopoly power. The Supreme Court has defined improper "exclusionary" conduct under Section 2 to "comprehend[] at the most behavior that not only (1) tends to impair the opportunities of rivals, but also (2) either does not further competition on the merits or does so in an unnecessarily restrictive way." Thus, a crucial distinction in Section 2 enforcement entails whether a firm's conduct represents competition on the merits or improper "exclusionary" conduct.[81]

78 *Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings*, ¶¶ 47-62, COM (2009) 864 final (Feb. 9, 2009).
79 *Id.* at ¶ 60.
80 *Id.*
81 Antitrust Modernization Commission Report and Recommendations, at 81 (2007) (citing *Aspen Skiing Co. v. Aspen Highlands Skiing Corp.*, 472 U.S. 585, 605 n.32 (1985) (quoting III Phillip E. Areeda & Donald F. Turner, Antitrust Law 78 (1978)).

Thus, in developing standards to evaluate single-firm conduct, courts and academics agree that such tests should strive to encompass those forms of conduct that unduly exclude competition, but are not so broad as to have a chilling effect on conduct that represents legitimate competition on the merits that benefits consumers. As the Supreme Court has stated with respect to evaluating pricing practices by a monopolist, a central goal of such standards should be to avoid "mistaken findings of liability [that] would 'chill the very conduct the antitrust laws are designed to protect.'"[82] As explained by the Antitrust Modernization Committee:

> The recognition of potential consumer harm from overdeterrence has led courts to try to avoid "false positives" – that is, finding Section 2 liability for a firm that has not engaged in unreasonably exclusionary conduct, but instead was simply competing aggressively on the merits. Nonetheless, it remains important to avoid underdeterrence that results in "false negatives" – that is, failing to condemn anticompetitive conduct – when the challenged conduct typically provides few or no benefits to consumer welfare and does not resemble competition on the merits. In an ideal world, of course, legal rules would avoid both underdeterrence and overdeterrence. In practical reality, however, such precision is often difficult to achieve. Thus, courts may need to make a trade-off between accuracy and the risks of either chilling precompetitive, or encouraging anticompetitive, conduct.[83]

Further, there are many practical difficulties in selecting standards that (1) provide enough guidance to courts and enforcement authorities to make consistent enforcement feasible, (2) limit, to the extent possible, the burden on litigants and courts of weeding out unmeritorious claims, and (3) give businesses objective standards by which they can evaluate the legality of their conduct *ex ante* so that procompetitive practices are not unnecessarily deterred. As summed up by the Antitrust Modernization Committee, "standards for applying Section 2 of the Sherman Act's broad proscription against anticompetitive conduct should be clear and predictable in application, administrable, and designed to minimize overdeterrence and underdeterrence, both of which impair consumer welfare."[84]

B. The Continuum of "Bundling" Conduct

The common thread running through the conduct at issue here is that it involves the sale of two or more distinct products in some sort of "package" for a single price. On one end of the continuum, a seller offers on a spot-purchase basis, two products as a "bundle" for a single price that is less than the price for the two items purchased separately. Simple examples include a hamburger and soft-drink, or shampoo and conditioner, sold together at a discounted price. At the other end of the continuum, a seller with market power in Product A refuses to sell Product A alone unless purchasers also buy Product B, for which the seller faces actual or potential competition. This is classic tying behavior.

But the realities of the business world are rarely so simple. There are a myriad of forms by which sellers package their products and services together, limited only by the

82 *Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber Co.,* 549 U.S. 312, 320 (2007) (quoting *Cargill, Inc. v. Monfort of Colo., Inc.,* 479 U.S. 104, 122 n.17 (1986)).

83 Antitrust Modernization Commission Report and Recommendations, at 90 (2007) (footnotes omitted).

84 *Id.* at 88; *see also id.* at 96 ("[W]hatever legal standards are adopted should be sufficiently clear to enable companies to conform their conduct to the law, be administrable by the courts, and avoid chilling procompetitive discounting.").

imagination and creativity of the free-market system.[85] These types of "bundles" might include, for example, loyalty or fidelity agreements in which a buyer is granted discounts or rebates in return for an agreed or *de facto* commitment to purchase a large or increasing share of its requirements for multiple products from the supplier. They may also include market share agreements or requirements contracts, where a buyer is granted price concessions for purchasing all or a certain proportion of its requirements across more than one product line from the seller.

From an analytical perspective, the challenge is in categorizing these arrangements as bundled discounts, exclusive agreements, ties or otherwise – and then applying the appropriate standard to achieve the desired goals. As widely noted, however, the lines between these categories of conduct often are not easily drawn.[86] And the legal standards appropriate for each category of conduct are far from clear.[87]

C. Proponents of a Uniform Test

In the face of the difficulties of appropriately categorizing bundling practices and their attendant potential for anticompetitive effects, some commentators advocate a uniform approach to all forms of conduct by a monopolist in which multiple products are bundled together.

Some have argued that all forms of bundling should be condemned *per se.*[88] However, this certainly would over deter procompetitive price-cutting conduct that benefits consumers and the consensus is that such an approach is clearly overbroad.

Another approach is to apply a rule of reason-type approach to all forms of bundling/tying to avoid the possibility of underdeterrence that may result from incorrect or inappropriate categorization of the conduct. For example, the amicus brief of the American Antitrust Institute in the *PeaceHealth* appeal argues that a rule of reason analysis should be applied to bundled discounts because even above-cost pricing can result in the exclusion of equally-efficient or less-efficient competitors that have an important role in constraining exclusionary conduct by the monopolist.[89] Under the proposed "structured rule of reason analysis," as set forth in the D.C. Circuit Court's opinion in *Microsoft*, the plaintiff would bear the initial burden of demonstrating that the monopolist's bundling arrangements resulted in harm to competition. The burden then shifts to the defendant to put forth evidence of the procompetitive efficiencies of the arrangements. Then the burden shifts back to the plaintiff to demonstrate that, on balance, the harm to competition outweighs the proffered procompetitive benefits. However, even the AAI recognized that the price-cost test of the variety later accepted by the *PeaceHealth* court could assist in determining whether plaintiff met its initial burden of demonstrating anticompetitive effects.[90]

---

85 *See, e.g.*, *id.* at 94 ("Large and small firms, incumbents, and new entrants use bundled discounts and rebates in a wide variety of industries and market circumstances.").

86 *See, e.g.*, M. Laurence Popofsky & Adam J. Gromfin, *Bundled Discounting: From LePage's to PeaceHealth, & Beyond*, 9 SEDONA CONF. J. 99, 109 and n.72 (2008).

87 *See, e.g.*, Antitrust Modernization Commission Report and Recommendations, at 91 (2007) (noting that "bundling" is an area of Section 2 law where "there is currently a lack of clear and consistent standards").

88 *See* Popofsky & Gromfin, *supra* note 86, at 100 (noting that this position was advocated by plaintiffs in the *PeaceHealth* appeal to the Ninth Circuit and "has found no support in the literature or courts").

89 Brief for American Antitrust Institute as Amici Curiae Supporting Appellants at 55-22, *Cascade Health Solutions v. PeaceHealth,* 515 F.3d 883 (9th Cir. 2008) (Nos. 05-36153 & 05-36202).

90 *Id.* at 22.

A similar approach, recently advocated by Economides and Lianos, calls for the application of a foreclosure analysis to all forms of bundling and tying.[91] They argue that this uniform approach will afford consistency in application and eliminate the tension between the discount attribution approach applied to bundling and the anticompetitive foreclosure analysis applied to all other forms of exclusionary practice.

Their proposed test would focus on anticompetitive foreclosure and the absence of objective justifications – much like the structured rule of reason approach promoted by the AAI. In this model, the distinct product element of the tying test is "reconsidered" and the coercion element of is eliminated. Economides and Lianos reject application of the price-cost test to bundling practices, primarily because the price-cost test can permit predation against less efficient competitors who still may constrain price, and can neglect the benefits to consumers in the form of lower prices and higher investment in quality and variety. They argue that the anticompetitive foreclosure test corrects a number of economic deficiencies caused by the price-cost test, including the difficulty of designating an appropriate measure of cost, disregard for the effect of bundled discounting on rivals' costs, and the difficulty of assessing the appropriate "discount" to be attributed to the competitive product(s).[92]

However, there is widespread disagreement about the competitive benefits of protecting less efficient competitors, particularly given the difficulties and burden of applying a full-blown rule of reason analysis. Noting that the price-cost standard would permit bundled discounts that exclude a less efficient competitor that provides some competitive restraint on the monopolist's pricing, the Antitrust Modernization Committee concluded that "[t]he difficulties of assessing such circumstances, the lack of predictability and administrability in any standard that would capture such instances, and the undesirability of a test that would protect less efficient competitors, however, counsel against the adoption of a screen that protects less efficient competitors."[93]

Moreover, even though there may be some benefits to consistency in analytical approach and evaluating each case through the lens of all the competitive factors at play, there are significant costs associated with applying a foreclosure or rule of reason standard to all forms of bundling conduct. The effort and expense required to bring and defend a lawsuit involving a rule of reason approach is massive and costly, largely due to the highly fact-intensive nature of the inquiry.[94] Closely related to the high cost of rule of reason litigation is the high degree of uncertainty associated with it. This serves to chill procompetitive pricing behavior, both by making it next to impossible for sellers to know *ex ante* what practices will be determined later to be illegal once all the evidence of marketplace conditions and rivals' cost structures are evaluated, and by raising the potential cost to them of defending such a suit.

These criticisms were articulated in response to the Third Circuit's decision in *LePage's*, which produced a wide range of commentary by courts and authorities. For example, the Antitrust Modernization Commission concluded that "[t]he lack of clear standards regarding bundling, as reflected in *LePage's v. 3M*, may discourage conduct that is procompetitive or competitively neutral and thus may actually harm consumer welfare."[95]

---

91 Nicholas Economides & Ioannis Lianos, *The Elusive Antitrust Standard on Bundling in Europe and in the United States in the Aftermath of the Microsoft Cases*, 76 ANTITRUST L.J. 483, 510-534 (2009).
92 *Id.* at 511-516.
93 Antitrust Modernization Commission Report and Recommendations, at 100 (2007).
94 For additional analysis of the high cost of rule of reason litigation, see Maurice E. Stucke, *Does the Rule of Reason Violate the Rule of Law?*, 42 U.C. DAVIS L. REV. 1375, 1460-66 (2009).
95 Antitrust Modernization Commission Report and Recommendations, at 94 (2007). "Because the court failed to evaluate whether 3M's program of bundled rebates represented competition on the merits, its decision offers no clear standards by which firms can assess whether their bundled rebates are likely to pass antitrust muster. Therefore, the Third Circuit's decision is likely to discourage firms from offering procompetitive bundled discounts and rebates to consumers." *Id.*

For these reasons, there are enormous benefits to a standard that provides a safe harbor from rule of reason analysis for pricing behavior that is very rarely likely to harm competition.

D. Benefits of a Diversified Approach

As described above, U.S. courts have applied different standards to different forms of bundling/tying arrangements, recognizing their respective potential for anticompetitive harm and the need for clear and applicable standards that strike a balance between overdeterrence and underdeterrence. Although this area of competition analysis is still evolving and lacks consensus, the Antitrust Modernization Committee has endorsed different standards for bundling and tying[96] and a growing number of commentators have concurred.

> Where price alone is the mechanism that compels purchases, courts confront lower prices that presumptively benefit consumers and reflect efficiencies – presumptions that should not be disregarded when the risk of false positives is inevitably high. In those cases, a high bar or broad safe harbor makes sense, and it is appropriate (indeed, necessary) to apply a derivative of [the price-cost test]. Because price discounting generally benefits consumers, the risk of false positives compels the application of a test that condemns only the rarest and most narrow of practices that would exclude an equally efficient competitor. Where, however, something more than price discounting is at issue, the risk of false positives dissipates and there is the need to balance foreclosure concerns against claimed efficiencies. Safe harbors are inappropriate in those situations; in place of a predatory pricing test, the more traditional Rule of Reason approach amplified by the D.C. Circuit in *Microsoft* should apply.[97]

A rational way to balance these various interests and goals is to evaluate different bundling/tying practices in three general categories.

1. Bundled Discounts

The first category of conduct involves the type of pure bundled discounts that are not based on a commitment with respect to overall requirements, market share or loyalty targets, or exclusivity over a period of time. For these types of discounts, the discount attribution test articulated by the Ninth Circuit in *PeaceHealth* strikes an appropriate balance for distinguishing between practices that are not likely to exclude efficient rivals and those that may have some exclusionary impact and should be further evaluated with respect to their impact on competition. As the Ninth Circuit put it, this test is "safer for consumers and our competitive economy" because it avoids the concern articulated by the Supreme Court in *Weyerhaeuser*: that "mistaken findings of liability would 'chill the very conduct the antitrust laws are designed to protect.'"[98]

---

96 *Id.* at 114 n.157 ("The Commission is not recommending application of this test outside the bundled pricing context, for example in tying or exclusive dealing cases.").
97 Popofsky & Gromfin, *supra* note 86, at 109-10.
98 *Cascade Health Solutions v. PeaceHealth*, 515 F.3d 883, 903 (9th Cir. 2008) (citing *Weyerhaeuser Co. v. Ross-Simmons Hardwood Lumber Co.*, 549 U.S. 312, 320 (2007).

The relative merits of this approach have been endorsed by the Antitrust Modernization Committee and numerous other commentators. "A test that compares incremental revenues with incremental costs . . . offers the most promising source of an economically sensible and administrable safe harbor for bundled rebates or discounts."[99] Noting that the price-cost test can operate as a "safe harbor," the Antitrust Modernization Committee suggested that this screen "would provide sufficient clarity to enable businesses to determine whether a particular bundled discount would be 'screened out' from further scrutiny."[100]

Bundling practices that fall outside the initial price-cost safe harbor may be further evaluated in a rule of reason-type analysis to measure their impact on competition and potential procompetitive benefits.

2. Tying

A second category of "bundled" arrangements (sometimes referred to as "contractual tying") involves the classic tying situation where a seller with market power requires customers to purchase two or more products together. As noted above, U.S. courts have traditionally treated tying arrangements as *per se* unlawful under Section 1 of the Sherman Act because they are almost always deemed to harm competition with no countervailing procompetitive justification. Those standards have evolved in *Microsoft* and other cases to permit the defendant in certain circumstances to demonstrate the efficiencies and procompetitive benefits resulting from the tie. The "separate products" and "coercion" elements of the tying standard serve to segregate conduct that may be procompetitive, and thus more appropriately analyzed pursuant to a more lenient standard, from conduct that is unlikely to be worthy of more extensive analysis. Although there is the possibility for over deterrence and courts continue to evolve in their approaches to tying cases, this standard aims to prohibit the most egregious forms of "bundling" – where customers are forced to purchase products they might otherwise choose to obtain from competitive sources – and provides predictability for practitioners and administrability for courts.

An unsettled question posed by the Ninth Circuit, the Antitrust Modernization Committee Report and others is whether the price-cost test can or should play any role in the tying analysis. In reversing the district court's grant of summary judgment for PeaceHealth on plaintiff's tying claim, the Ninth Circuit remanded the issue to the district court to evaluate "whether, to establish the coercion element of a tying claim through a bundled discount, McKenzie must prove that PeaceHealth priced below a relevant measure of cost."[101] Professors Areeda and Hovenkamp also have suggested that the price-cost test can be employed in a traditional tying analysis, either to determine whether the "separate product" element of the tying standard is met or in evaluating the element of "coercion."[102] Those issues are worthy of further study and consideration.

---

99 Antitrust Modernization Commission Report and Recommendations, at 95 (2007).

100 *Id.* at 100. Note that the Antitrust Modernization Committee also advocated two other "prongs" to its proposed test for bundled discounting: (1) that "the defendant is likely to recoup" short term losses incurred through the offering of bundled discounts and (2) that "the bundled discount or rebate program has had or is likely to have an adverse effect on competition." *Id.* at 99. The *PeaceHealth* court rejected the application of these additional prongs, holding that the recoupment element is unnecessary because "exclusionary bundling does not necessarily involve any loss of profits for the bundled discounter," and that a separate requirement of competitive effect "is redundant because it is no different than the general requirement of 'antitrust injury' that a plaintiff must prove in any private antitrust action." *PeaceHealth*, 515 F.3d at 910 n.21 (internal citations omitted).

101 *PeaceHealth*, 515 F.3d at 916.

102 *See* AREEDA & HOVENKAMP, 32 & ¶ 749b2(B), at 257 (Supp. 2007).

3. Exclusive Dealing and Other Forms of Bundling

Bundling practices that do not fall within either of the categories described above may present the greatest challenge from an analytical perspective. Most often these arrangements involve discounts or rebates that are not tied to a single sale, but involve an incentive or commitment for the purchaser to buy all or a large portion of its requirements across several product lines, typically over a period of time, from a seller who has market power in at least one product in the bundle. These may take the form of market share or loyalty agreements, and often the discounts/rebates are tiered and apply to all purchases within a defined reference period.

The concern generally articulated with respect to these types of arrangements is that they foreclose competition by acting as *de facto* exclusive agreements. Buyers are not "coerced" into purchasing the bundle of products in the same way as the classic tying case, but the economic incentives are such that exclusivity is almost assured. Accordingly, the standards applied to exclusive agreements may be most appropriate. The fact-finder will be concerned with assessing whether the arrangements operate as *de facto* exclusive agreements, the degree of market foreclosure that results, and the competitive impact of that foreclosure weighed against any procompetitive justifications.[103] In these cases, a rule of reason/foreclosure analysis may be best suited to determine the unique competitive effects of the arrangement, notwithstanding the costs, burdens and uncertainty associated with that standard.

V. CONCLUSION

A multiple-standard approach to cases of tying and bundled discounting may strike the most appropriate balance between efficiency for courts and the parties, predictability and certainty for sellers and buyers, and deterrence of the most harmful forms of anticompetitive behavior. However, given the relatively limited experience courts have with bundled discounting, and the recent evolution of tying standards to incorporate some elements of the rule of reason test, no standard should be so hidebound as to preclude change as courts and regulators gain more experience. As has been noted by Professors Areeda and Hovenkamp, a similar rationale underlay courts' early decision to apply below-cost pricing standards to cases of predatory pricing:

> It is one thing to develop a theory showing that a particular practice can be anticompetitive. It is quite another to show that this theory explains a particular practice without producing an unacceptably high number of false positives. In the case of predatory pricing, the result of this concern was the development of tests requiring prices to be below marginal cost or average variable cost. . . . The reason these tests for predatory pricing were adopted was *not* because there is widespread consensus that above-cost pricing strategies can never be anticompetitive in the long run. Rather, it is because our measurement tools are too imprecise to evaluate such strategies without creating an intolerable risk of chilling competitive behavior.[104]

---

103 For an exposition of the analysis applied to exclusive dealing cases, see *United States v. Dentsply Int'l, Inc.*, 399 F.3d 181, 187 (3d Cir. 2005).

104 AREEDA & HOVENKAMP, ¶ 749, at 144 (Supp. 2008).

As has been the case in the predatory pricing context, it can be expected that the experience and sophistication of the courts, regulators, academics and practitioners will grow and evolve over time.[105] As that takes place, it remains possible that a full anticompetitive foreclosure analysis under the rule of reason will become more efficient and afford more certainty. Until then, the relative benefits of differentiated standards for different types of tying and bundled discounting should guide the way.

105 *See* Antitrust Modernization Commission Report and Recommendations, at 100 (2007) (encouraging further empirical economic research in this area, suggesting that "[t]he courts, the antitrust agencies, and antitrust practitioners generally would benefit from a more thorough and empirically based understanding of the likely competitive effects of bundled discounts in a variety of settings").

THE PLAUSIBILITY OF PLEADINGS AFTER *TWOMBLY* AND *IQBAL*

Robert D. Owen and Travis Mock[1]
Fulbright & Jaworski LLP
New York, NY

The Supreme Court's pleadings standards decisions in *Bell Atlantic v. Twombly*[2] and *Aschcroft v. Iqbal*[3] have ignited a firestorm of judicial and academic analysis. *Twombly* is already one of the 20 most cited cases of all time in the federal courts, and *Iqbal* averages over 300 new citations per month.[4] But this abundance of analysis has so far failed to coalesce around a concrete and workable interpretation of the "plausibility standard" introduced by these two important decisions.

A review of the analysis to date reveals a broad range of theories and narratives, which often appear to be shaped by the authors' pre-existing beliefs about the proper role of pleadings in federal civil litigation. However, a brief look at some of the key cases and academic analysis can highlight the primary areas of confusion and conflict to focus the analysis and enable practitioners to negotiate these new uncertainties.

I. *TWOMBLY* AND *IQBAL*

Bell Atlantic v. Twombly involved a class action antitrust claim against the so-called "Baby Bells," massive telecommunications companies known as Incumbent Local Exchange Carriers (ILECs).[5] The complaint alleged that the Baby Bells violated the Sherman Act § 1 by engaging in anticompetitive parallel conduct.[6] The plaintiff class encompassed approximately 90 percent of all subscribers of local telephone and high-speed Internet service.[7] The Southern District of New York dismissed the complaint for failure to state a claim, and the Second Circuit reversed.[8] In a 7-2 decision, the Supreme Court reversed, holding that the allegations of parallel conduct, without more, were insufficient to sustain a claim under the Sherman Act.[9] The Court noted that the Sherman Act bans only anticompetitive conduct that is the result of "a contract, combination, or conspiracy."[10] Parallel conduct may be consistent with such illegal behavior, but it is "just as much in line with a wide swath of rational and competitive business strategy unilaterally prompted by common perceptions of the market."[11] Thus, the Court held that "an allegation of parallel conduct and a bare assertion of conspiracy will not suffice."[12] Rather,

---

1 Robert D. Owen is admitted to practice in New York and is a partner of Fulbright & Jaworski L.L.P. Travis Mock is an associate in the New York office of Fulbright & Jaworski L.L.P. awaiting admission to the New York bar.
2 550 U.S. 544 (2007).
3 129 S.Ct. 1937 (2009).
4 *See* Adam Steinman, *The Pleading Problem*, 62 STANFORD L. REV. 1293, 1360 (May 2010).
5 550 U.S. at 549.
6 *Id.* at 548.
7 *Id.* at 559.
8 *Id.* at 552-53.
9 *Id.* at 570.
10 *Id.* at 553 (citing *Copperweld Corp. v. Independence Tube Corp.*, 467 U.S. 752, 775 (1984)).
11 *Id.* at 554.
12 *Id.* at 556.

the allegations of a complaint must "plausibly suggest" (not be "merely consistent with") illegal conduct.[13] Because plaintiffs had not provided the "further factual enhancement" necessary to "nudge[] their claims across the line from conceivable to plausible, their complaint must be dismissed."[14]

The *Twombly* Court also addressed the potential for oppressive litigation in antitrust cases, noting that the mere threat of such discovery can be used to coerce settlement, even with weak claims.[15] For a time, courts and commentators attempted to limit the *Twombly* holding to antitrust cases or complex litigation involving a similar risk of expansive discovery.[16] The Court's per curiam decision in *Erickson v. Pardus*,[17] which was issued two weeks after *Twombly*, seemed to validate the narrow interpretation of *Twombly* by applying traditional notice pleading principles to a prisoner's Eighth Amendment claim of improper medical treatment, without any mention of *Twombly* or its plausibility standard.[18]

The Court explained and expanded *Twombly* in *Ashcroft v. Iqbal*. In *Iqbal*, a Pakistani man brought a Bivens action against former Attorney General Ashcroft and former FBI Director Mueller alleging that he was detained under harsh conditions after September 11.[19] The plaintiff further alleged that his detention was the result of an "unconstitutional policy that subjected [Iqbal] to harsh conditions of confinement on account of his race, religion, or national origin."[20] The Eastern District of New York denied defendants' motion to dismiss, and the Second Circuit affirmed.[21] In a 5-4 decision, the Supreme Court reversed.[22] The Court held that Iqbal's general factual allegations of a policy of restrictive confinement for detainees of high interest in the wake of the September 11 attacks did not sufficiently allege the individual discriminatory intent required to sustain a Bivens action.[23] The Court also explicitly applied *Twombly* to "all civil actions and proceedings in the United States district courts."[24]

The *Iqbal* Court further explained *Twombly* by establishing a two-part test for evaluating complaints under the plausibility standard. First, it invited the district courts to identify and set aside wholly conclusory allegations, which are not entitled to the presumption of truth.[25] The Court alternately described these wholly conclusory allegations as "labels and conclusions," "naked assertions," "formulaic recitation of the elements of a cause of action," or "unadorned, the-defendant-harmed-me accusation[s]."[26] Then, the Court instructed the district courts to examine whether the remaining allegations, accepted as true, plausibly state a claim for relief.[27]

Like *Twombly*, the *Iqbal* decision was also rooted in concerns over discovery. Specifically, the Court voiced concern that intrusive discovery procedures would interfere with vital government functions.[28] However, the Court again declined to respond to this

13 *Id.* at 557.
14 *Id.* at 557, 570.
15 *Id.* at 558-59.
16 *See* Steinman, 62 STANFORD L. REV. at 1305 (noting the interpretations limiting the application of *Twombly*).
17 551 U.S. 89 (2007).
18 *Id.*
19 *Iqbal*, 129 S.Ct. at 1943.
20 *Id.* at 1942.
21 *Id.* at 1942.
22 *Id.*
23 *Id.* at 1952.
24 *Id.*
25 *Id.* at 1950.
26 *Id.* at 1949-50.
27 *Id.* at 1950.
28 *Id.* at 1953.

concern by arming district courts with more effective ways to manage discovery.[29] The Court's indirect approach to addressing discovery concerns has created uncertain effects on the role and scope of discovery in civil litigation in federal courts. But before we can begin to consider *Twombly* and *Iqbal*'s effects on discovery, we must first examine how this new plausibility standard is being applied in the lower courts.

II. THE CIRCUIT COURTS ADDRESS THE "PLAUSIBILITY STANDARD"

The circuit courts have largely taken *Twombly* and *Iqbal* in stride, but there are significant and problematic differences of interpretation over several key questions.

A. Does **Conley***'s Concept of Notice Pleading Still Exist?*

The courts have been more restrained than some commentators in their analysis of *Twombly* and *Iqbal*. Though their interpretations and applications vary, most courts have gone about the business of incorporating plausibility into more familiar concepts of notice pleading.[30]

The Third Circuit, in contrast, has made dramatic statements asserting the death of notice pleading. However, its opinions appear to be more cautious than its language would suggest. In *Fowler v. UPMC Shadyside*,[31] the Third Circuit addressed *Swierkiewicz v. Sorema N.A.*,[32] a pre-*Twombly* decision in which the Supreme Court reiterated *Conley*'s liberal pleading standard, declaring that complaints need not demonstrate a likelihood of success on the merits or a likelihood of later discovery of evidence to support plaintiff's claims.[33] The Third Circuit summarily declared that "because *Conley* has been repudiated . . . , so too has *Swierkiewicz*, at least insofar as it concerns pleading requirements and relied on *Conley*."[34] The court's application of this pronouncement turned out to be less dramatic. The *Fowler* court upheld the validity of a Rehabilitation Act complaint, finding it was sufficient that the plaintiff pleaded an impairment and alleged that the impairment was a disability under the Rehabilitation Act, that the defendant hospital was aware of the impairment, and that the hospital failed to accommodate the impairment.[35] In other words, "[t]he complaint pleads how, when, and where [defendant] allegedly discriminated against [plaintiff].[36]

B. Do **Twombly** *and* **Iqbal** *Establish a Heightened Fact Pleading Requirement?*

The Sixth Circuit has interpreted *Twombly* and *Iqbal* as introducing an elevated pleading standard. In *Hensley Mfg. v. ProPride, Inc.*,[37] the court upheld the dismissal of a trademark action because the complaint did not "allege facts sufficient to show that ProPride's use of the 'Hensley' name create[d] a likelihood of confusion as to the source of the products."[38] This elevated standard resembles the summary judgment standard. The *Iqbal* Court did appear at times to engage in an elevated, probability-type analysis.[39]

29 *Id.*
30 *See Riley v. Vilsck*, 665 F. Supp. 2d 994, 1001 (W.D. Wis. 2009) (remarking that "little changed in this circuit as a result of *Twombly*," and that the Seventh Circuit has continued to adhere "to the view that [Federal Rule of Civil Procedure 8] required nothing more than "fair notice."); *Brooks v. Ross*, 578 F.3d 574, 581 (7th Cir. 2009) (maintaining that *Erickson* "put to rest" any "doubt that *Twombly* had repudiated the general notice-pleading regime of Rule 8").
31 578 F.3d 203 (3d Cir. 2009).
32 534 U.S. 506 (2002).
33 *Id.* at 511-12, 515.
34 *Fowler*, 578 F.3d at 211.
35 *Id* at 212.
36 *Id.*
37 579 F.3d 603 (6th Cir. 2009).
38 *Id.* at 610.
39 *See, e.g.*, *Iqbal*, 129 S.Ct. at 1951.

However, the Sixth Circuit's broad reading seems at odds with the Supreme Court's explicit rejection of any requirements of evidentiary support or calculations of probability of success at the motion to dismiss stage of the case.[40]

The Fifth Circuit has exhibited more nuance in its approach. In *Floyd v. City of Kenner, Louisiana*,[41] the court rejected assertions that plaintiff must provide evidentiary support at the pleading stage. "At a later stage, [plaintiff] will be required to produce specific support for his claim of unconstitutional motive. But at the pleading stage, his allegation that [defendant's] actions were spurred by [an associate's] ill will suffices."[42] The court affirmed the dismissal of plaintiff's claims against a different defendant, observing that "[u]nlike his allegations [against the first defendant], this bare assertion does not provide any detail about what [defendant, in his official capacity] *did*.[43] Here, the problem was not pleading of insufficient facts, but failure to plead any facts that described the defendant's wrongful actions.

The Eight Circuit has maintained a liberal pleading standard, as exemplified by its decision in *Braden v. Wal-Mart Stores, Inc.*[44] The *Braden* court reversed the district court's dismissal of plaintiff's claim for violation of fiduciary duties under the Employee Retirement Income Security Act (ERISA).[45] The court emphasized that the plausibility standard does not change the rule that plaintiffs are entitled to all reasonable inferences supported by the facts alleged.[46] In addition, it is improper for courts to draw inferences in the defendant's favor, thereby faulting the plaintiff for failing to plead facts tending to contradict those inferences.[47] The court specifically rejected the idea that plaintiffs must plead specific facts regarding the ways in which they were wronged by defendant.[48] Rather, "indirect facts showing unlawful behavior" are sufficient, as long as they give notice and allow reasonable inferences to be drawn in plaintiff's favor that show entitlement to relief.[49] The court further observed that at the plausibility stage, the complaint is to be viewed as a whole, not as individual allegations.[50] This broad view of the plausibility standard may be the most like traditional notice pleading of any of the circuit interpretations.

As noted above, the Seventh Circuit also takes a limited view of *Twombly* and *Iqbal*. Nevertheless, in what is perhaps intended to be a compromise of sorts, the Seventh Circuit has adopted a sliding scale for pleading standards.[51] However, "the plausibility standard has its most force when special concerns exist about the burden of litigation on the defendant or when the theory of the plaintiff seems particularly unlikely."[52] In addition, although the *Smith* court found that the plaintiff's fraud claim had "no merit" and was dismissible "under any reasonable interpretation of Rule 12(b)(6)," the court implied that *Twombly* and *Iqbal* may not apply to all cases.[53] This holding is curious, given that *Iqbal* expressly applied *Twombly* to "all civil actions."[54]

---

40 *See Twombly*, 550 U.S. at 569 n.14 ("In reaching this conclusion, we do not apply any 'heightened' pleading standard . . . ."); *Iqbal*, 129 S.Ct. at 1949 ("The plausibility standard is not akin to a 'probability requirement' . . . .").
41 2009 WL 3490278 (5th Cir. Oct. 29, 2009).
42 *Id.* at *5.
43 *Id.* at 8 (emphasis added).
44 588 F. 3d 585 (8th Cir. 2009).
45 *Id.* at 603.
46 *Id.* at 595.
47 *Id.*
48 *Id.* (citing *Erickson*, 551 U.S. at 93.)
49 *Id.*
50 *Id.* at 594 (citing *Vila v. Inter-Am. Inv. Corp.*, 570 F.3d 274, 285 (D.C. Cir. 2009)).
51 *Cooney v. Rossiter*, 583 F.3d 967, 971 (7th Cir. 2009) ("[T]he height of the pleading requirement is related to circumstances."). "[I]n the ordinary case, the burden remains low." *Riley*, 665 F. Supp. 2d at 1003-04.
52 *Id. See also Smith v. Duffey*, 576 F.3d 336 (7th Cir. 2009) (distinguishing *Twombly*'s complex factual landscape from the case before the court).
53 *Id.* at 340 (noting that "maybe neither *Bell Atlantic* nor *Iqbal* governs here").
54 *Iqbal*, 129 S.Ct. at 1951.

In contrast, it appears that at least some courts are taking seriously *Iqbal*'s instruction that *Twombly* applies to all civil cases. The Western District of Virginia recently dismissed a slip-and-fall claim because it failed to satisfy the plausibility standard.[55] It was not enough, the court reasoned, for the plaintiff to plead that she slipped on liquid on the store's floor.[56] To survive a motion to dismiss, the plaintiff must also plead that the owner of the store caused the liquid to be on the floor or had actual or constructive notice of the wet floor and that the owner failed to either remove the liquid in a reasonable time or to warn the plaintiff.[57]

C. What Does "Plausibility" Really Mean?

The First Circuit appears to have embraced a highly subjective and expressly comparative version of plausibility. In *Chao v. Ballista*,[58] the Massachusetts District Court upheld a complaint for sexual abuse of a prison inmate by a guard. The court distinguished the required showing for qualified immunity in *Chao* from that in *Iqbal*.[59] However, in so holding, the court announced a surprising interpretation of the plausibility standard. The court held that an allegation is conclusory when it "recites only the elements of the claim and, at the same time, the court's commonsense credits a far more likely inference from the available facts."[60] The court's subjective inquiry is highly context-specific and "depends on the full factual picture, the particular cause of action, and the available alternative explanations."[61] However, "a complaint should only be dismissed at the pleading stage where the allegations are so broad, and the alternative explanations so overwhelming, that the claims no longer appear plausible."[62] This formulation incorporates directly into the legal standard *Iqbal*'s invitation for courts to employ their experience and common sense. This is not particularly controversial, but the court's full-throated endorsement of a comparative element resembles the probability analysis that the *Iqbal* Court expressly rejected.[63] It will likely be small comfort to plaintiffs, who must now refute myriad alternative explanations of the case at the motion to dismiss stage, that those alternatives must be "far more likely" and "so overwhelming."[64]

The Second Circuit, in contrast, has heard cases factually similar to both *Twombly* and *Iqbal* and has taken a more conservative approach to plausibility. In *Starr v. Sony*,[65] the court upheld a Sherman Act antitrust complaint. The court distinguished the facts of the case before it from those in *Twombly*, noting that the *Twombly* complaint based its claims of illegal antitrust activity purely on the presence of parallel dealing among the defendant phone companies.[66] In contrast, the Second Circuit noted that the *Starr* complaint also alleged facts regarding the underlying agreement between the defendants.[67] The court found the complaint to be plausible, even though plaintiff did not allege specific dates or times that the defendants' conspiracy supposedly took place. The context of the defendant's parallel actions raised the suggestion of illegal behavior, because the defendants' parallel conduct would have been harmful to their individual interests absent an agreement to act in

55 *Branham v. Dolgencorp, Inc.,* Civil No. 6:09-CV-00037 (W.D. Va. Aug. 24, 2009).
56 *Id.24,*
57 *Id.*
58 630 F. Supp. 2d 170 (D. Mass. 2009).
59 *Id.* at 178, n.2.
60 *Id.* at 177 (citing *Maldonado v. Fontanes*, 568 F.3d 263, 268 (1st Cir. 2009)).
61 *Id.*
62 *Id.* (citing *Thomas v. Rhode Island*, 542 F.3d 944, 948 (1st Cir. 2008)).
63 *See Iqbal*, 129 S.Ct. at 1949 (citing *Twombly*, 550 U.S. at 556).
64 *Chao*, 630 F. Supp. 2d at 177.
65 2010 U.S. App. LEXIS 768 (2d Cir. Jan. 10, 2010).
66 *Id.* at *17.
67 *Id.* at *18-23.

concert.[68] The Second Circuit also decided a case since *Iqbal* that involved Bivens claims against former Attorney General Ashcroft related to post-September 11 law enforcement activity.[69] The Second Circuit rejected the complaint as insufficient under *Iqbal*.[70] The court stated that the plaintiff's passive allegations of conspiracy against undifferentiated plaintiffs did not satisfy the requirement that Bivens actions allege unconstitutional discriminatory intent on the part of each individual defendant.[71] The court's vote to rehear *Arar* en banc, a very rare decision in the Second Circuit, may be an indication of the importance the court placed on establishing a clear pleading standard post-*Iqbal*.[72]

The Ninth Circuit has also adopted a cautious approach to *Twombly* and *Iqbal*. The court has rejected arguments that *Twombly* and *Iqbal* impose a heightened pleading standard, and in *al-Kidd v. Ashcroft*,[73] it upheld a Bivens complaint against former Attorney General Ashcroft that was similar to the complaint in *Iqbal*. The court distinguished *Iqbal*, noting that the complaint included facts that plausibly alleged that the defendant had the requisite knowledge and intent.[74] Importantly, the Ninth Circuit also accepted the Supreme Court's invitation to employ its own subjective expertise in evaluating plausibility. "Drawing on our 'judicial experience and common sense,' as the Supreme Court urges us to do, we find that al-Kidd has met his burden of pleading a claim for relief that is plausible."[75]

In the Tenth Circuit, plausibility relates to the scope of the allegations. "[I]f [the allegations] are so general that they encompass a wide swath of conduct, much of it innocent, then the plaintiffs have not nudged their claims across the line from conceivable to plausible."[76] This formulation finds support in *Twombly*'s treatment of allegations of parallel conduct.[77] Like some other circuit interpretations, the "plausibility as scope" approach seems to impose a heightened burden on the plaintiff to plead facts that not only support its own claims but also exclude all possible competing explanations.

As these cases illustrate, the conclusory nature of allegations and the plausibility of claims are close questions that involve a degree of subjectivity. Courts are likely to come to different conclusions even on similar sets of facts. But a realistic look at notice pleading reflects similar ambiguities at the margins. Therefore, the direction of pleading standards going forward will likely depend more on the content given to the standard than on whether the standard is inherently liberal or restrictive.[78]

D. Carve-Outs From Plausibility Standard

Courts have declined to extend *Iqbal* and *Twombly* to several areas.

In *Mitchell v. Federal Bureau of Prisons*,[79] the Court of Appeals for the District of Columbia declined to extend *Iqbal* to a case in which the court sought to determine

68 *Id.* at 20-21, 32.
69 *Arar v. Ashcroft*, 585 F.3d 559 (2d Cir. 2009) (*en banc*).
70 *Id.* at 563.
71 *Id.* at 569.
72 *See* Michael B. de Leeuw & Samuel P. Groner, *En Banc Review In the Second Circuit*, 242 N.Y. L. J. 115 (Dec. 18, 2009).
73 580 F.3d 949 (9th Cir. 2009).
74 *Id.* at 975 ("Here, unlike Iqbal's allegations, al-Kidd's complaint 'plausibly suggest[s]' illegal conduct, and does more than contain bare allegations of an impermissible policy." (quoting *Iqbal*, 129 S.Ct. at 1950)).
75 *Id.* at 978 (quoting *Iqbal*, 129 S.Ct. at 1950).
76 *Robbins v. Oklahoma*, 519 F.3d 1242, 1247 (10th Cir. 2008) (pre-*Iqbal)*; *Phillips v. Bell*, 2010 WL 517629, at *5 (10th Cir. Feb. 12, 2010) (post-*Iqbal*).
77 *See Twombly*, 550 U.S. at 554.
78 *See* Steinman, 62 STANFORD L. REV. at 1324.
79 587 F.3d 415 (D.C. Cir. 2009).

whether a prisoner's *in forma pauperis* claims satisfied the imminent danger exception to the Prison Litigation Reform Act.[80] Following *Ibrahim v. District of Columbia*,[81] the court held that it would accept all of the plaintiff's allegations as true, regardless of whether they were conclusory or plausible.[82] The court held that Iqbal had "no applicability to [*in forma pauperis*] proceedings where we are exercising our discretion to grant or withhold a privilege made available by the courts. [*In forma pauperis*] proceedings are nonadversarial and implicate none of the discovery concerns lying at the heart of *Iqbal*."[83] The court noted, however, that if IFP status was granted, defendants could then rely on *Iqbal* in seeking to dismiss the underlying complaint.[84] The court's emphasis on the discovery concerns in *Iqbal* resembles the Seventh Circuit's rationale for its sliding scale approach.

In the *NuvaRing* multi-district litigation,[85] the Eastern District of Missouri recently held the master complaint in a multi-district litigation cannot be challenged under *Iqbal*. *Iqbal* does not change the precedent that master complaints are "administrative tools" not intended to be subject to motions to dismiss aimed at dismissing the entire action.[86]

III. ACADEMIC COMMENTARY

The commentary surrounding *Twombly* and *Iqbal* has been as varied as the judicial interpretation.

The *Twombly* decision generated a great deal of academic criticism alleging that it was overturning decades of precedent and flying in the face of the basic principles of notice pleading and the Federal Rules of Civil Procedure. *Iqbal* has done little to change critics' minds. In fact, *Iqbal*'s introduction of a two-part test and its invitation for judges to apply their "personal experience and common sense" has fueled criticism that may be even harsher than that levied against *Twombly*.[87]

However, scholars have also now had some time to attempt to harmonize the case law, and substantial emerging commentary argues that *Twombly* and *Iqbal* are not inherently radical departures from notice pleading and earlier case law.[88]

Adam Steinman's draft article is particularly interesting. In it, Steinman emphasizes the importance of the sequence of the Supreme Court's two-step analysis in *Iqbal*, noting that it is improper for courts to jump immediately to considerations of plausibility.[89] Rather, courts must first engage in the analysis of whether the allegations in a complaint are conclusory.[90] If the complaint pleads non-conclusory allegations for each element of the claim, the plausibility analysis is unnecessary, since a claim logically must be plausible if all of its allegations are entitled to the assumption of truth.[91] In order to evaluate

80 *Id.*
81 463 F.3d 3, 6 (D.C. Cir. 2006).
82 *Id.* at 420.
83 *Id.* (internal citations omitted).
84 *Id.*
85 *In re Nuvaring Prods. Liab. Litig.,* MDL No. 4:08-md-1964 (E.D. Mo. Aug. 6, 2009).
86 Memorandum and Order, at 3, *In re Nuvaring*, PACER Doc. No. 231.
87 *See Pleading Standards*, 123 HARV. L. REV. 252 (2009); Kenneth S. Klein, Ashcroft v. Iqbal *Crashes Rule 8 Pleading Standards on to Unconstitutional Shores*, 88 NEB. L. REV. 261 (2009); Robert L. Rothman, Twombly *and* Iqbal*: A License to Dismiss*, 35 NO. 3 LITIGATION 1 (2009).
88 *See, e.g.,* Robert G. Bone, *Pleading Rules, and the Regulation of Court Access*, 94 IOWA L. REV. 873 (2009) (asserting that it is incorrect to assume that *Twombly* tightens pleading standards); Edward A. Hartnett, *Taming* Twombly*, Even After* Iqbal, 158 U. PA. L. REV. 473 (2010) (emphasizing *Twombly*'s connection to prior case law); 62 STANFORD L. REV. 1293, n.14 (advancing a theory of "transactional pleading").
89 *Id.* at 24.
90 *Id.*
91 *Id.* at 26.

the conclusory nature of allegations, Steinman advances a theory he calls "plain pleading," in which the facts in a complaint focus on the transaction underlying the claim.[92] Steinman argues that both notice pleading and plausibility pleading are different methods of describing pleading the underlying transaction.[93] This theory of pleading may go a long way toward creating continuity throughout the evolving pleading standards.[94]

IV. The Impact on Discovery

The plausibility standard's impact on discovery is unclear. Although the *Twombly* Court voiced concern about the cost and abuses of discovery, the Court's "gatekeeper" remedy of addressing pleading standards only impacts discovery issues indirectly.

Nevertheless, the plausibility standard may moderate discovery in a few important ways. Most obviously, some discovery will be eliminated when complaints that fail to meet the plausibility standard are dismissed. In the case of *Twombly*, for example, discovery involving 90 percent of the local telephone subscribers in the nation was rendered unnecessary by the dismissal. As commentators have noted, however, this benefit comes at a cost. As an initial matter, the plausibility standard will create a roadblock for at least some meritorious claims, particularly those claims containing an element of conspiracy or scienter. Evidence regarding elements such as these is typically only in the hands of the defendant. Therefore, a heightened pleading standard will guarantee that plaintiff's allegations of those elements will always be conclusory. Therefore, plaintiffs may find themselves inevitably cut off from the very discovery that they need in order to prove up their allegations.[95] Second, the uncertainties of the plausibility standard are likely to increase the amount of litigation at the pleading stage, something that notice pleading was intended to avoid.[96]

The heightened pleading standard may also indirectly limit discovery by incentivizing parties to plead more facts. Since the permissible scope of discovery under Federal Rule of Civil Procedure 26(b) is determined according to the foundation laid by the facts and allegations in the complaint, supplementing allegations with additional facts may subsequently limit plaintiffs to discovery on their narrowed facts. The uncertainty surrounding the plausibility standard will lead cautious plaintiffs to plead more factual detail than necessary, narrowing the scope of potential discovery even further.

On the other hand, by linking discovery management to pleadings standards, *Twombly* and *Iqbal* may have the perverse effect of strengthening some plaintiffs' demands for comprehensive discovery. Instead of reviewing each discovery request on its merits, courts may deem claims that survive the plausibility analysis to be stronger and therefore deserving of full discovery. The Supreme Court's manifest lack of faith in the district courts' ability to control the discovery process further reinforces an all-or-nothing approach to discovery.[97]

---

92 *Id.* at 37-57.
93 *Id.* at 44-49.
94 *See, e.g., Fowler*, 578 F.3d at 212 ("Under the 'plausibility paradigm' . . . these averments are sufficient to give UPMC notice of the basis for Fowler's claim. The complaint pleads how, when, and where UPMC allegedly discriminated against Fowler.") (citiations omitted).
95 *See* Robert Rothman, *Twombly and Iqbal: A License to Dismiss*, 35 No.3 LITIGATION 1, 2 (Spring 2009).
96 *See* Steinman, 62 STANFORD L. REV. at 1355 ("[A] stricter pleading standard has the potential to encourage costly, time-consuming litigation over pleading sufficiency . . . .").
97 *See Twombly*, 550 U.S. at 560 ("Given the system that we have, the hope of effective judicial supervision [of discovery] is slim."); *Iqbal*, 129 S.Ct. at 1953-54 (rejecting the "careful-case-management approach" and finding that it "provides especially cold comfort" in cases involving qualified immunity claims of high-level Government officials).

V. THE ROAD TO RESOLUTION

The cases and literature compiled thus far do not lend themselves to easy reconciliation. It seems likely that the Supreme Court will be required to take up this issue again, if only to attempt to clarify its intentions in *Twombly* and *Iqbal*.

EFFECTIVENESS OF THE 2006 RULES AMENDMENTS

Emery G. Lee III [1]
Federal Judicial Center
Washington, DC

On December 1, 2006, electronic discovery amendments to the Federal Rules of Civil Procedure ("Rules") went into effect. How effective have these amendments been in achieving their goals? At the 12th Annual Sedona Conference® on Complex Litigation, on April 8, 2010, a panel discussed that question. In preparation for that panel, I prepared this brief outline of the legislative history of the goals of the 2006 amendments and pulled together existing information that might assist in the discussion of whether the amendments have, about three years into their existence, achieved those goals. The empirical evidence on the effectiveness of the 2006 amendments is limited, but there is some support for the proposition that the amendments have succeeded, at least in a limited sense, in getting the parties in cases involving electronic discovery to pay early attention to potential problems.

I. GOALS OF THE 2006 AMENDMENTS

To evaluate the effectiveness of the 2006 amendments, it is first necessary to determine what goals they were intended to accomplish. In what follows, I have drawn on the report of the Committee on Rules of Practice and Procedure ("Standing Committee") to the Judicial Conference of the United States ("Judicial Conference"), from September 2005,[2] and from the report of the Advisory Committee on Civil Rules to the Standing Committee, from May 2005.[3] Overall, the 2006 electronic discovery amendments were intended to provide adequate "guidance to litigants, judges, and lawyers in determining discovery rights and obligations in particular cases" and to prevent the development of "a patchwork of [disparate local] rules" addressing electronic discovery issues.[4] In presenting the amendments to the Judicial Conference, the Standing Committee placed the 2006 amendments in the context of the 2000 amendments, which were aimed at reducing costs of discovery, increasing its efficiency, increasing uniformity of practice, and encouraging more active judicial case management, when appropriate.[5] In a sense, the 2006 amendments were intended to fill in the gaps in the discovery rules "to make the rules apply better to electronic discovery problems."[6]

1 Senior Researcher, Federal Judicial Center. Affiliation provided for identification purposes only; views expressed herein are those of the author and do not represent those of the Federal Judicial Center or any other entity in the judicial branch. Tom Willging, Angelia Levy, and Jill Gloekler provided useful comments on the Article.
2 Report of the Committee on Rules of Practice and Procedure to the Judicial Conference of the United States, September 2005 (hereinafter "Standing Report").
3 Report of the Advisory Committee on Civil Rules to the Committee on Rules of Practice and Procedure, May 2005 (revised July 2005) (hereinafter "Advisory Report").
4 Standing Report, *supra* note 2, at 23.
5 *Id.* at 24 ("The proposed amendments to make the rules apply better to electronic discovery share the same focus.")
6 *Id.* The phrase "fill in the gaps" was suggested by my colleague Tom Willging, who was Federal Judicial Center liaison to Civil Rules throughout this period.

Both reports also list five sets of amendments with more specific purposes. First, the amendments to Rules 16, 26(a), and (f) were to "present a framework for the parties and the court to give early attention to issues related to electronic discovery," especially "frequently-recurring" preservation and privilege issues.[7]

- "Rule 16 is amended to invite the court to address the disclosure or discovery of electronically stored information in the Rule 16 scheduling order."[8]

- "The proposed amendment to Rule 26(a) clarifies a party's duty to include in its initial disclosures electronically stored information . . . ."[9]

- Finally, under this set of amendments, the 26(f) conference is "to include discussion of any issues relating to disclosure or discovery of electronically stored information," including "the form of produc[tion], a distinctive and recurring problem," and preservation issues.[10]

Second, the proposed amendments to Rules 33 and 34 (referred to as "discovery workhorses"[11]) are meant to "clarify" how they apply to electronically stored information.

- "The proposed amendment to Rule 33 clarifies that a party may answer an interrogatory involving review of business records by providing access to the information if the interrogating party can find the answer as readily as the responding party can."[12]

- "Under the proposed amendment to Rule 34, electronically stored information is explicitly recognized as a category subject to discovery that is distinct from 'documents' and 'things.'"[13]

- "Rule 34 is also amended to authorize a requesting party to specify the form of production . . . and for the responding party to object. Under the proposed amended rule, absent a court order, party agreement, or a request for a specific form of production, a party may produce responsive electronically stored information in the form in which the party ordinarily maintains it or in a reasonably usable form. Absent a court order, a party need only produce . . . in one form."[14]

- Rule 45 and Form 35 were also amended to conform to the proposed changes.[15]

Third, the proposed amendment to Rule 26(b)(5) was intended to deal with assertions of privilege after inadvertent production.

- "The proposed amendment to Rule 26(b)(5) provides a procedure for asserting privilege after production . . . . [T]he volume of electronically stored information searched and produced in response to discovery can be enormous,

7 *Id.* at 26.
8 *Id.*
9 *Id.*
10 *Id.*
11 *Id.* at 27.
12 *Id.*
13 *Id.* at 28.
14 *Id.*
15 *Id.*

and certain features of the forms in which [it] is stored make it more difficult to review for privilege and work-product protection than paper. The inadvertent production of privileged or protected material is a substantial risk."[16]

- "By providing a clear procedure to allow the responding party to assert privilege after production, the amendment helpfully addresses the parties' burden of privilege review, which is particularly acute in electronic discovery."[17]

Fourth, the proposed amendment to Rule 26(b)(2) was amended to "clarif[y] the obligation of a responding party . . . [when] information is not reasonably accessible, an increasingly disputed aspect of such discovery."[18]

- "The proposed amendment to Rule 26(b)(2) responds to distinctive problems encountered in discovery of electronically stored information that has no close analogue in the more familiar discovery of paper documents. . . . [S]ome forms of computer storage make it very difficult to access, search for, and retrieve information."[19]

- "Lawyers sophisticated in these problems are developing a two-tier practice in which they first obtain and examine the information that can be provided from easily accessed sources and then determine whether it is necessary to search the difficult-to-access sources."[20]

- "The Rule 26(b)(2)(B) proposal authorizes a party to respond to a discovery request by identifying sources of electronically stored information that are not reasonably accessible because of undue burden or cost. If the requesting party seeks discovery from such sources, the responding party has the burden to show that the sources are not reasonably accessible. Even if that showing is made, the court may order discovery if . . . the requesting party shows good cause. The court may specify conditions for discovery."[21]

Fifth, the proposed amendment to Rule 37(f)[22] was intended to "provide guidance in a troublesome area distinctive to electronic discovery[23]… the application of sanctions rules in a narrow set of circumstances distinctive to electronic discovery."[24]

- "The proposed amendment provides limited protection against sanctions under the rules for a party's failure to provide [ESI] in discovery. . . . [A]bsent exceptional circumstances, sanctions may not be imposed . . . if [ESI] sought in discovery has been lost as a result of the routine operation of an electronic information system, as long as the operation is in good faith."[25]

16 *Id.* at 29.
17 *Id.* at 30.
18 *Id.*
19 *Id.* at 31.
20 *Id.*
21 Advisory Report, *supra* note 3, at 40. The Advisory Report adds that "[i]n many circumstances, the two-tier approach will be worked out by negotiations." *Id.* at 41.
22 This is now Rule 37(e).
23 Standing Report, *supra* note 2, at 32.
24 Advisory Report, *supra* note 3, at 16.
25 Standing Report, *supra* note 2, at 33.

- "The proposed rule also recognizes that suspending or interrupting these features can be prohibitively expensive and burdensome. . . ."[26]

In sum, the 2006 amendments were intended to clarify the discovery rights and obligations of the parties in cases with electronic discovery and, in so doing, provide guidance to litigants, lawyers, and judges. The Advisory Committee identified recurring problems in such cases—including disputes over the form of production of ESI, the cost of privilege review and the risk of inadvertent production, the preservation of ESI in dynamic information systems, and the risk of sanctions for spoliation—and then crafted amendments intended to provide rules and procedures to address them. The amendments also clearly envision that if the parties (and in some instances, the court) pay early attention to potential problems related to electronic discovery, those problems can be prevented or at least lessened. The goal of preventing inconsistent and conflicting local rules to address these issues is also present in the legislative history.

In providing guidance and clarity, however, the 2006 amendments may "not provide the sort of specificity that some who decry uncertainty seek . . . . Particularly with such a new topic, it is not likely that great certainty will come from rule changes."[27] The degree of specificity that rules *can* provide, and the degree of specificity that rules *should* provide, is an interesting and open question.

With that caveat in mind, it may be useful to ask, have the 2006 amendments achieved these goals? Are, for example, parties paying early attention to potential electronic discovery issues and, if so, is that having the desired effect? Are disputes over electronic discovery decreasing? Moreover, a number of recent surveys have asked litigants and lawyers for their views on whether the 2006 electronic discovery amendments have been effective. The next few sections present some information on these questions.

II. EARLY ATTENTION TO E-DISCOVERY ISSUES

Are attorneys and litigants paying early attention to electronic discovery issues as a result of the amendments? There is some available information on this question. A 2009 Federal Judicial Center ("FJC") survey of attorneys in civil cases found that, in about 1 in 3 cases, the attorneys discussed electronic discovery issues in a Rule 26(f) discovery planning conference.[28] Limiting the analysis to cases with one or more requests for production of electronically stored information, the attorneys discussed electronic discovery issues in more than half of the Rule 26(f) discovery planning conferences—53.2 percent of plaintiff attorneys and 51.5 percent of defendant attorneys reported doing so.

The most commonly discussed topic related to collection was the parties' practices with respect to preservation of electronically stored information, followed by the scope, method, and duration of preserving ESI, the potential cost or burden of collecting, reviewing, and producing ESI, and the possibility of restricting the scope of or altogether avoiding discovery of ESI.[29] The most commonly discussed topic related to production was the format of production of electronically stored information, followed by confidential, trade secret, and privilege issues, the media of production, and privilege log issues.[30]

26 *Id.*
27 Richard L. Marcus, *Only Yesterday: Reflections on Rulemaking Responses to E-Discovery*, 73 FORDHAM L. REV. 1, 17 (2004).
28 EMERY G. LEE III & THOMAS E. WILLGING, FEDERAL JUDICIAL CENTER NATIONAL, CASE-BASED CIVIL RULES SURVEY: PRELIMINARY REPORT TO THE JUDICIAL CONFERENCE COMMITTEE ON CIVIL RULES, AT 15, FIG. 5 (Federal Judicial Center, Oct. 2009) (hereinafter "Preliminary Report").
29 *See id.* at 16, Table 2.
30 *See id.* at 18, Table 3.

Given the sampling frame of the FJC closed-case survey (attorneys of record in civil cases terminated in the last quarter of 2008), about one third of the cases were filed prior to December 1, 2006, the effective date of the amendments, and two thirds after. This makes it possible to compare whether discussion of electronic discovery issues increased after the amendments went into effect. The answer is yes, but not dramatically. Prior to the effective date, 46.6 percent of plaintiff attorneys and 41.7 percent of defendant attorneys reported discussing electronic discovery issues at the Rule 26(f) conference. After the effective date, 54.9 percent of plaintiff attorneys and 53.8 percent of defendant attorneys reported having done so.[31]

While statistically significant, the substantive significance of these findings is open for discussion. After the amendments' effective date, in electronic discovery cases in which a Rule 26(f) conference was held, slightly more than half of respondents reported discussing electronic discovery issues. These issues are not being discussed in many cases in which the Rules envision that they should be. With additional time and experience with the amendments, perhaps the percentage of attorneys raising these issues at the Rule 26(f) conference will increase? There is clearly some reluctance. One plaintiff attorney, interviewed by FJC researchers said, "At Rule 26(f) conferences, I have never met an attorney who wants to get into the electronic issues."[32]

Many attorneys may not have a choice; "the electronic issues" simply cannot be avoidable in certain kinds of cases. It is possible that these issues are being discussed in cases where electronic discovery disputes are more likely to arise. In cases in which the respondent did not report discussing electronic discovery issues at the Rule 26(f) conference, there was a reported dispute over ESI in 21 percent of plaintiff attorneys' cases and in 9.9 percent of defendant attorneys' cases (before and after the effective date). In cases in which such issues were discussed, there was a reported dispute in 31.3 percent and 29.7 percent of cases, respectively. The percentage of cases with one or more reported disputes, in other words, is higher for both plaintiff and defendant attorney respondents in cases in which electronic discovery issues were discussed at an early stage in the case.

This finding is, again, open to discussion. It is possible that discussing electronic discovery issues at the Rule 26(f) conference actually makes future disputes over the same more likely. But it is also possible, as suggested above, that attorneys are more likely to discuss electronic discovery issues in cases in which, even at an early stage of the case, they think that a dispute over the same is likely to occur. In other words, the discussion and the dispute have a common cause—the one does not cause the other. It would be useful to know more about *why* attorneys raise electronic discovery issues at the Rule 26(f) conference, when they do, and why they do not, even in cases in which there is a request for production of electronically stored information.

The discussion at the Sedona Complex Litigation Conference revealed a great deal of consensus that the "early attention" rule amendments have been effective. A similar result was obtained from the magistrate judges' survey, which is also discussed in this issue.[33]

---

31 The percentage of respondents answering "I can't recall" was higher for the cases filed prior to December 1, 2006, than for the cases filed after, as one might expect. The survey also asked whether the discovery plan (if one was adopted) included provisions related to ESI. Overall, plaintiff attorneys in electronic discovery cases reported that the discovery plan included provisions related to ESI in 29 percent of electronic discovery cases and defendant attorneys reported the same in 32.2 percent. Plaintiff attorneys were more likely to report such provisions after the amendments' effective date, but defendant attorneys were not.

32 Quoted in THOMAS E. WILLGING & EMERY G. LEE III, IN THEIR WORDS: ATTORNEY VIEWS ABOUT COSTS AND PROCEDURES IN FEDERAL CIVIL LITIGATION 16 (Federal Judicial Center, Mar. 2010).

33 Emery G. Lee III & Kenneth J. Withers, *Survey of U.S. Magistrate Judges on the Effectiveness of the 2006 Amendments to the Federal Rules of Civil Procedure,* 11SEDONA CONF. J. 201 (2010).

III. E-DISCOVERY DISPUTES

Have the 2006 amendments reduced the number of electronic discovery disputes requiring intervention by the court? Under the heading, "E-discovery Disputes in Court," a 2009 report from Fulbright & Jaworski states:

> Respondents reported that over time they were less likely to have had an e-discovery dispute become the subject of a hearing. In 2006, 58 percent of respondents reported some incidence—from "rarely" to "always"—of an e-discovery matter becoming the subject of a motion, hearing, or ruling from a tribunal. For the largest companies [with annual gross revenues over $1 billion], however, the figure was 75 percent.
>
> In 2007, the figure for the total sample declined slightly to 56 percent, while the large company sample rose slightly to 78 percent. But by 2008, the corresponding figures were down dramatically—to just one third of the total sample and 40 percent of the largest companies. Although we didn't survey respondents on the reasons for this decline, commentators have suggested it resulted from respondents' increased readiness, deeper understanding, and successful use of a well-informed meet-and-confer process.[34]

In 2008, 67 percent of all respondents indicated that no e-discovery issue became the subject of a motion, hearing, or ruling from a tribunal, and 22 percent indicated that this "rarely" happened. Nine percent indicated that this happened "sometimes," two percent answered "frequently," and zero percent answered "always." For U.S. respondents only, 37 percent of respondents indicated at least one such dispute in 2008.[35] The takeaway would be that a majority of respondents had not had an electronic discovery issue or dispute that was the subject of a motion, hearing, or tribunal ruling in the last year.

This may indicate that early attention to electronic discovery issues is having the desired effect. The 2008 Fulbright & Jaworski report gave some credit to the 2006 amendments: "[t]his most likely reflects the efforts of the judiciary to update and clarify rules concerning e-discovery, as well as the desire of many litigants to resolve e-discovery issues through the 'meet and confer' process rather than in the courtroom."[36]

These figures are not difficult to square with the FJC's finding that disputes over electronically stored information occurred in less than 30 percent of cases in which one or more request for its production was made.[37] Few cases involved multiple types of disputes over electronic discovery.[38] The most common type of dispute, reported by about one in 10 respondents in electronic discovery cases, was one over the burden of production of ESI that could not be resolved without court action.[39]

Even if disputes over ESI are not the norm in the general run of electronic discovery cases, when they occur, they are expensive. The FJC found that each reported type of dispute over ESI increased a party's overall litigation costs by 10 percent, even after

34 FULBRIGHT & JAWORSKI L.L.P., E-DISCOVERY TRENDS: E-DISCOVERY FINDINGS FROM THE 2005–2009 FULBRIGHT & JAWORSKI LITIGATION TRENDS SURVEYS 3 (2009) (internal citations omitted).
35 *Id.*
36 *Id.* at 123.
37 Preliminary Report, *supra* note 28, at 24, Fig. 11.
38 *Id.*
39 *Id.* at 23, Fig. 10.

controlling for other case factors (including stakes and time to disposition).[40] So a case with a dispute over cost or burden, a dispute over the reasonable accessibility of requested information, and a dispute over spoliation, for example, would be 30 percent more expensive than a similar case without those disputes. To the extent that disputes are declining—and we really don't have that much information on this, to be blunt[41]—then costs in electronic discovery cases should be trending down, as well.

IV. SURVEYS

Another way of assessing the effectiveness of the 2006 amendments is to simply ask attorneys or judges whether, in their opinion, the amendments have been effective in achieving their goals. In addition to the magistrate judges' survey, discussed elsewhere in this issue, there are a few relevant surveys.

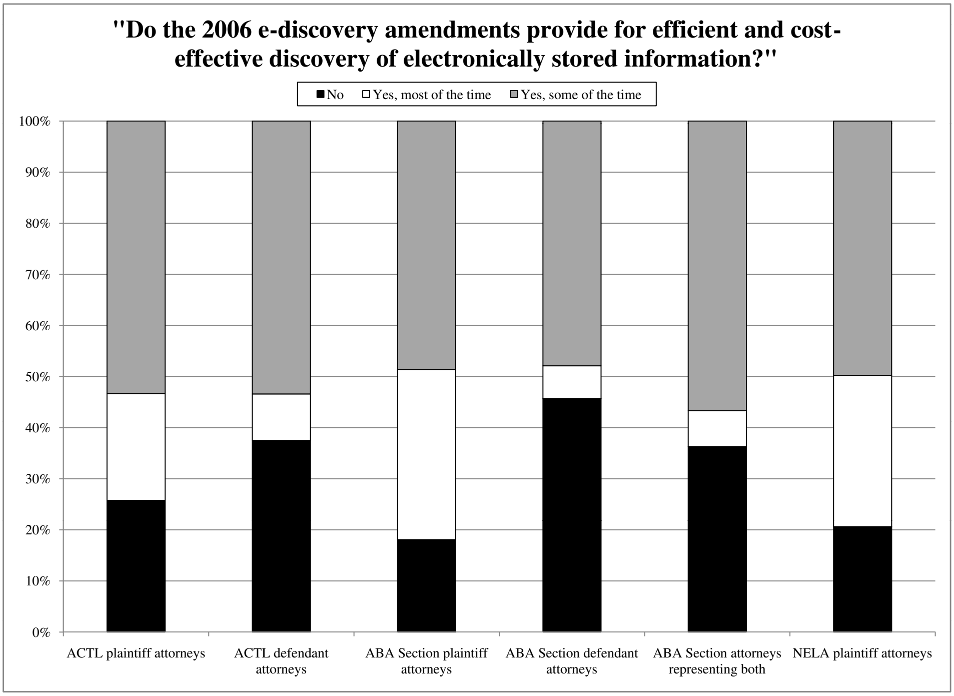
In its 2007 survey, Fulbright & Jaworski asked corporate in-house counsel (U.S. only), "How have the new federal e-discovery rules affected the ease of your company's handling of issues in federal litigation?" This question, at a minimum, gets at the disruption in corporate practices that the 2006 amendments may have caused. Overall, 27 percent of respondents (U.S. residents) answered that "It's more difficult now," 55 percent answered "Not much change," 13 percent answered "It's somewhat easier now," and 5 percent answered "It's much easier now."[42] On the one hand, more than half of respondents said that the rules had not affected them, and almost twice as many respondents answered "Not much change" as answered "more difficult," suggesting that, as of 2007, in-house counsel were not seeing much impact of the amendments. On the other hand, the percentage of respondents answering "It's more difficult now" was *much higher* for those from companies with annual gross revenues of more than $1 billion (35 percent) and of between $100 and $999 million (31 percent) than for those from companies with gross revenues of less than $100 million (two percent).[43] In 2007, at least, it was the biggest firms reporting the most difficulty with the new Rules. Fulbright & Jaworski have not repeated that particular question on subsequent surveys, however.

Three recent surveys of attorneys (one of the fellows of the American College of Trial Lawyers (ACTL), one of members of the American Bar Association Section of Litigation ("ABA Section"), and one of the members of the National Employment Lawyers Association ("NELA") have addressed the effectiveness of the 2006 amendments.[44] All three surveys asked respondents (limited to those who had dealt with electronic discovery cases since December 1, 2006), "[d]o the 2006 e-discovery amendments provide for efficient and cost-effective discovery of electronically stored information?" The response options were "No," "Yes, most of the time," and "Yes, some of the time." Interpreting the results, one should probably take "No" to mean that, in the respondent's view, the amendments *never* provide for efficient and cost-effective discovery of electronically stored information, "most of the time" to mean that the amendments provide for efficient and cost-effective discovery of electronically stored information *in a majority of cases*, and "some of the time" to mean that the amendments provide for efficient and cost-effective discovery of electronically stored information *in less than a majority of cases*.

40 EMERY G. LEE III & THOMAS E. WILLGING, LITIGATION COSTS IN CIVIL CASES: MULTIVARIATE ANALYSIS 5, 7 (Federal Judicial Center, Mar. 2010). The effect was similar for both plaintiffs' and defendants' reported costs.
41 The FJC data does not speak directly to whether disputes have been less likely over time. To the extent that disputes cause cases to take longer to terminate, it is possible that the sampling frame (cases terminated in the last quarter of 2008) does not include the many dispute-ridden cases filed after the amendments' effective date.
42 E-Discovery Trends, *supra* note 33, at 108, Table 54.
43 *Id.*
44 For more information on the surveys, *see* EMERY G. LEE III & THOMAS E. WILLGING, ATTORNEY SATISFACTION WITH THE FEDERAL RULES OF CIVIL PROCEDURE (Federal Judicial Center, Mar. 2010).

The distribution of responses is summarized in Figure 1. Almost half of the ABA Section defendant attorneys (45.7 percent) responded "No," that the amendments never provide for effective and cost-effective discovery of ESI, a view shared by more than a third of ACTL defendant attorneys (37.5 percent) and ABA Section attorneys representing both plaintiffs and defendants (36.3 percent). The percentage of respondents answering "Yes, most of the time" ranges from 33.3 percent of ABA Section plaintiff attorneys, 29.6 percent of NELA plaintiff attorneys, and 20.9 percent of ACTL plaintiff attorneys, to 9.1 percent of ACTL defendant attorneys, seven percent of ABA attorneys representing both plaintiffs and defendants about equally, and 6.4 percent of ABA Section defendant attorneys. The percentage of respondents answering "some of the time": 56.7 percent of ABA attorneys representing both plaintiffs and defendants about equally, 53.4 percent of ACTL plaintiff and defendant attorneys, 49.7 percent of NELA plaintiff attorneys, 48.7 percent of ABA Section plaintiff attorneys, and 47.9 percent of ABA Section defendant attorneys. "Yes, some of the time" was the most common response for every group—accounting for roughly half of all responses.

The pattern here is suggestive, if not surprising. Attorneys primarily representing defendants tend to have a more negative view of the effectiveness of the 2006 amendments than do those primarily representing plaintiffs. Those representing both plaintiffs and defendants tended to respond like defendant attorneys, although they were the group most likely to respond "some of the time."



V. CONCLUSION

It is probably too soon to determine whether the 2006 amendments have been effective in achieving their goals in any kind of global sense. There is, however, some empirical support for the proposition that the effort to get the parties to pay early attention to electronic discovery issues has been at least a limited success. Moreover, disputes over electronically stored information do not appear to be the norm. Still, especially on the defendants' side, there is continuing dissatisfaction with the costs and burdens associated with electronic discovery. These difficulties, as seen in the 2007 Fulbright & Jaworski survey, may be more common for the largest companies.

SURVEY OF UNITED STATES MAGISTRATE JUDGES ON THE EFFECTIVENESS OF THE 2006 AMENDMENTS TO THE FEDERAL RULES OF CIVIL PROCEDURE

By Emery G. Lee III & Kenneth J. Withers[1]
Federal Judicial Center The Sedona Conference®
Washington, DC Phoenix, AZ

Introduction

In the past two years, several surveys have been administered by several different groups attempting to quantify experiences with – and attitudes towards – the American civil litigation system, particularly in federal courts. Starting with the survey of members of the American College of Trial Lawyers (ACTL) administered by the Institute for the Advancement of the American Legal System (IAALS);[2] continuing with surveys of members of the American Bar Association Section of Litigation (ABA),[3] the National Employment Lawyers Association (NELA),[4] and the Association of Corporate Counsel (ACC);[5] and including a survey of 400 senior corporate counsel commissioned by the law firm of Fulbright & Jaworsky,[6] these surveys paint a picture of a legal profession, both defendant and plaintiff, largely dissatisfied with a civil litigation system viewed as costly and time-consuming, burdened by overbroad discovery and governed by ineffective rules.[7] In contrast to these surveys of professional attitudes towards civil litigation in general, the Federal Judicial Center (FJC) conducted a much broader and more scientific survey of both plaintiff and defendant counsel in 3,000 recently concluded federal cases, asking questions about their experience in those particular cases.[8] While the FJC survey yielded strikingly

1 Emery G. Lee III is a Senior Research Associate at the Federal Judicial Center in Washington, D.C. Kenneth J. Withers is Director of Judicial Education and Content at The Sedona Conference® in Phoenix, AZ. The views expressed in this article are those of the authors, and do not necessarily reflect the views of the Federal Judicial Center or The Sedona Conference®.

2 IAALS, "Final Report on the Joint Project of the American College of Trial Lawyers Task Force on Discovery and the Institute for the Advancement of the American Legal System" (2009), *available at* http://www.du.edu/legalinstitute/form-ACTL-Final-Report.html.

3 ABA, "ABA Section of Litigation Member Survey on Civil Practice: Detailed Report" (Dec. 11, 2009), *available at* http://civilconference.uscourts.gov/LotusQuickr/dcc/Main.nsf/$defaultview/EE0CFE1E19ED62808525768D00502C34/$File/ABA%20Section%20of%20Litigation%2C%20Survey%20on%20Civil%20Practice.pdf?OpenElement.

4 NELA, "Summary of Results of Federal Judicial Center Survey of NELA Members, Fall 2009" (2009), *available at* http://civilconference.uscourts.gov/LotusQuickr/dcc/Main.nsf/$defaultview/FE4312C5C76A7B6D852576F70051149D/$File/NELA%2C%20Summary%20of%20Results%20of%20FJC%20Survey%20of%20NELA%20Members.pdf?OpenElement.

5 IAALS, "Civil Litigation Survey of Chief Legal Officers and General Counsel Belonging to the Association of Corporate Counsel" (2010), *available at* http://civilconference.uscourts.gov/LotusQuickr/dcc/Main.nsf/$defaultview/AEA457B143CF3A9E852577050046F931/$File/IAALS%2C%20General%20Counsel%20Survey.pdf?OpenElement.

6 Fulbright & Jaworsky, LLP, "Fulbright's 6th Annual Litigation trends Report," (2009), *available at* http://civilconference.uscourts.gov/LotusQuickr/dcc/Main.nsf/$defaultview/26EBEE9A1CED6C14852576740055DE8F/$File/Fulbright's%206th%20Annual%20Litigation%20Trends%20Survey%20Report.pdf?OpenElement.

7 For a comparison of the ACTL, ABA, and NELA surveys, *see* Emery G. Lee III and Thomas E. Willging, "Attorney Satisfaction with the Federal Rules of Civil Procedure: Report to the Judicial Conference Advisory Committee on Civil Rules," Federal Judicial Center (2010), *available at* http://www.fjc.gov/public/pdf.nsf/lookup/costciv2.pdf/$file/costciv2.pdf.

8 Federal Judicial Center, "National Case-Based Civil Rules Survey: Preliminary Report to the Judicial Conference Advisory Committee on Civil Rules" (Oct. 2009), *available at* http://www.fjc.gov/public/pdf.nsf/lookup/dissurv1.pdf/$file/dissurv1.pdf.

different results in terms of lawyer satisfaction, with most lawyers reporting that they achieved just results at costs proportionate to the stakes of their case,[9] the FJC survey also found that costs increased dramatically in particular categories of cases.[10]

In preparation for the 12th Annual Sedona Conference on Complex Litigation, held April 8 and 9, 2010, in Phoenix, a group of panelists noted that one important constituency hadn't been surveyed – judges. And while judges were not in a position to report on discovery costs or provide useful opinions on whether outcomes were "fair," they could provide insight on whether the current rules of civil procedure, and particularly the 2006 amendments addressing discovery of electronically stored information, have been "effective" in a number of specific ways.

In March of 2010, The Sedona Conference® approached the Federal Magistrate Judges Association (FMJA) and obtained permission to circulate a 62-question online survey to its entire email list of 594 members, which is slightly broader than the currently authorized 523 full-time and 41 part-time United States Magistrate Judge positions in the federal court system.[11] The questions were originally drafted by Complex Litigation XII faculty members Robert D. Owen, Dawson Horn, III, and William P. Butterfield, with input from Ariana J. Tadler, Paul C. Saunders, and Judge Shira A. Scheindlin. The survey questions were extensively revised by Kenneth J. Withers, after input from members of the FMJA and others, including a test survey.[12] We received 87 responses to the survey between March 4 and March 19, 2010.

The responses indicate that, by and large, the rules are working to achieve the "just, speedy, and inexpensive determination of every action" as dictated by Rule 1 of the Federal Rules of Civil Procedure, to the extent that litigants are using the rules to full advantage. It is safe to say that the amendments to Rules 26(f) and 16(b), which prompt the parties and the court to pay "early attention" to potential e-discovery issues, are rated as the most effective amendments by the judges answering the survey. This is consistent with the dialogue that took place at the April 2010 Complex Litigation Conference in Phoenix. More surprisingly, more than 6 in 10 of the judges who responded to the survey reported that the proportionality provisions in Rules 26(b)(2)(C) and 26(c) were being invoked and that, when invoked, were effective in limiting the cost and burden of e-discovery. Similarly, a majority of responding judges reported that the much-maligned scope of discovery, found in Rule 26(b)(1), is effective at least some of the time in assisting the parties and the court to define the appropriate scope of e-discovery. Almost two-thirds of respondents reported that the parties invoked the scope rule at least some of the time to argue for or against the scope of an e-discovery request, in their experience.

One interesting pattern that emerges from the survey is that some of the rules addressed in the survey are not being invoked on a consistent basis. This is especially true of Rule 26(g), which will receive additional discussion in a later section. But it is also true of the use of Rules 26(f) and 16(b) to address preservation issues or the use of Rule 16(b) scheduling orders to set explicit timeframes for the completion of specific e-discovery tasks. Similarly, many respondents expressed no opinion as to whether Rule 34 was effective in avoiding or resolving disputes over the form of production.

---

9 *Id.* at 27-28, 30.
10 *Id.* at 35-44.
11 For a thorough description of the role of United States Magistrate Judges in the federal judicial system, see Federal Magistrate Judges Association, *available at* http://www.fedjudge.org/.
12 It should be noted that while The Sedona Conference® is grateful for the cooperation of individual members of the FMJA in drafting and administering the survey, the FMJA as an organization at no time has endorsed or sponsored the survey or its results.

Use and Effectiveness of 2006 Rules Amendments

Figure 1 summarizes the respondents' answers to a series of questions regarding how often various 2006 Rules amendments are invoked in the e-discovery proceedings over which they had presided. Rule 26(b)(1) ("Scope in General") was reported being invoked "frequently" by 23.4% of respondents, "sometimes" 42.9%, "rarely" 24.7%, and "never" 9.1%. In other words, the scope rule was reported as being invoked sometimes or frequently by almost two-thirds of respondents, 66.2%. About one-third of respondents reported that the scope rule was invoked rarely or never.

Only 13% of responding judges indicated that Rule 26(b)(2)(B) is being invoked "frequently," with 41.6% of respondents indicating "sometimes," 32.5% "rarely," and 11.7% "never." In other words, over half of respondents, 54.6%, responded that Rule 26(b)(2)(B) is invoked at least sometimes, and 44.2% responded that it is invoked rarely or never.

Rules 26(b)(2)(C) and 26(c) were reported being invoked "frequently" by 20.8% of respondents, "sometimes" 42.9%, "rarely" 22.1%, and "never" 13.0%. In other words, more than 6 in 10 respondents, 63.7%, reported that the proportionality rules were invoked, at least sometimes; 35.1% reported that they were invoked rarely or never.

Respondents reported receiving reports of Rule 26(f) discovery planning conferences prior to Rule 16(b) scheduling conferences "always" 28.6%, "frequently" 14.3%, "sometimes" 16.9%, "rarely" 22.1%, and "never" 18.2%. In other words, about 4 in 10 respondents indicated that they always or frequently received reports of the 26(f) conference prior to the 16(b) conference, and about 4 in 10 indicated that they rarely or never received such reports prior to the 16(b) conference. About 2 in 10 indicated that they sometimes received such reports prior to the 16(b) conference.

Interestingly, 40.3% of respondents indicated that they "never" explicitly set out timeframes for completion of e-discovery tasks in Rule 16(b) scheduling orders, with 35.1% indicating doing this "rarely," 16.9% "sometimes," 5.2% "frequently," and 2.6% "always." In other words, more than three-quarters of responding judges reported that they rarely or never set out explicit timeframes for completion of e-discovery tasks in their Rule 16(b) scheduling orders.

On preservation issues, only 2.6% of respondents indicated that they "always" receive reports from Rule 26(f) conferences explicitly addressing ESI preservation, with 16.9% "frequently" receiving such reports, 37.7% "sometimes," 29.9% "rarely," and 11.7% "never." In other words, fewer than 1 in 5 respondents, or 19.5%, frequently or always receive 26(f) reports addressing preservation issues, and more than 4 in 10, or 41.6%, rarely or never do.

Accordingly, Rule 16(b) orders include explicit provisions regarding preservation of ESI "always" in 5.2% of respondents' Rule 16(b) scheduling orders, "frequently" in 11.7%, "sometimes" in 22.1%, "rarely" in 39.0%, and "never" in 20.8%. In other words, almost 6 in 10 respondents rarely or never address ESI preservation issues in Rule 16(b) scheduling orders in proceedings involving e-discovery issues.

Disputes over the form of production of ESI that required judicial intervention to resolve, governed by Rule 34, were reported as occurring "frequently" by 13% of respondents, "sometimes" 45.5%, "rarely" 32.5%, and "never" 9.1%. In other words, almost 6 in 10 respondents, 58.5%, reported that disputes over the form of production of ESI requiring judicial intervention occurred sometimes or frequently, and slightly more than 4 in 10 reported that such disputes occurred rarely or never.

Finally, respondents were divided evenly with respect to how often assertions of privilege have been raised after production of ESI under Rule 26(b)(5). Only 9.1% of respondents indicated that such assertions were made “frequently,” 40.3% “sometimes,” 28.6% “rarely,” and 20.8% “never.” In other words, about half of respondents, 49.4%, reported that such assertions were made sometimes or frequently in ESI cases, and 49.4% reported that such assertions were made rarely or never.

Figure 1: Frequency of use of 2006 rules amendments, based on magistrate judges' experiences: "how often [invoked or used]?" (Questions 4, 7, 9, 15, 16, 24, 25, 37, 47)

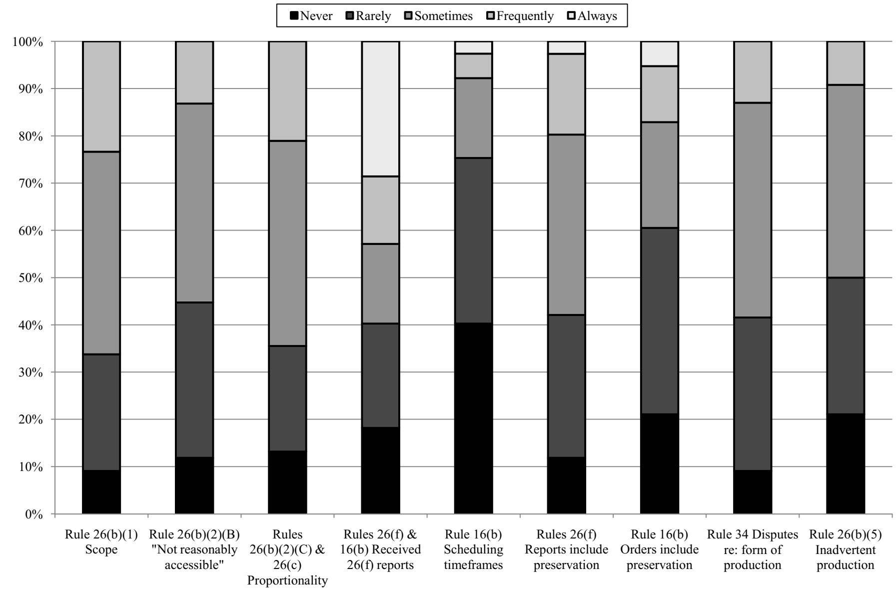


Figure 2 summarizes respondents' answers to a series of questions asking how often 2006 rules amendments had been effective in achieving their goals. With respect to the scope of discovery, Rule 26(b)(1), 18.2% of respondents reported that it had been "frequently" effective, 40.3% "sometimes," 29.9% "rarely," 3.9% "never," with 7.8% not having experience with the rule in e-discovery cases. In other words, 58.5% answered that the rule was effective sometimes or frequently in assisting the parties and the court define the appropriate scope of e-discovery, with 33.8% indicated that it rarely or never did so.

When asked how often Rule 26(b)(2)(B) had been effective in limiting the cost and burden of discovery, 22.1% of respondents answered "frequently," 33.8% "sometimes," 20.8% "rarely," and 5.2% "never," with 16.9% not having experience with the rule. In other words, more than half of respondents, 55.9%, found Rule 26(b)(2)(B) effective in limiting the cost and burden of e-discovery, at least sometimes, more than a quarter of respondents, 26%, answered that it rarely or never did so, and about 1 in 6 respondents did not express an opinion.

With respect to Rules 26(b)(2)(C) and 26(c), 19.5% of respondents answered that they had "frequently" been effective in limiting the cost and burden of e-discovery, 44.2% "sometimes," 15.6% "rarely," 2.6% "never," and 18.2% indicated no opinion. In other words, 63.7% of respondents reported that Rules 26(b)(2)(C) and 26(c) had been effective in limiting the cost and burden of e-discovery, at least sometimes, about 1 in 5, 18.2%, reported that they were rarely or never effective in doing so, and another 1 in 5 had little or no experience with these rules in e-discovery cases.

Turning to the "early attention" Rules 26(f) and 16(b), 9.1% of respondents answered that these amendments had been effective in limiting the cost and burden of discovery "frequently," 54.5% "sometimes," 15.6% "rarely," 1.3% "never," and 18.2% expressed no opinion. Fully 63.6% of respondents found these rules effective, at least sometimes, with just 16.9% finding them effective rarely or never. Still, about 1 in 5 respondents were unable to answer because of limited experience with the rules.

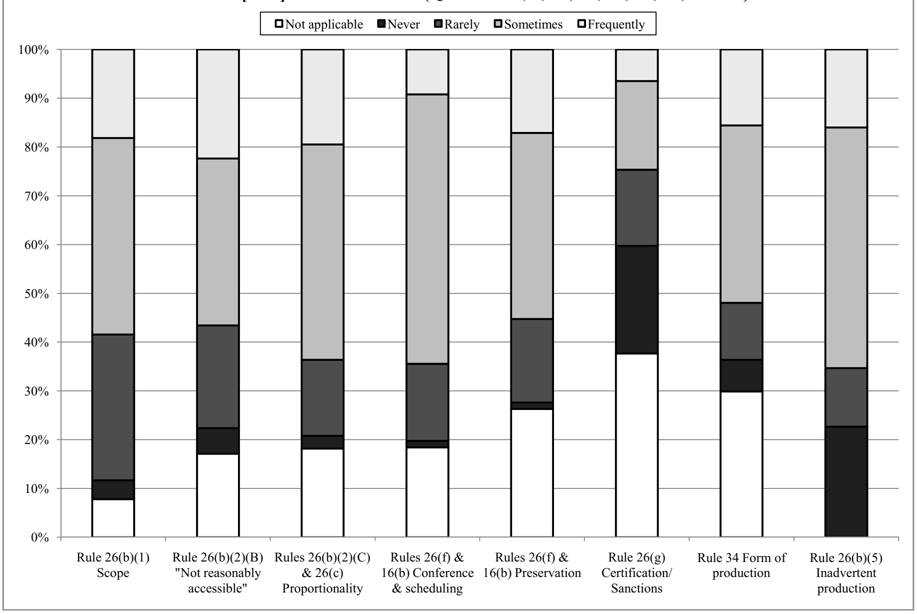
With respect to Rules 26(f) and 16(b) and preservation issues, 16.9% of respondents reported that the "early attention" rules had been effective in addressing preservation issues "frequently," 37.7% "sometimes," 16.9% "rarely," 1.3% "never," with 26% not expressing an opinion. In other words, slightly more than half of respondents, 54.6%, reported that the "early attention" rules had been effective in addressing preservation issues sometimes or frequently, and less than 1 in 5, 18.2%, reported that they had been effective rarely or never. But again, a relatively large percentage of respondents, more than 1 in 4, had little or no experience with cases in which preservation issues were addressed early in e-discovery cases.

As can be seen clearly in Figure 2, Rule 26(g) certification has not been very effective as a deterrent against improper or disproportionately burdensome discovery requests, incomplete discovery responses, and/or improper objections. Only 6.5% of respondents reported that it had been effective in doing so "frequently," 18.2% "sometimes," 15.6% "rarely," and 22.1% "never." Almost 4 in 10 respondents were unable to answer because of a lack of experience with Rule 26(g)—a topic addressed below, in connection with Figure 5.

Rule 34 was reported as effective in avoiding or resolving disputes over appropriate forms in which ESI should be produced “frequently,” by 15.6% of respondents, “sometimes” 36.4%, “rarely” 11.7%, “never” 6.5%, with 29.9% of respondents not expressing an opinion. Similar to responses to other “effectiveness” questions, about half of respondents, 52% in this case, reported the rule effective sometimes or frequently, less than 1 in 5 found the rule effective rarely or never, and a relatively large percentage of respondents not able to express an opinion—about 3 in 10 respondents.

Finally, 15.6% of respondents reported that Rule 26(b)(5) had been effective in avoiding or resolving disputes over inadvertent production of privileged documents in e-discovery “frequently,” 48.1% “sometimes, 11.7% “rarely,” 22.1% “never.”[13] In other words, fully 63.7% reported that Rule 26(b)(5) had been effective, at least sometimes, with 33.8% reporting that the rule was rarely or never effective in preventing or resolving privilege disputes.

Figure 2: Effectiveness of 2006 amendments, based on magistrate judges' experiences: "how often has [rule] been effective?" (Questions 6, 8, 12, 23, 28, 33, 40, and 48)

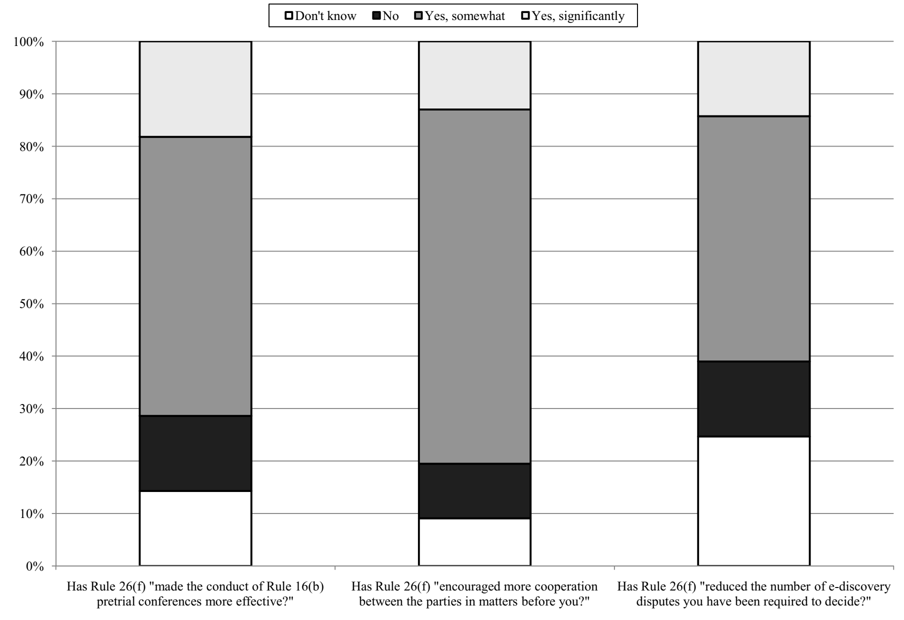


13 Two respondents did not answer the question.

Rule 26(f) Discovery Conferences

Figure 3 summarizes respondents' answers to a further series of questions on Rule 26(f) discovery conferences. When asked whether Rule 26(f) had "made the conduct of Rule 16(b) pretrial conferences more effective," 18.2% of respondents answered, "yes, significantly," 53.2% "yes, somewhat," 14.3% "no," and 14.3% "don't know." When asked whether Rule 26(f) had "encouraged more cooperation between the parties in matters before you," 13% answered "yes, significantly," 67.5% "yes, somewhat," 10.4% "no," and 9.1% "don't know." And when asked whether Rule 26(f) had "reduced the number of e-discovery disputes you have been required to decide," 14.3% answered "yes, significantly," 46.8% "yes, somewhat," 14.3% "no," and 24.7% "don't know." In short, 71.4% of judges responding to the survey had found Rule 26(f) had been at least somewhat effective in improving the conduct of Rule 16(b) pretrial conferences, 70.5% had found Rule 26(f) at least somewhat effective in encouraging the parties to cooperate, and 61.1% had found that Rule 26(f) had reduced the number of e-discovery disputes that they had been called upon to resolve. About 1 in 4 judges expressed no opinion as to whether Rule 26(f) had reduced e-discovery disputes.

Figure 3: Effectiveness of Rule 26(f) discovery conferences (Questions 17, 18, and 19)

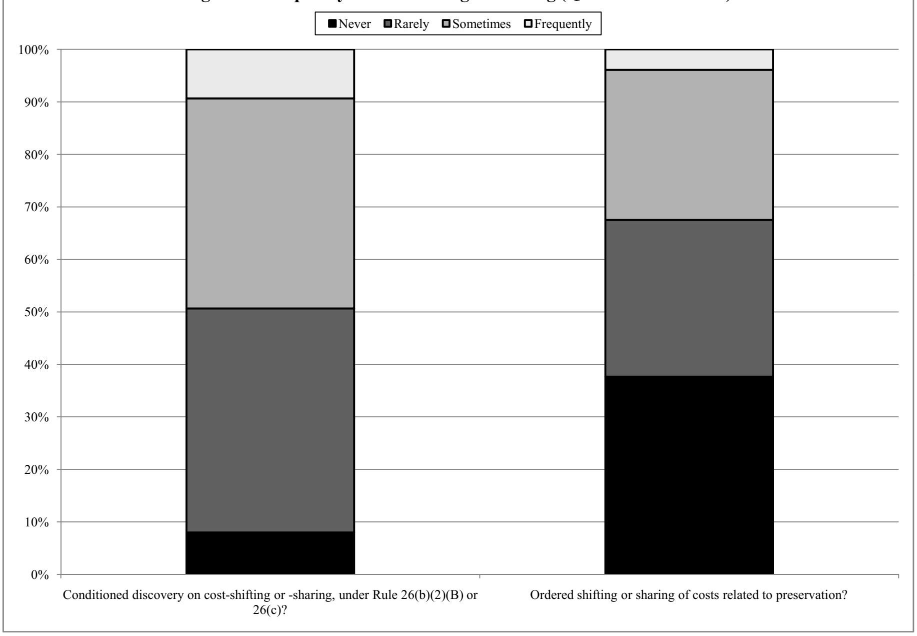


Cost-Shifting

Two questions on the survey asked the judge respondents about the frequency of cost-shifting or cost-sharing in e-discovery cases. Figure 4 summarizes their responses to these questions. There is a relatively stark split in practices related to cost-shifting, in general. Only 9.1% of respondents indicated that they "frequently" condition discovery on cost-shifting, under Rule 26(b)(2)(B) or Rule 26(c); 39% indicated that they "sometimes" do so, 41.6% "rarely," and 7.8% "never." In other words, about half the respondents (48.1%) frequently or sometimes condition discovery on cost-shifting, and about half (49.4%) rarely or never do.

Fewer respondents yet answered that they frequently or sometimes order cost-shifting related to preservation. Only 3.9% of respondents indicated that they "frequently" order cost-shifting related to preservation, and 28.6% indicated that they "sometimes" do so. In contrast, 29.9% of respondents indicated that they "rarely" order cost-shifting related to preservation, and 37.7% indicated that they "never" do so. In other words, more than 2 in 3 respondents indicated that they rarely or never order shifting or sharing of costs related to preservation.

Figure 4: Frequency of cost-shifting or sharing (Questions 11 and 26)



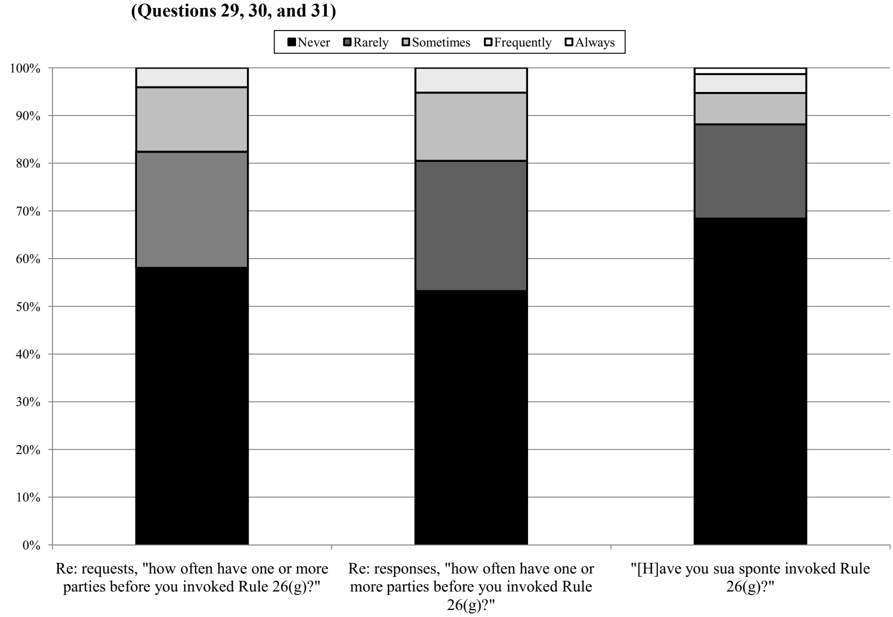
Rule 26(g) Sanctions

As mentioned above, there can be little doubt that Rule 26(g) is an underutilized provision in the Rules. Figure 5 summarizes the responding judges' answers to a series of questions on this rule. With respect to the propriety of requests for production of ESI, only 3.9% of respondents indicated that parties invoke Rule 26(g) "frequently," and just 13% indicated that parties "sometimes" do so. Almost 1 in 4 respondents indicated that parties "rarely" invoke Rule 26(g) with respect to requests for production of ESI, and more than half, 55.8%, indicated that parties "never" do so. With respect to responses to requests for production of ESI, only 5.2% of respondents indicated that parties invoke Rule 26(g) "frequently," and only 14.3% indicated that they "sometimes" do so. More than 1 in 4, on the other hand, indicated that the parties "rarely" do so, and, again, more than half, 53.2%, indicated that parties "never" invoke Rule 26(g) with respect to responses to requests for production.

Moreover, judges generally do not invoke Rule 26(g) *sua sponte*. Almost 1 in 5 respondents, 19.5%, indicated that they "rarely" invoke Rule 26(g) *sua sponte*, and more than 2 in 3, 67.5%, indicated that they never invoke the rule *sua sponte*. Only 1.3% of judges responding to the survey indicated that they "always" do so, with 3.9% indicated that they "frequently" do so, and 6.5% "sometimes.

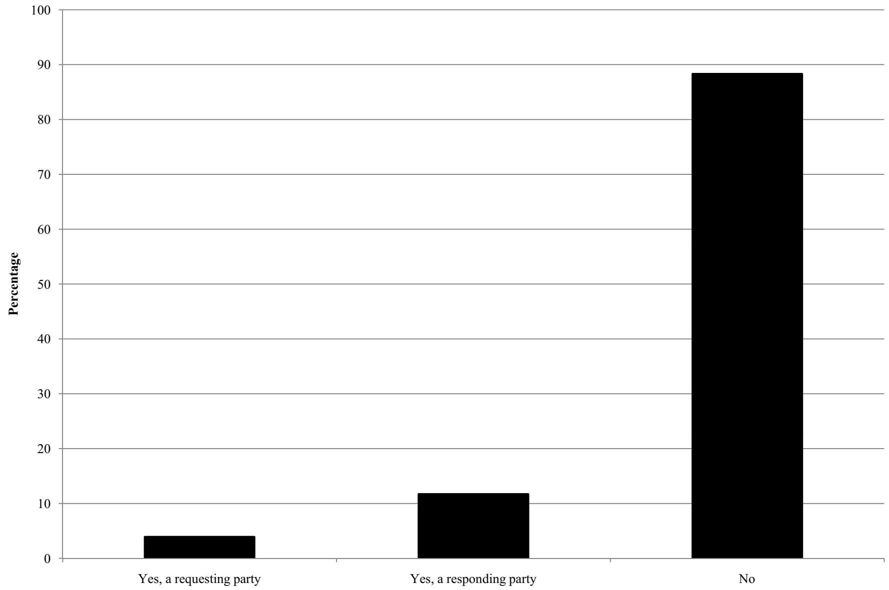
Moreover, sanctions are not commonly imposed for violations of Rule 26(g), as can be seen in Figure 6. Only 3.9% of responding judges indicated that they had sanctioned a requesting party for a violation of Rule 26(g), and only 11.7% indicated that they had sanctioned a responding party. Almost 9 in 10 of the judges responding to the survey, 88.3%, had never sanctioned a party for a violation of Rule 26(g).[14]

Figure 5: Reported use of Rule 26(g) in e-discovery cases: "how often invoked?" (Questions 29, 30, and 31)



14 Figures do not add to 100% because respondents could report imposing sanctions on both kinds of parties.

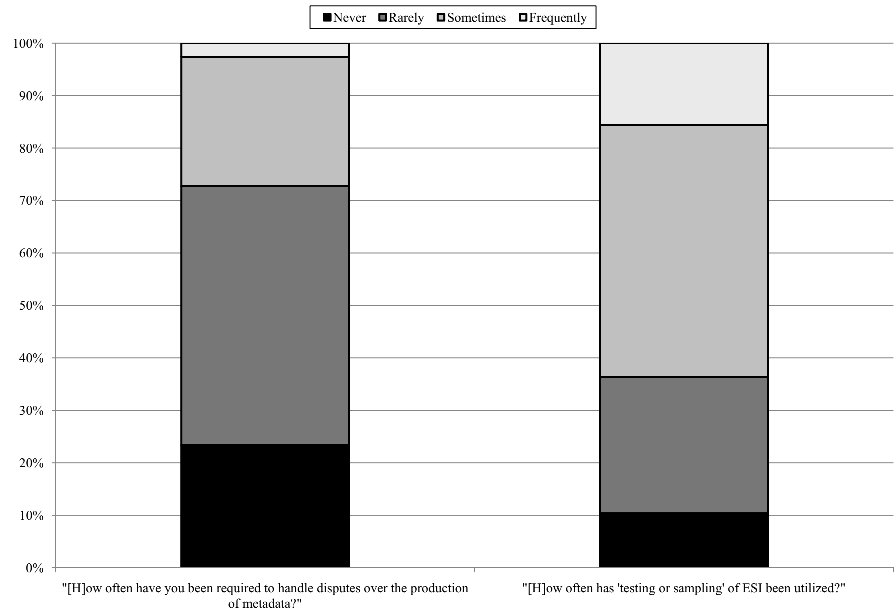
Figure 6: Distribution of responses to question 32: "In e-discovery proceedings over which you have presided, have you ever sanctioned a party for violating Rule 26(g)?"



Rule 34

Two questions asked respondents about issues with respect to Rule 34 and forms of production in e-discovery cases—metadata and testing or sampling of ESI. The responses to these questions are summarized in Figure 7. Only 2.6% of responding judges indicated that they are "frequently" required to handle disputes over the production of metadata, with 24.7% indicated "sometimes," 49.4% "rarely," and 23.4% "never." In other words, almost 3 in 4 respondents, 74.1%, are rarely or never called upon to resolve disputes over metadata. The picture that emerges with respect to the use of sampling or testing of ESI is different, with 15.6% of respondents indicated that testing or sampling is "frequently" used, and 48.1% that it is "sometimes" used. Still, slightly more than 1 in 3 respondents reported that testing or sampling is "rarely," 26%, or "never," 10.4%, used in their e-discovery cases. The takeaway would seem to be that testing or sampling is being used, although perhaps not as often as it might be.

Figure 7: Frequency of special issues in Rule 34 form of production (Questions 38 and 39)

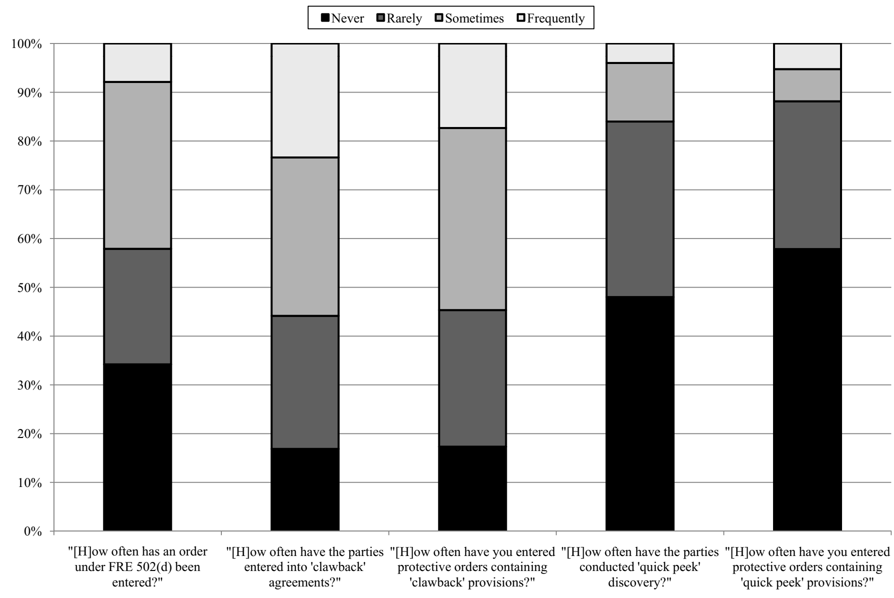


Privilege and Work-Product Issues

The survey asked the judges a series of questions related to recently enacted Federal Rule of Evidence 502 and the use of "clawback" and "quick peek" agreements between the parties in their e-discovery cases. Figure 8 summarizes the responses to these questions.

Consistent with anecdotal impressions, FRE 502(d) appears to be an underutilized rule, at least at this early stage in its life. Only 7.8% of responding judges indicated that they have "frequently" entered orders under FRE 502(d), with 33.8% indicating "sometimes," 23.4% "rarely," and 33.8% "never." Almost 6 in 10 respondents, 57.2%, indicated that the parties rarely or never employ FRE 502(d). It may actually be an encouraging sign that about 1 in 3 judges responding to the survey reported that FRE 502(d) orders are entered "sometimes," but it is probably too early to tell. The responses to a follow-up question on the effectiveness of FRE 502 shows that the most common response was "not applicable/don't know." These responses are summarized in Figure 9.

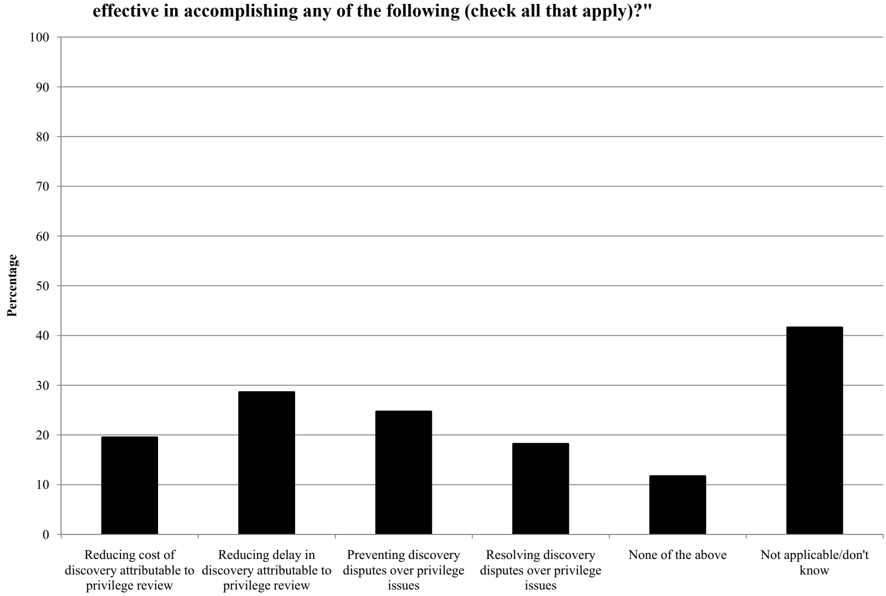
Figure 8: Use of FRE 502 and'clawback' and 'quick peek' agreements (Questions 41, 43, 44, 45, 46)



Even prior to enactment of FRE 502, parties in e-discovery cases had other means to attempt to reduce the expense of privilege and work-product review. Fully 23.4% of the judges responding to the survey indicated that the parties before them had "frequently" entered into clawback agreements, with 32.5% indicated that they had "sometimes" done so, 27.3% "rarely," and 16.9% "never." In other words, more than half of respondents, 55.9%, reported that parties are entering into clawback agreements frequently or sometimes. Moreover, judges appear to be issuing protective orders containing clawback provisions, when the parties enter into such agreements. Fully 16.9% of respondents indicated that they "frequently" enter protective orders containing clawback provisions, with 36.4% indicating that they "sometimes" have done so, 27.3% "rarely," and 16.9% "never." Again, about half of respondents, 53.3%, are entering protective orders with clawback provisions frequently or some of the time.

"Quick peek" discovery procedures, in which open-file disclosure is conducted prior to any privilege review under a reservation of rights, appears to be less common than clawback agreements. Only 3.9% of judges responding to the survey indicated that the parties before them are conducting quick peek discovery "frequently," with 11.7% indicating that the parties have done so "sometimes," 35.1% "rarely," and 46.8% "never." In other words, more than 4 in 5 respondents reported that quick peek discovery is rarely or never used. And it follows that judges are not entering protective orders containing quick peek provisions. Only 5.2% of respondents indicated that they "frequently" enter protective orders containing quick peek provisions and only 6.5% indicated that they "sometimes" do so. Fully 29.9% of respondents indicated that they "rarely" enter protective orders containing quick peek provisions, and 57.1% indicated that they "never" do so. In short, 87% of responding judges rarely or never issue protective orders containing quick peek provisions.

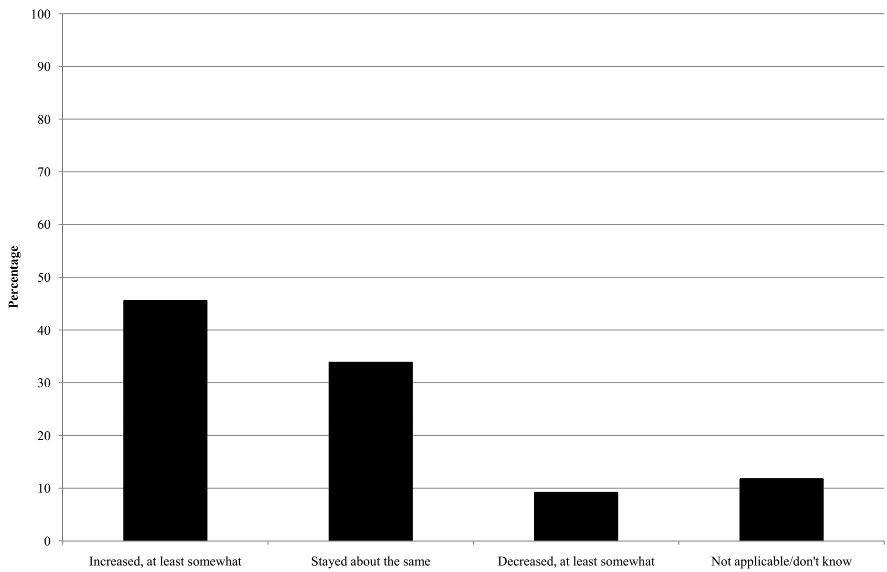
Figure 9: Distribution of responses to Question 50, "In your opinion, based on discovery proceedings over which you have presided in which Rule 502 was invoked, was the rule effective in accomplishing any of the following (check all that apply)?"



Spoliation and Preservation

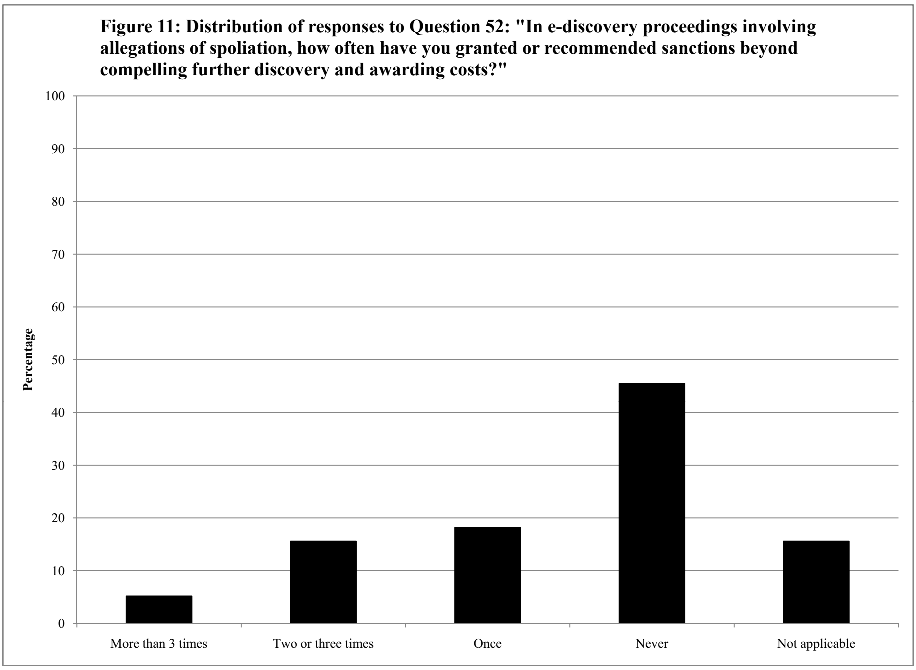
This survey may provide some support for the anecdotal impression that spoliation claims have increased in recent years. About 4 in 5 responding judges indicated that, in their opinion, discovery disputes involving allegations of spoliation had either increased at least somewhat over the past five years (45.5%) or stayed about the same (33.8%). Only 9.1% of respondents expressed the view that such disputes had decreased, at least somewhat, in the past five years, and 11.7% were unwilling or unable to express an opinion. These responses are summarized in Figure 10.

Figure 10: Distribution of responses to Question 51: "In your opinion, have discovery disputes involving allegations of spoliation increased or decreased over the past five years?"



However, the largest group of judges in the survey responded that they had "never" granted or recommended sanctions (beyond compelling discovery and awarding costs) in e-discovery proceedings involving allegations of spoliation. As can be seen in Figure 11, fully 45.5% of responding judges answered "never" to that question. About 1 in 5 respondents, 18.2%, indicated that they had imposed sanctions for spoliation (beyond compelling discovery and costs) only once, and 15.6% indicated that they had done so two or three times. Only 5.2% of respondents indicated that they had imposed sanctions for spoliation (beyond compelling discovery and costs) more than three times.[15]

Another way to get at the rarity of sanctions: of judges who have imposed or recommended sanctions in e-discovery cases involving allegations of sanctions, only 13.3% had imposed such sanctions more than three times.



15 Another 15.6% responded "not applicable."

Overall Efficacy of Civil Litigation

Figure 12 summaries the judges' reactions to six statements that parallel the statements found in the surveys of practitioners. Six in 10, 60.9%, agreed with the first statement that cases "are generally being resolved or settled on the merits," with only 2.3% disagreed, 19.5% neither agreeing nor disagreeing, and 17.2% "don't know." On the other hand, an almost equal percentage, 59.8%, agreed with the statement that costs "are affecting the parties' assessment … [of] whether to file or settle a case." 8% disagreed, 13.8% neither agreed nor disagreed, and 18.4% "don't know."

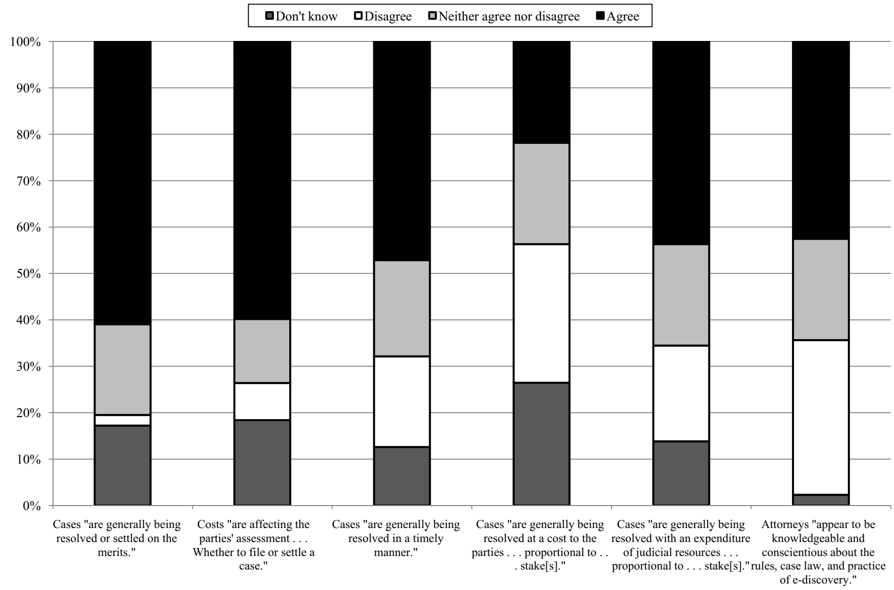
Slightly less than half of the responding judges, 47.1%, agreed with the statement that cases "are generally being resolved in a timely manner," with 19.5% disagreeing, 20.7% neither agreeing or disagreeing, and 12.6% "don't know."

Rather dramatically, a mere 21.8% of the responding judges agreed with the statement that cases "are generally being resolved at a cost to the parties … proportional to the stake[s]." Slightly more, about 29.8%, disagreed with that statement, with 21.8% neither agreeing nor disagreeing and 26.4% "don't know."

In a question unique to this survey, only 43.6% of the responding judgers agreed with the statement that cases "are generally being resolved with an expenditure of judicial resources … proportional to" the stakes. 20.6% disagreed, 21.8% neither agreed nor disagreed, and 13.8% "don't know."

Finally, judges were asked about the attorneys who appear before them in cases involving e-discovery. Less than half of the responding judges, 42.5%, agreed with the statement that attorneys "appear to be knowledgeable and conscientious about the rules, case law, and practice of e-discovery." A substantial 33.3% of the judges disagreed with that statement. 21.8% neither agreed nor disagreed, and only 2.3% indicated that they didn't know, indicating that 1/3 of judges believe that attorneys are unprepared, to one degree or another, for e-discovery.

Figure 12: Magistrate judges' reactions to statements regarding e-discovery cases and costs (Questions 55–60)



PRESERVATION RULEMAKING AFTER THE 2010 LITIGATION CONFERENCE

Thomas Y. Allman[1]
Cincinnati, OH

At the 2010 Civil Litigation Conference held at the Duke Law School in May 2010 (the "Litigation Conference"), a panel of jurists and practitioners representing a cross-section of e-discovery involvement (the "E-Discovery Panel")[2] reported its consensus recommendation that a rule addressing preservation (spoliation) would be a valuable addition to the Federal Rules. The Panel consisted of two federal judges (a district judge and a magistrate judge), a plaintiff's employment counsel, two defense counsel, a former general counsel and was moderated by the president-elect of the American College of Trial Lawyers.

The Panel also outlined considerations which might be involved in drafting such a rule, should the effort be undertaken.

The Advisory Committee Report, issued shortly after the Litigation Conference ("Conference Committee Report"), noted that preservation "rulemaking was considered [in the 2006 Civil Rules Amendment process] but put aside, apart from the protection against sanctions included in Rule 37(e)."[3] For the reasons stated below, it is submitted that this is an appropriate time for such rulemaking despite the proximity in time to the 2006 Amendments.

I. THE 2010 LITIGATION CONFERENCE

The announced purpose of the 2010 Litigation Conference was to "explore the current costs of civil litigation, particularly discovery and e-discovery, and to discuss possible solutions."[4] The Conference included a series of panel discussions designed to reflect a broad spectrum of views, with participants invited from academia, the judiciary and the practicing bar. The discussions expanded upon the written submissions made by participants, all of which remain available on a website established for that purpose.[5]

The organizers of the Litigation Conference asked a number of entities to submit "suggested amendments to the Federal Rules of Civil Procedure." Thus, submissions were made by the ACTL/IAALS Task Force,[6] Lawyers for Civil Justice[7] and a Special Committee

---

1 ©2010 Thomas Y. Allman. The author, a former General Counsel, was a member of the E-Discovery Panel at the Litigation Conference held by the Civil Rules Advisory Committee on May 10-11, 2010 at the Duke Law School.
2 The formal title of the Panel segment was: "E-Discovery: Discussion of the Cost Benefit Analysis of E-Discovery and the Degree to Which the New Rules are Working or Not."
3 REPORT OF THE CIVIL RULES ADVISORY COMMITTEE, May 17, 2010, at 12, *available at* http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/Reports/CV05-2010.pdf (conceding the potential advantages of such rules "for planning [an entity's] affairs and for achieving some uniformity").
4 Memorandum, Hon. John G. Koeltl to Participants in the 2010 Conference (Aug. 4, 2009 (copy on file with author).
5 Submissions referred to in this Paper ("CONFERENCE PAPERS") are *available at* http://civilconference.uscourts.gov/LotusQuickr/dcc/Main.nsf/h_Toc/B896CCD29A7DE0C88525764100492D17/?OpenDocument.
6 The ACTL/IAALS Task Force presented model Rules which evolved from the principles expressed its Final Report. *See* ACTL/IAALS Pilot Project Rules, CONFERENCE PAPERS.
7 Lawyers for Civil Justice (on behalf of DRI, the Federation of Defense & Corporate Counsel and the International Association of Defense Counsel), *White Paper: Reshaping the Rules of Civil Procedure for the 21st Century*, CONFERENCE PAPERS.

of the ABA ("ABA Special Committee"),[8] with opposing views expressed by members of the class action bar[9] and the Center for Constitutional Litigation.[10] A comprehensive summary of those submissions was circulated prior to the Conference.[11] In addition, the Seventh Circuit Electronic Discovery Pilot Principles now being implemented by Standing Orders as a pilot project were described[12] by the Chief Judge of the Northern District of Illinois. Additional rulemaking concepts were discussed by many of the participants and audience members.

The Conference was also presented with the results of a survey[13] administered to members of the American Bar Association Section on Litigation ("ABA Survey"),[14] the American College of Trial Lawyers[15] and the National Employment Lawyers Association ("NELA Survey").[16] A survey of opinions held by chief legal officers and general counsel by the Association of Corporate Counsel ("ACC Survey") was also made available[17] as was a study of the costs of litigation undertaken by the Searle Institute (the "Searle Survey").[18]

Finally, the Federal Judicial Center presented results from a survey of trial counsel involved in recently closed cases (the "FJC Survey").[19]

Some of the key results:

1. Cooperation. There was widespread agreement among Conference participants with the observation of the ABA Special Committee that "[w]hen lawyers are collaborative and cooperative the case costs less for clients."[20] The Sedona Conference® efforts promoting its *Cooperation Proclamation* were duly noted and supported.[21] One participant, expressing mild skepticism, referred to the approach as the "Kumbaya Campaign."[22] The Conference Committee Report also referred to "rather wistful suggestions" for revising Rule 1.[23]

Greater cooperation and collaboration on preservation was the intended focus of the 2006 amendment to Rule 26(f) mandating discussion of preservation at the meet and confer prior to the Rule 16(b) scheduling conference. Clearly, early agreement on

---

8 ABA Special Committee, *Civil Procedure in the 21st Century* (Apr. 24, 2010), CONFERENCE PAPERS.

9 One Paper asserted that it is "far too early, and the current data too flawed" to begin efforts to revise the [2006 Amendments] relating to e-discovery. *See* Milberg LLP and Hausfeld LLP, *E-Discovery Today: The Fault Lines Not in Our Rules*, at 2-3, CONFERENCE PAPERS ("Milberg and Hausfeld").

10 Center for Constitutional Litigation, PC, *Nineteenth Century Rules for Twenty-First Century Courts: An Analysis and Critique* (Mar. 2010), CONFERENCE PAPERS ("CCL Analysis").

11 Rules Administrative Office, *Summary Comparison of Bar Association Submissions to the Duke Conference Regarding the Federal Rules of Civil Procedure* (Apr. 26, 2010), CONFERENCE PAPERS.

12 7TH CIRCUIT PILOT PROGRAM REPORT ON PHASE ONE (2010) ("PHASE ONE REPORT") (summarizing results from October 2009 through March 2010), CONFERENCE PAPERS.

13 The original survey instrument was administered in 2008 to members of the American College of Trial Lawyers and subsequently adapted for use by the ABA Section of Litigation and the National Employment Lawyers Association by the Federal Judicial Center.

14 American Bar Association Section of Litigation, *Member Survey on Civil Practice* (Dec. 11, 2009), CONFERENCE PAPERS.

15 The results of the survey was described in the 2009 joint "ACTL/IAALS Final Report" advocating sweeping changes in the civil litigation practices because the civil justice system "is in serious need of repair." American College of Trial Lawyers and the Institute for the Advancement of the American Legal System Final Report, at 2, CONFERENCE PAPERS.

16 National Employment Lawyers Association, *Summary of Results of [a] Survey of NELA Members, Fall 2009*, CONFERENCE PAPERS.

17 Association of Corporate Counsel, *Civil Litigation Survey of Chief Legal Officers and General Counsel*, CONFERENCE PAPERS.

18 The survey was conducted on behalf of Lawyers for Civil Justice, the Civil Justice Reform Group and the U.S. Chamber Institute for Legal Reform. *See* Searle Survey, *Litigation Cost Survey of Major Corporations,* CONFERENCE PAPERS.

19 Emery G. Lee III & Thomas J. Willging, Federal Judicial Center, *National Case-Based Civil Rules Survey, Preliminary Report* (2009), CONFERENCE PAPERS.

20 ABA Special Committee, *supra*, at 5, CONFERENCE PAPERS.

21 Hon. Paul Grimm, *The State of Discovery Practice in Civil Cases, Must The Rules Be Changed to Reduce Costs and Burdens, or Can Significant Improvements Be Achieved Within The Existing Rules,* at 19-21, CONFERENCE PAPERS (describing ongoing projects to facilitate cooperation in discovery).

22 Elizabeth J. Cabraser, *Uncovering Discovery*, at 35, CONFERENCE PAPERS ("[e]veryone would wish this leap to succeed, no one who deserves to be a lawyer or judge would wish to be seen subverting it, and everyone would volunteer to be the second to jump").

23 Conference Committee Report, *supra,* at 11.

preservation issues by counsel is ideal. However, many preservation issues must be resolved *prior* to the meeting at a time when opposing counsel are not available to commit to the ultimate scope of discovery.[24]

According to the FJC Survey, the topic of "retention" was discussed in only about 17% of the cases surveyed.[25] One unfortunate result of the increasing tendency to hold counsel responsible for client mishaps[26] has been, in absence of agreement, that outside counsel are reluctant to advise clients to do anything other than err on the side of over-preservation.[27]

A clear desire for greater guidance on preservation obligations was expressed by potential producing parties and reflected in the proposals identified by the E-Discovery Panel in its "Elements of a Preservation Rule."[28] This includes guidance for actions which must be undertaken prior to the commencement of litigation.

2. Judicial Management. Participants at the 2010 Litigation Conference generally echoed the call in the ACC Survey for "greater court involvement in 'crafting an e-discovery plan prior to a dispute.'"[29] The ABA Special Committee opined that "[j]udges should play an active role in supervising the discovery process and should work to assure that discovery costs are proportional to the dispute."[30] The Conference Committee Report speaks of the "virtual, perhaps absolute, unanimity" with which "[p]leas for universalized case management" were greeted.[31]

Many participants emphasized the adequacy of existing tools such as Rule 16 and Rule 26(c),[32] noting the ability to utilize creative methods of case management such as staggered discovery and judicially mandated "tracks" based on the need for increased judicial involvement in e-discovery.[33] Cases assigned to a more intensive track could be subject to a standing order with preservation management provisions such as those in the ongoing Seventh Circuit Electronic Discovery Pilot program.[34]

However, while case management could undoubtedly be improved, Rule 26(c) does not currently empower courts to deal with unduly burdensome preservation demands,[35] with the possible exception of those relating to inaccessible sources identified under Rule 26(b)(2)(B).[36] Rule 26(f) is also silent on the need to report on open

---

24 Kenneth J. Withers, *"Ephemeral Data" and the Duty To Preserve Discoverable Electronically Stored Information*, 37 U. BALT. L. REV. 349, 377 (Spring 2008) ("By the time the parties sit down at the Rule 26(f) conference, the preservation issues surrounding ephemeral data may be moot and the fate of the responding party may already be sealed, if sanctions are later found to be warranted").

25 Retention was listed as discussed in only 35% of the cases where ESI was discussed, which constituted about 50% of the cases surveyed. *See FJC Civil Rules Survey*, CONFERENCE PAPERS, at 15- 24. The FJC did not survey the extent to which preservation agreements were reached as a result of such discussions.

26 *See, e.g., In re A&M Florida Properties II, LLC*, 2010 WL 1418861 (Bankr. S.D.N.Y. Apr. 7, 2010) (sanctioning counsel and client for conduct which resulted in delayed production).

27 Thomas Y. Allman, *Preservation and Spoliation Revisited: Is it Time for Additional Rulemaking?* at 3, CONFERENCE PAPERS.

28 *See generally*, E-Discovery Panel, *Elements of a Preservation Rule,* CONFERENCE PAPERS.

29 ACC Survey, *supra*, at 3, CONFERENCE PAPERS.

30 ABA Special Committee, *supra,* at 4, CONFERENCE PAPERS.

31 Conference Committee Report, *supra,* at 8 (discussing need for "one-case-one judge," firm deadlines, regular and prompt access and "substantially successful" use of techniques stemming from the 2006 Amendments).

32 Hon. Paul Grimm, *supra,* at 21- 33, CONFERENCE PAPERS (suggesting use of discovery budgets, local rules and "common sense" systems such as innovative use of Rule 502).

33 Conference Committee Report, *supra*, at 14 ("Some version of tracking could be added" by "building into the present sequence or by adding a separate set of 'simplified' or 'tracking' rules" which could be mandatory as "long as jury trial is preserved").

34 *See* 7TH CIRCUIT PILOT PROGRAM REPORT ON PHASE ONE (2010) ("PHASE ONE REPORT"), *supra,* CONFERENCE PAPERS.

35 *Kemper Mortgage v. Russell*, 2006 WL 4968120 (S.D. Ohio May 4, 2006) (refusing to opine on whether the producing party was being "overly cautious" in plans for a litigation hold).

36 FINAL REPORT OF THE ADVISORY COMMITTEE (May 2005) "FINAL REPORT (2005)," at 51 [Changes Made after Publication and Comment](" [Rule 26(b)(2)(B)] has been changed to recognize that the responding party may wish to determine its search and potential preservation obligations by moving for a protective order"), *available at* http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/Reports/CV5-2005.pdf.

preservation issues after "meet and confers"[37] and many trial courts do not use Rule 16 conferences for any form of discovery planning purposes. In rare cases, courts treat the issue of preservation planning by local initiatives.[38]

3. Pre-Litigation Access to Court. The risks associated with spoliation sanctions for those who "guess wrong" can force over-preservation and case-dispositive decisions prior to an action being commenced.[39] In *Texas v. City of Frisco*,[40] the State of Texas was denied an opportunity to seek pre-commencement relief from a broad preservation demand relating to a recently announced highway project. The magistrate judge held that a justiciable controversy did not exist even though the municipality making the demand clearly intended to bring a suit.

The Conference Committee Report endorsed the possibility of enacting "more explicit provisions" for dealing with preservation issues in Rule 26(c), "possibly including preservation before an action is filed."[41] Rule 27 "Depositions to Perpetuate Testimony," which authorizes limited pre-litigation proceedings "before an action is filed" to "prevent a failure or delay of justice" could serve as the vehicle. It could be amended to permit issuance of preservation or protective orders on the initiative of a person who "expects" to be a party to an action but "cannot presently bring it or cause it to be brought."[42] Rule 27 actions do not require an independent basis for federal jurisdiction as long as the contemplated action is itself authorized by statute.[43]

4. E-Discovery Costs. The Searle Survey documented a significant increase in the costs of discovery, including e-discovery. A substantial majority of defense and mixed practice lawyers surveyed by the ABA agreed that "the costs of litigation have risen disproportionately due to e-discovery."[44] Similar percentages believed that "e-discovery is overly burdensome."[45]

In some cases, the remedies proposed involved better case management and increased use of technology. Counsel representing plaintiffs in employment disputes agreed that costs are higher but asserted that "properly managed e-discovery can reduce the overall costs."[46] Milburg and Hausfeld argued that upwards of 80% of privilege and review costs—a primary cause of increased costs—will be reduced by use of "search technology."[47]

Those entities responding to the Searle Study, however, supported the recommendations initially made by Lawyers for Civil Justice that "each party [should] pay the costs of the discovery it seeks" which would "encourage each party to manage its own discovery expenses by shifting the cost-benefit decision onto the requesting party – the best cost avoider."[48]

37 *See also* Form 52, *Report of the Parties' Planning Meeting* (2007) (no mention of preservation issues) *and compare to* ACT/IAALS Pilot Rule 8.1(b) (suggesting that courts discuss "production, continued preservation, and restoration of [ESI]"), *supra,* CONFERENCE PAPERS.

38 *See* Standing Order, ¶ 6 (N.D. Cal.) (requiring report on arrangements for evidence preservation to court before the scheduling conference).

39 *TIG Insur. Co. v. Giffin Winning,* 444 F.3d 587, 392 (7th Cir. 2006) (settlement occurred only after expending $1.2M in defending spoliation motion).

40 2008 WL 828055 (E.D. Tex. Mar. 27, 2008).

41 Conference Committee Report, *supra*, at 13 (noting the need to amend rules to allow for emergency application on filing the complaint).

42 Subsection (a)(1) of Rule 27 could be amended to authorize a movant to seek relief to "respond to preservation demands concerning discovery before an action is filed."

43 Jay E. Grenig, *Taking and Using Depositions Before Action or Pending Appeal in Federal Court,* 27 AM. J. TRIAL ADVOC. 451, 454-55 (Spring 2004).

44 ABA Survey, *supra*, at 5, CONFERENCE PAPERS.

45 *Id.* at 11.

46 NELA Survey, *supra*, at 6, CONFERENCE PAPERS.

47 Milberg and Hausfeld, at 46, CONFERENCE PAPERS ("the use of FRE 502 could reduce the cost of privilege review by as much as 80% in some cases").

48 Searle Study, *supra,* at 7 & 16, CONFERENCE PAPERS ("out of over 743 e-discovery disputes reported between 2004 and 2009, there was only *one* case where cost shifting was utilized to resolve a dispute") (emphasis in original).

The participants did not separately identify the need to shift e-discovery preservation costs, since, as the RAND representative explained, they are hard to identify but are nonetheless very real. Some courts already endorse cost-shifting as a viable option for incremental preservation costs,[49] as do the ACTL/IAALS Pilot Project Rules[50] and some local rules.[51]

5. Spoliation Sanctions. The emergence of e-discovery has coincided with a substantial growth in allegations that spoliation has occurred. A survey presented at the Conference (the "Sanctions Survey")[52] confirmed the author's findings that reported decisions have increased from an average of 10 or less per year prior to 2005 to at least 71 in 2009.[53] As the survey put it, "[allegations of] [f]ailure to preserve is the most prevalent" source of the disputes.[54] This is occurring despite the fact that a majority of entities surveyed by the ACC report that they have implemented some form of a litigation hold "mechanism" and records retention/destruction policies as well as other steps to enhance e-discovery compliance.[55]

In one sense, the increase in motion practice is understandable, given the inconsistencies among the circuits and the rigid requirements imposed by some courts[56] As noted in the Sanctions Survey, "[l]itigants and their lawyers [facing demands for e-discovery] must immediately identify, promptly preserve, comprehensively collect, fairly filter, properly process, rigorously review, and produce ESI in appropriate format[s] without sluggishness, purposeful or otherwise." It is not surprising that challenges have generally increased.[57]

Unfortunately, despite the addition of Rule 37(e) to address the inconsistent treatment of culpability among the circuits, many courts simply ignore the plain meaning of the rule.[58] Some courts—but not all[59]—have concluded that Rule 37(e) is inapplicable if a preservation duty existed at the time of the loss at issue, regardless of the culpability involved.[60]

Thus, the ABA Special Committee suggested that the "federal courts should adopt a uniform standard to address when sanction may be imposed for the deletion of ESI after a duty to preserve ESI has attached."[61] The author, Lawyers for Civil Justice, the

---

49 *Treppel v. Biovail,* 233 F.R.D. 363, 372-373 (S.D.N.Y. Feb. 6, 2006).
50 *ACT/IAALS Pilot Rules*, Comment, Rule 8.1 (referring to potential of court for shifting "any or all costs associated with the preservation, collection and production of [ESI] if the interests of justice and proportionality so require."), *supra,* CONFERENCE PAPERS,
51 Local Rule 26.1, District of New Jersey (2007) (requiring parties to meet and confer to attempt to agree on "[w]ho will bear the costs of preservation, production, and restoration (if any) of any digital discovery").
52 Willoughby and Jones, *Sanctions for E-discovery Violations: By the Numbers*, CONFERENCE PAPERS.
53 The number of reported cases found by the author was 32 in 2006, 68 in 2007 and 62 in 2008 (copies on file with author). There undoubtedly are many more that have escaped the author's unscientific tracking methods. *See also Symposium on Ethics and Professionalism in the Digital Age*, 60 MERCER L. REV. 863, 899 (2009) (high volumes of spoliation motions were almost unheard of before e-discovery).
54 Sanctions Survey, *supra*, CONFERENCE PAPERS.
55 ACC Survey, *supra*, at 8, CONFERENCE PAPERS.
56 Lee H. Rosenthal, *A Few Thoughts on Electronic Discovery After December 1, 2006*, *supra,* 116 YALE L. J. 167, 190 (2006 Pocket Part) ("judge seeking effective control over electronic discovery may impose unrealistically stringent demands on litigants and lawyers, which will predictably lead to an increase in sanctions motions if parties cannot meet the demands").
57 A less attractive explanation is that there may be an element of strategic gamesmanship involved, since experience has shown that once allegations of spoliation are made, they tend to become the primary focus of the litigation, not the merits.
58 Some decisions show a remarkable reluctance to apply the plain meaning of the rule. *Wilson v. Thorn Energy LLC,* 2010 WL 1712236 (S.D.N.Y. Mar. 15, 2010) (finding that it was not "good faith" within Rule 37(e) to fail to make a copy of a flash drive before it inadvertently failed).
59 *Olson v. Sax,* 2010 WL 2639853, at 3 (E.D. Wisc. June 25, 2010) (no sanctions imposed because of overwriting after duty to preserve attached since "no evidence that [defendant] engaged in the 'bad faith' destruction of evidence for the purpose of hiding adverse evidence").
60 *Doe v. Norwalk Community College,* 248 F.R.D. 372 (D. Conn. 2007).
61 ABA Special Committee*, supra,* at 12,CONFERENCE PAPERS.

ACTL/IAALS Pilot Rules and, to a lesser extent, the Conference Committee Report[62] have also suggested addressing this issue.

6. Standard of Care. The Federal Rules do not currently articulate a standard of care for the implementation of preservation obligations under Rule 34. As one court presciently put it, "[a]bsent from Rule 34 is a procedure to preserve documents, things or land from damage or destruction that could compromise the integrity of the very existence of the evidence requested."[63]

As noted earlier, the E-Discovery Panel has recommended decoupling the duty to preserve from its evidentiary roots and incorporating it, taking into consideration a number of appropriate elements, presumably as part of Rule 34.[64] This would implement a similar proposal, made by the American College of Trial Lawyers at the time of the 2006 Amendments, that "[i]t would enhance . . . the entire body of the Federal Rules" if the Rules were amended "to state a standard of care for production and preservation—which we think should be reasonableness."[65]

The linkage between reasonable conduct and proportionality suggests that both characteristics should be referenced in any rule. Indeed, if "one word came to express the quest for speedier and less expensive procedures" at the Conference it was increased use of "proportionality."[66] Principle Five of the *Sedona Principles*[67] and the Seventh Circuit Pilot Program on E-discovery[68] both support this approach, as do thoughtful decisions such as *Rimkus Consulting v. Cammarata,*[69] where the district court held that "[w]hether preservation or discovery conduct is acceptable in a case depends on what is *reasonable,* and that in turn depends on whether what was done—or not done—was *proportional* to that case and consistent with clearly established applicable standards." (emphasis in original).[70]

II. PRE-LITIGATION CONDUCT

The principal argument against preservation rulemaking is that the Supreme Court is barred from enacting rules relating to pre-litigation matters.[71] The Committee Conference Report, in its discussion of preservation rulemaking, stated that the "first issue" is whether "a rule addressing discovery obligations and sanctions can attach to conduct before an action is filed in federal court."[72]

A preservation rule that fails to deal with pre-litigation conduct risks being irrelevant. Much of the "preservation action" of concern occurs before suit is filed. In *Phillip*

62 Conference Committee Report, *supra,* at 13 ("Rule 37(e) might be amended so as to bar sanctions against an attorney in the circumstances that now bar sanctions against a party").
63 *Capricorn Power Company v. Siemens Westinghouse,* 220 F.R.D. 429, 433 (W.D. Pa. 2004).
64 Statutory or regulatory requirements may also trigger preservation actions depending upon the intent of Congress or the regulators, a subject beyond the scope of these remarks. *See, e.g.,* Committee Note, Rule 26(f) (2006) (the obligation "may arise from many sources, including common law, statutes, regulations, or a court order in the case").
65 Letter, Robert L. Byman, Chairman to Peter G. McCabe, Secretary, *Proposed Amendments To the Federal Rules of Civil Procedure* (Jan. 25, 2005) (on file with author).
66 Conference Committee Report, *supra,* at 7 ("[h]ow to achieve it is the question").
67 Principle 5, THE SEDONA PRINCIPLES (2d ed. 2007) ("The obligation to preserve electronically stored information requires reasonable and good faith efforts to retain information that may be relevant to pending or threatened litigation. However, it is unreasonable to expect parties to take every conceivable step to preserve all potentially relevant electronically stored information.").
68 7TH CIRCUIT PILOT PROGRAM ("Every party to litigation and their counsel are responsible for taking reasonable and proportionate steps to preserve relevant and discoverable ESI within its possession, custody or control.").
69 2010 WL 645353 (S.D. Tex. Feb. 19, 2010).
70 *Id.* at *6.
71 Memorandum (Jan. 27, 2004), Advisory Committee on Civil Rules, at 35-36, *available at* http://www.uscourts.gov/RulesAndPolicies/FederalRulemaking/ResearchingRules/AgendaBooks.aspx ("the Civil Rules only address pending actions").
72 Conference Committee Report, *supra*, at 12.

M. Adams & Associates v. Dell, Inc., for example, the court held that the duty to preserve was triggered many years before suit was filed because other "computer and component manufacturers [in the industry] were sensitive" to the issue.[73] As one participant has observed, "[p]rospective litigants are at serious risk of committing spoliation—passively, actively, unintentionally—before litigation commences (if it ever does) because they have no codified benchmarks' to which to conform their behavior."[74]

Upon closer examination, the Enabling Act concern seems overblown. While Rule 1 states that the civil rules govern procedure "*in* the United States district courts," reliance on that language conveniently ignores the existence of Rule 27 (involving pre-commencement depositions used to perpetuate testimony)[75] which was approved by the Supreme Court and Congress - and the fact that courts routinely assess pre-litigation conduct wearing their inherent power hat. In *Silvestri v. General Motors*[76] and in *Goodman v. Praxair Services*,[77] for example, courts issued sanctions despite the fact that the discoverable evidence at issue was disposed of before the lawsuits were filed.

The test of Enabling Act jurisdiction,[78] or, for that matter, the use of inherent judicial power, is the relationship of the conduct to be regulated to the functioning of the courts. The mere fact that an action has not yet been commenced is not decisive. In *Chambers v. NASCO,*[79] the majority approved, over a dissent by Justice Kennedy, sanctions relating to pre-commencement conduct which was intimately related to the appropriate resolution of the case. *See* 501 U.S. at 55, n. 17 ("[a]lthough the fraudulent transfer of assets took place before the suit was filed, it occurred after Chambers was given notice, pursuant to court rule, of the pending suit. Consequently, the sanctions imposed on Chambers were aimed at punishing not only the harm done to NASCO, but also the harm done to the court itself.").

Rules seeking to limit the adverse impact of conduct on the functioning of discovery are well within the rulemaking power. As the Supreme Court noted in *Business Guides, Inc. v. Chromatic Comm. Enterprises, Inc.*,[80] Rule 11 is authorized since its "main objective" is to promote the judicial process by curbing abuses.[81] In *Shady Grove v. Allstate*,[82] the Court more recently noted that a rule which "regulate[s] only the process for enforcing [parties] rights" and not "the available remedies, or the rules of decision by which the court adjudicated" is clearly permissible.

Finally, Congress has, under the Enabling Act, reserved the power to review, revise and adopt changes to any rules that are proposed, including those that touch on pre-litigation conduct. Rules which survive that review have the same force of law as if directly enacted by Congress.[83] In *Chambers*, *supra,* the Supreme Court noted that "the exercise of the inherent power of lower federal courts can be limited by statute and rule, for '[t]hese

73 621 F. Supp. 2d 1173, 1191 (D. Utah 2009).
74 Gregory Joseph, *Electronic Discovery and Other Problem,* at 2-3, CONFERENCE PAPERS.
75 Rule 27 provides for limited discovery "before an action is filed" to "prevent a failure or delay of justice."
76 271 F. 3d 583, 590 (4th Cir. 2001).
77 632 F. Supp. 2d 494, 505 (D. Md. July 7, 2009).
78 28 U.S.C. § 2072 (a-b) (The Supreme Court shall have the power to prescribe "general rules of practice and procedure" provided they do not modify "substantive" rights).
79 501 U.S. 32 (1991). Justice Kennedy refused to accept this approach. *See* Kennedy, J., dissenting, at 74 ("By exercising inherent power to sanction pre-litigation conduct, the District Court exercised authority where Congress gave it none.").
80 498 U.S. 533 (1991).
81 *Id.* at 553.
82 130 S. Ct. 1431 (2010).
83 As Judge Posner has noted, "when a domain of judicial action is covered by an express rule, such as Rules 26 and 37 of the civil rules, the judge will rarely have need or justification for invoking his inherent power." *Fidelity National Title Insurance Co. v. Intercounty National Title,* 412 F.3d 745, 752 (7th Cir. 2005).

courts were created by act of Congress.'"[84] The Court has also indicated a strong preference for resolving fundamental civil discovery obligations by rulemaking, given the benefits of the practical and transparent process involved.[85]

III. Possible Rules

Reliance on ad hoc inherent power to articulate the duty to preserve has resulted in contradictory rulings and different formulaic approaches in different Circuits. In *Pension Comm. v. Bank of Am. Sec., LLC,*[86] for example, the court exempted moving parties from having to demonstrate that relevant and discoverable evidence was missing in order to seek sanctions merely because of the lack of written litigation holds. Other courts have reached diametrically opposed conclusions on similar facts.[87]

In addition, by relying on inherent power, not the specific and targeted provisions of Rule 37, the "wrong reason[s]" may be advanced[88] for the imposition of sanctions in the absence of a showing of egregious conduct. [89]

It is time for the Civil Rules to include duties relating to preservation.[90] This would not be the first time that rulemaking has superseded court-developed common law applied by inherent powers. In 1983, the Supreme Court acted to provide rule-based guidance in order "to obviate dependence upon" the "court's inherent power to regulate litigation."[91] However, as in the case of all rulemaking, the "devil is in the details," and great care must be taken not to exacerbate the very trends which have made preservation such a problem in the world of modern discovery.

A. Trigger of the Obligation

To help address the confusion inherent in assessing the "forseeability" of litigation, which helps illuminate when a party should have been aware of the need to preserve, the E-Discovery Panel recommended articulation of specific actions which would unequivocally trigger knowledge.[92] These examples could be included in a Committee Note. Some typical examples could include the service or delivery of a document such as a request or demand to preserve, a subpoena, CID or similar inquiry. When the shoe is on the other foot—i.e., when a party intends to initiate litigation or submit a counterclaim—the Rule or Committee Note could identify as triggering conduct the steps taken in anticipation of asserting or defending claims, such as preparation of reports, hiring of experts, presenting claims to regulators, hiring counsel and the like.[93]

---

84 501 U.S. 32 at 48.

85 *Mohawk Industries, Inc. v. Carpenter,* 130 S. Ct. 599, 609 (2009) (rulemaking "draws on the collective experience of bench and bar" for "measured, practical solutions").

86 685 F. Supp. 2d 456 (S.D.N.Y. Jan. 15, 2010) (holding that it was gross negligence to fail to issue written litigation holds). The opinion was subtitled "*Zubulake* Revisited: Six Years Later," and was amended on May 28, 2010.

87 *Kinnally v. Rogers Corporation,* 2008 WL 4850116, at *6 (D. Ariz. Nov. 7, 2008) ("the absence of a written litigation hold . . . does not in itself establish [a violation]" (emphasis in original). A strict liability approach was also rejected in the context of Rule 26(b)(2)(B) by *Major Tours v. Colorel* ("Major Tours III"), 2010 WL 2557250, at *28 (D.N.J. June 22, 2010) ("The Rules compel [a] discretionary balancing . . . not a bright line requirement of production if a party "fails to adequate preserve every byte of previously accessible data.").

88 John M. Barkett, *Walking the Plank, Looking Over Your Shoulder, Fearing Sharks Are in the Water: E-Discovery in Federal Litigation?*, Conference Papers, at 35.

89 *Rimkus Consulting v. Cammarta,* 2010 WL 645353, at *5 and *7 (S.D. Tex. Feb. 19, 2010) (the Supreme Court's decision in *Chambers* may require a degree of culpability greater than negligence since, in that case "the inherent power was linked to the bad-faith conduct that affected the litigation."); *cf. United Medical Supply v. United States,* 77 Fed. Cl. 257, 268 (Fed. Cl. 2007) (it is a "huge logical leap" to suggest that Chambers "limit[s] sanctions to cases in which there is a showing of bad faith").

90 Conference Committee Report, supra, at 12 ("it may be possible to focus on provisions that address" this narrow source of preservation duties).

91 Rule 16, Committee Note, Subdivision (f) (1983) (dealing with failure to comply with Rule 16).

92 *See generally* E-Discovery Panel, "Elements of a Preservation Rule" ("Preservation Elements"), ¶ 1(Trigger), Conference Papers.

93 *Id.* at ¶ 1(b)(v).

Some have argued that a neutral cut-off point would be preferable to one based on forseeability. Thus, Professor Martin Redish has suggested establishing the trigger at a fixed point, such as the service of a discovery request or, if opposed, issuance of a discovery order.[94] The New York City Bar[95] has also suggested an objective retroactive limitation on preservation obligations. Under that proposal, "no sanctions [would be possible] for loss of data occurring more than one year prior to receipt of (i) a preservation demand letter; or (ii) the filing of a complaint, which ever comes first."

However, regardless of the method utilized, courts should concentrate on assessing the culpability of parties *at the time of loss* since "[t]he ultimate focus for imposing sanctions for spoliations of evidence is the intentional destruction of evidence indicating a desire to suppress the truth, not the [mere] prospect of litigation." [96]

B. Components of the Duty to Preserve

A preservation rule should require reasonable efforts, not extraordinary, excessive, disproportionate or unduly burdensome actions.[97] A "cost-benefit" balance should be applied so that the efforts are not "disproportionate to the potential value" of the information at issue. There are, admittedly, difficulties in applying the cost-benefit approach to preservation issues, given that costs are not always the sole factors at issue in litigation.[98] Nonetheless, it makes sense to embody the principle in the rule in addition to the requirement of a reasonable effort.

Thus, a standalone provision (e.g., Rule 34.1)[99] could provide:

> "Parties with actual or constructive notice of the likelihood that relevant and discoverable evidence is or will be sought in discovery shall undertake reasonable and proportionate efforts to preserve any such evidence within its possession, custody or control subject to the considerations of Rule 26(b)(2)(C) and Rule 37(e)."

However, experience also indicates the need for more explicit guidance to provide a substantial measure of certainty for preservation planning purposes and to form a "checklist" for disclosures and Rule 26(f) and 16(b) discussions.

One approach would be to specify types of ESI which need not be preserved absent agreement or a court order. Such a provision has been successfully implemented as a key part of the Seventh Circuit E-discovery Pilot Program. It would be analogous to Rule 26(b)(2)(B), added in 2006, which exempts *production* of ESI from inaccessible sources in the absence of a showing of good cause and has served as a model for preservation obligations.[100]

---

94 *See* Martin R. Redish, *Electronic Discovery and the Litigation Matrix,* 51 DUKE L.J. 561, 624-25 (2001) (advocating trigger of the duty to preserve upon receipt of discovery requests unless destruction took place before time when otherwise normally scheduled for destruction).

95 New York City Bar, *Proposals for the 2010 Duke Conference Regarding the Federal Rules of Civil Procedure*, at 5 (unnumbered), CONFERENCE PAPERS.

96 *Greyhound Lines v. Wade,* 485 F.3d 1032 (8th Cir. 2007).

97 The Sedona Conference® *Commentary on Preservation, Management and Identification of sources of Information That Are Not Reasonably Accessible*, 10 SEDONA CONF. J. 281, 282, 292 - 293 (2009) ["Step Five Analysis - The Proportionality Principle"] (the nature of the storage media and the characteristics of the information play a role in determining if burdens and costs of preservation outweigh the potential benefits).

98 *See* CCL Critique, at 4, CONFERENCE PAPERS ("[c]ost is a legitimate concern in adjudicating disputes, but mandating cost-benefit analysis 'at all times' [citing to ACTL/IAALS proposed Pilot Rule 1.2] is neither desirable nor practical.").

99 A proposed Rule 34.1 ("Duty to preserve") was also distributed for discussion purposes to attendees at the E-Discovery Conference in February, 2004 prior to the 2006 Amendments. *See* FORDHAM E-DISCOVERY CONFERENCE PARTICIPANT MEMO (2004), at 35, *available at* http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/E-Discovery_Conf_Agenda_Materials.pdf.

100 *Cache La Poudre Feeds, LLC v. Land O Lakes, Inc.,* 2007 WL 684001 at *15 (D. Colo. Mar. 2, 2007) (the duty to preserve "would not automatically include information maintained on inaccessible computer backup tapes.").

Thus, a second provision could provide (in the Rule or in a Committee Note) that in the absence of a court order or prior agreement, the necessity of preservation of the following categories would not be required:

(1) Deleted, slack, fragmented or unallocated data on hard drives
(2) Random access memory (RAM) or other ephemeral data
(3) On-line access data such as temporary internet files
(4) Data in metadata fields that are frequently updated; such as last opened dates
(5) Backup data that is substantially duplicative of more accessible data available elsewhere, and
(6) Other forms of ESI which require extraordinary affirmative measures not utilized in the ordinary course of business."[101]

This provision would help reduce the obvious unfairness of "sandbagging" a producing party with unanticipated preservation burdens.[102] Courts would be urged to deal promptly with any disputes at the Rule 16 conference or in response to a motion for a preservation or protective order.[103] The right to seek sanctions would be waived by parties failing to take advantage of the opportunity to discuss and resolve contested issues.[104] Something like this approach was successfully applied in the landmark decision in *Columbia Pictures v. Bunnell*,[105] where a duty to preserve information temporarily stored in RAM was held to arise only after motion and a showing of the necessity for the retention.

A related approach would be to provide presumptive limitations on the total number of "key custodians" and information systems whose relevant information must be preserved and produced.[106] For example, up to 10 custodians (depending upon the value of the case as determined by the demand or by the court) and an equal number of information systems could be a presumptive maximum that a potential producing party would be responsible to address.[107] This would provide a strong incentive for parties and counsel to take advantage of the Rule 26(f) meeting process. The numbers, of course, would be subject to modification in individual cases.

C. Sanctions

Spoliation sanctions are but one form of discovery sanctions,[108] yet under the current regime they are routinely imposed without guidance from Rule 37. Once the preservation obligation is decoupled from the common law spoliation doctrine and

101 For a similar approach, *see Lawyers for Civil Justice White Paper,* CONFERENCE PAPERS, at 36-37 (Rule 26(h) Specific Limitations on Electronically Stored Information).
102 *Frey v. Gainey Trans. Services,* 2006 WL 2443787, at *9 (N.D. Ga. Aug. 22, 2006) (refusing spoliation sanctions where demand letter arguably was intended to "sandbag" party if ignored).
103 *See* 7TH CIR. PILOT PROGRAM, PRINCIPLE 2.04 (Scope of Preservation) ("if any party intends to request the preservation or production of these categories, then that intention should be discussed at the meet and confer or as soon thereafter as practicable").
104 *Healthcare Advocates v. Hardin, Earley, Follmer & Frailey,* 497 F. Supp. 2d 627 (E.D. Pa. 2007) (finding no duty to preserve contents of cache files where preservation letter did not alert them to the need to do so).
105 245 F.R.D. 443 (C.D. Cal. 2007), denying motion to reverse order regarding preservation of server log data, 2007 WL 2080419 (C.D. Cal. May 29, 2007) (emphasizing "its relevance and the lack of other available means to obtain it").
106 This approach has worked well in the production context. *See, e.g.,* Rule 30(a)(2)(A)(no more than 10 depositions); Rule 33(a)(no more than 25 written interrogatories); *see also* Rule 30(d)(1) (deposition limited to 1 day of 7 hours unless otherwise stipulated or ordered by the court). *See generally*, Gregory S. Weber, *Potential Innovations in Civil Discovery: Lessons for California From the State and Federal Courts*, 32 MCGEORGE L. REV. 1051 (Summer 2001).
107 *Cf. Lawyers for Civil Justice White Paper*, *supra,* at 32 CONFERENCE PAPERS (suggesting a presumptive limitation of "a reasonable number of custodial or other information sources for production, not to exceed 10"). CPR (the International Institute for Conflict Prevention & Resolution) has incorporated a similar approach in its model Economical Litigation Agreement (2010)(copy on file with author).
108 *Casale v. Kelly,* 2010 WL 1685582, at *10 (S.D.N.Y. Apr. 26, 2010) (describing remedies for "failing to preserve" as a form of "discovery sanctions").

incorporated in the Civil Rules, it would be comparatively easy to adapt Rule 37 to treat preservation infractions under that Rule. Rule 37 already incorporates most of the traditional discretionary remedies needed to address such challenges.[109]

Thus, Rule 37(b)(2)(A) could provide that "[if a party] fails to obey an order to *preserve evidence or* provide or permit discovery," it would apply. Rule 37(c)(1) could be amended so that it would apply "[if a party] fails to *preserve or* provide information as required by these rules or identify a witness as required by rule 26(a) or (e)."

Rule 37 would then provide the basis for promoting the uniform treatment of the sanction issues among the Circuits. Rule 37 instructs courts to assess whether sanctions are "substantially justified"[110]or were not "unjust."[111] Examples could be given in the Committee Notes of whether or how to identify the relevance of the information alleged to have been lost and the prejudice suffered by the loss.[112]

The Committee Note could also address the issue of whether there is conduct that presumptively satisfies the requisite state of mind to justify sanctions under Rule 37 standards. In *Scalera v. Electrograph Systems,*[113] for example, internal counsel orally instructed key players to retain email, collected relevant files and was satisfied that backup media was secure and could be accessed. The court denied sanctions.

D. Rule 37(e)

In most circuits, proof of bad faith—evidencing a subjective intent to interfere with access to discoverable evidence—is required to justify imposition of serious sanctions.[114] However, the contrary is true in other circuits, where mere "negligence" is sufficient. Rule 37(e) was enacted as part of the 2006 Amendments to clarify that in the case of "routine" losses—which can occur before or after a duty to preserve attaches—rule-based sanctions are inapplicable provided the party acted in good faith, an "intermediate" standard which provides for the absence of bad faith. Thus, conduct falling within that scope is not to be treated differently even if the governing law of the circuit in which the action is pending would do so, absent Rule 37(e).

Unfortunately, some courts have interpreted an ambiguous Committee Note to Rule 37(e)[115] as a mandatory duty to take specific action, regardless of the need to so to effectuate preservation, thereby barring application of Rule when a duty to preserve is identified and the action is not taken. The author of the *Zubulake* opinions is quoted as arguing that "it can't be routine and good-faith not to suspend your process once you know

109 Rule 37(b)(2)(A), for example, authorizes the issuance of orders establishing or opposing "designated facts," the striking of "pleadings" or the entry of a "default judgment" or "dismissal." Rule 37(c) bars use of information or a witness to supply evidence on a motion, at a hearing, or at a trial and mandates payment of reasonable expenses, including attorney's fees, under many circumstances.

110 *Devaney v. Continental American Insurance,* 989 F.2d 1154, 1163 (11th Cir. 1993) (determination turns on whether "reasonable people could differ as to the appropriateness of the contested action.").

111 *Lewis v. Ryan,* 2009 WL 3486702 at *6 (S.D. Cal. Oct. 23, 2009) (not "unjust" to sanction where defendants knew documents were relevant but allowed them to be destroyed); *Webb v. The District of Columbia,* 146 F.3d 964, 971 (D.C. Cir. 1998) (the same "considerations" apply to Rule-based sanctions as have traditionally been applied to sanctions issued under authority of inherent powers).

112 Preservation Elements, *supra*, ¶ 7(d), CONFERENCE PAPERS.

113 262 F.R.D. 162, 178-179 (E.D.N.Y. 2009).

114 *Bakhtiari v. Lutz,* 507 F. 3d 1132 (8th Cir. 2007) (refusing to sanction pre-suit destruction of email through automatic deletion since the "ultimate focus" should be on proof of intentional destruction).

115 The Committee Note to Rule 37(e) provides that "[w]hen a party is under a duty to preserve information because of pending or reasonably anticipate litigation, intervention in the routine operation of an information system is one aspect of what is often called a 'litigation hold.'"

there is litigation."[116] However, "if the party cannot avail itself of the safe harbor because it had a duty to preserve data in the first instance, then Rule 37 does little to change the state of the pre-existing common law."[117]

Thus, Rule 37(e) should be clarified to provide that:

> "Absent exceptional circumstances, a court may not impose sanctions under these rules on a party for failing to provide electronically stored information or tangible things lost as a result of the routine, good-faith operation of a system or process in the absence of a showing of intentional actions designed to avoid known preservation obligations."[118]

This would be consistent with the Private Securities Litigation Act (the "PSLRA")[119] and recommendations by Lawyers for Civil Justice and the ACTL/IAALS Pilot Project Rules.[120] It shuld also be broadened to include all forms of discoverable evidence, not just electronically stored information.[121]

IV. CONCLUSION

The preservation doctrine belongs in the Federal Rules, where it can be linked to and supportive of discovery obligations and whose compliance can be assessed by the provisions of an expanded Rule 37. Additional rulemaking involving practical standards which are "up to the task"[122] is feasible and should be discussed at the upcoming Advisory Committee meetings. Despite the hesitancy of the Advisory Committee to act in 2006, the time has come for action.

116 Panel Discussion, *Sanctions in Electronic Discovery Cases: Views from the Judges,* 78 FORDHAM L. REV. 1, 30-31 (Oct. 2009)("what this toothless thing [Rule 37(e) really tells you is the flip side of a safe harbor. It says if you don't put in a litigation hold when you should there's going to be no excuse if you lose information.").

117 Emily Burns, Michelle Greer Galloway & Jeffrey Gross, *E-Discovery: One Year of the Amended Federal Rules of Civil Procedure,* 64 NYU ANN. SURV. AM. L. 201, 217 (2008).

118 *See also Lawyers for Civil Justice White Paper*, CONFERENCE PAPERS, at 38.

119 *Danis v. USN Communications,* 2000 WL 1694325, at *32, n. 20 (N.D. Ill. Oct. 23, 2000) (noting that under the PSLRA [15 U.S.C. §78u-4(b)(3)(C)(i)] sanctions may be imposed under the PSLRA only for willful document destruction.").

120 ACTL/IAALS Pilot Project Rules, *supra*, at 7 CONFERENCE PAPERS ("only upon a showing of intent to destroy evidence or recklessness"); *cf.* CCL Critique, *supra,* at 24, CONFERENCE PAPERS.

121 Letter, American College of Trial Lawyers to Advisory Committee (Jan. 25, 2005), at 4 ("[i]f a safe harbor is introduced into the Rules [which the ACTL then supported in principle], it should extend to all types of information").

122 *Chambers v. NASCO,* 501 U.S. 32, 50 (1991).

A NUTSHELL ON NEGOTIATING E-DISCOVERY SEARCH PROTOCOLS

Jason R. Baron & Edward C. Wolfe[1]
Washington, DC Bloomfield, MI

The aim of this Nutshell is to provide practical guidance to counsel on the subject of conducting search negotiations as part of the Fed. R. Civ. P. 26(f) meet and confer process or in the context of negotiating a Rule 16 pre-trial order. The outline is intended to highlight key concepts and approaches that have emerged since the revisions to the Federal Rules of Civil Procedure went into effect in December 2006. Due to the pace of technological change and new case developments, the authors understand that what is said here may be in need of revision and updating shortly after its publication date, and thus counsel is cautioned to remain vigilant in keeping up with emerging case law and commentary in this area.

A. General Guidance

1. Conducting a Reasonable, Comprehensive Search Pursuant to Fed. R. Civ. P. 34

The Sedona Conference® *Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery* ("*Sedona Search Commentary*") addressed the subject of how lawyers search for electronically stored information (ESI), providing a roadmap to counsel on the practical limitations of keyword searching as practiced by the profession, with practice points given on how to approach the task of satisfying one's professional obligations to conduct a reasonable, comprehensive search in response to a Rule 34 request for documents and ESI.[2] The *Sedona Search Commentary's* eight practice points, as set out and further explained therein, are as follows:[3]

> *Practice Point 1. In many settings involving electronically stored information, reliance solely on a manual search process for the purpose of finding responsive documents may be infeasible or unwarranted. In such cases, the use of automated search methods should be viewed as reasonable, valuable, and even necessary.*
>
> *Practice Point 2. Success in using any automated search method or technology will be enhanced by a well-thought out process with substantial human input on the front end.*

---

1 Jason R. Baron serves as Director of Litigation at the National Archives and Records Administration. Edward C. Wolfe was with the Office of the General Counsel of General Motors Company in Detroit for over 30 years. Both authors presently serve on the Steering Committee for The Sedona Conference® Working Group on Electronic Document Retention and Production (WG1). The views expressed herein are the authors alone and do not necessarily represent the views of the institutions with which they are or have been affiliated.

2 The Sedona Conference,® *The Sedona Conference® Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery*, 8 SEDONA CONF. J. 189 (2007), *available at* http://www.thesedonaconference.org/content/miscFiles/publications_html.

3 *Id.*

Practice Point 3. The choice of a specific search and retrieval method will be highly dependent on the specific legal context in which it is to be employed.

Practice Point 4. Parties should perform due diligence in choosing a particular information retrieval product or service from a vendor.

Practice Point 5. The use of search and information retrieval tools does not guarantee that all responsive documents will be identified in large data collections, due to characteristics of human language. Moreover, differing search methods may produce differing results, subject to a measure of statistical variation inherent in the science of information retrieval.

Practice Point 6. Parties should make a good faith attempt to collaborate on the use of particular search and information retrieval methods, tools, and protocols (including as to keywords, concepts, and other types of search parameters).

Practice Point 7. Parties should expect that their choice of search methodology will need to be explained, either formally or informally, in subsequent legal contexts (including in depositions, evidentiary proceedings, and trials).

Practice Point 8. Parties and the courts should be alert to new and evolving search and information retrieval methods.[4]

2. Employing Quality Control Techniques

Additional guidance on the subject of employing quality control techniques as part of the search process, including sampling and iterative methods, was provided in The Sedona Conference® *Commentary on Achieving Quality in the E-Discovery Process* ("*Sedona Achieving Quality Commentary*").[5] In particular, Principle 3 in the *Sedona Achieving Quality Commentary* states in relevant part that "[i]mplementing a well thought out e-discovery 'process' should seek to enhance the overall quality of the production" in terms of time, cost, accuracy and completeness. As further explained, this includes "using iterative and adaptive processes that allow for learning and correction, and, where appropriate, making use of statistically valid metrics in order to monitor progress and obtain valid measures of the accuracy of the effort."[6]

3. Cooperation During the Discovery Process

In 2009, The Sedona Conference® *Cooperation Proclamation* declared that "cooperation by all parties in the discovery process" promotes achievement of the goal of a "just, speedy and inexpensive determination of every action," consistent with the dictates of Fed. R. Civ. P. 1. Among the methods included in the *Cooperation Proclamation* aimed at accomplishing cooperation among counsel are

---

4 *See generally Victor Stanley, Inc. v. Creative Pipe, Inc.,* 250 F.R.D. 251, 261-263 (D. Md. May 29, 2008) (discussing the *Sedona Search Commentary* practice points).

5 The Sedona Conference,® *The Sedona Conference® Commentary on Achieving Quality in the E-Discovery Process*, 10 SEDONA CONF. J. 299 (2009), *available at* http://www.thesedonaconference.org/content/miscFiles/publications_html.

6 *See Pension Committee of the Montreal Pension Plan v. Banc of America,* 2010 WL 93124, *3 (S.D.N.Y. Jan. 15, 2010) (Scheindlin, J.) ("failure to assess the accuracy and validity of selected search terms" constitutes "negligence"); *Victor Stanley, supra,* 250 F.R.D. at 256 (defendants "regrettably vague" on how keywords searched "were developed, how the search was conducted, and what quality controls were employed to assess their reliability and accuracy"); *William A. Gross Construction Associates, Inc. v. American Manufacturers Mutual Insurance Co.,* 2009 WL 724954 (S.D.N.Y. Mar. 19, 2009) ("this Opinion should serve as a wake up call to the Bar … about the need for careful thought, quality control, testing and cooperation with opposing counsel in designing search terms or 'keywords' to be used"); *In re Seroquel Prods. Liab. Litig.,* 244 F.R.D. 650, 660 n.6, 662 (M.D. Fla. 2007) (criticizing defendant's failure to "assure reasonable completeness and quality control" in search for relevant material).

- "Exchanging information on relevant data sources, including those not being searched..."
- "Jointly developing automated search and retrieval methodologies to cull relevant information."[7]

"The Case for Cooperation," published as a supplement to volume 10 of *The Sedona Conference Journal*,® goes on to say:[8]

> "[W]orking cooperatively with opposing counsel to identify a reasonable search protocol, rather than making boilerplate objections to the breadth of a requested protocol or unilaterally selecting the keywords used without disclosure to opposing counsel, may help avoid sanctions or allegations of intentional suppression. Indeed, because knowledge of the producing party's data is usually asymmetrical, it is possible that refusing to 'aid' opposing counsel in designing an appropriate search protocol that the party holding the data knows will produce responsive documents could be tantamount to concealing relevant evidence." [9]

4. Placing "Search Negotiations" in Context: Custodians, Dates, and Scope

In order to appropriately search for responsive and relevant ESI, the responding party will need to identify key custodians familiar with the allegations in the pleadings filed in a case. They are often in the best position to identify what types of information relate to the matter and where they are stored. Information gleaned from these custodians can greatly assist, in addition to the allegations in the pleadings, in determining preservation requirements and whether the information retention schedules would suggest what information is likely to be extant.

B. Search Methods 101: Keywords and Their "Alternatives"

All lawyers are familiar with keyword searching from Westlaw and Lexis searches performed against structured bodies of case law and legislation. However, as the *Sedona Search Commentary* addresses at length, and a growing body of case law recognizes, the exponential growth of ESI coupled with the ambiguities of human language pose profound challenges to constructing efficacious keyword searches for the purpose of finding all or most relevant documents in a given collection. *See Sedona Search Commentary, passim.* Accordingly, any negotiations over search terms must start with the assumption that simple recitation of "keywords," without further refinement, will necessarily end up being both under-inclusive (as there may be relevant documents that fail to contain the string of letters comprising a given keyword), and over-inclusive (in that a potential huge number of false positive, nonrelevant "hits" can be expected with the input of any common term). Parties should recognize the danger in proposing (or accepting a search based on) large numbers of keywords.[10]

---

7 The Sedona Conference,® *Cooperation Proclamation,* 10 SEDONA CONF. J. 331, 332 (2009 Supp.).

8 The Sedona Conference,® *The Case for Cooperation,* 10 SEDONA CONF. J. 339 (2009 Supp.).

9 *See also Dunkin' Donuts Franchised Restaurants, Inc. v. Grand Central Donuts, Inc.,* 2009 WL 1750348 (E.D.N.Y. June 19, 2009) (citing the *Cooperation Proclamation*, parties "are directed to meet and confer on developing a workable search protocol").

10 *See In re Fannie Mae Litigation,* 552 F.3d 814 (D.C. Cir. 2009) (appellate court upholds contempt citation against government agency, where it had failed to meet court-imposed deadlines after agreeing to produce non-privileged documents found responsive to 400 keyword search terms, where the production set consisted of 660,000 recovered documents that needed to be, but could not be, reviewed in time); *see also William A. Gross Construction, supra* (court notes 1,000 proposed search terms put forward by one party, including very generic ones); *Kipperman v. Onex Corporation,* 2008 WL 4372005 (N.D. Ga. Sept. 19, 2008) (defendants' motion for relief from having to review and produce all results from plaintiffs' proposed searches denied, where defendants had failed to timely object to scope of email search both with respect to named end users and search terms).

As the *Sedona Search Commentary* also recognized, due to errors inherent in the transcribed words in texts, keyword searches that are limited to correct spellings, without also including commonly misspelled variants of words, or variations of words using different word stems, are at risk of being incomplete. Counsel should be alert to the need to account for possible variants of words, and consideration should be given to software that employs principles derived from "fuzzy logic."[11]

As the *Sedona Search Commentary* goes on to note, "[l]awyers are beginning to feel more comfortable using alternative search tools to identify potentially relevant" ESI, including language-based and statistical tools that fall under the umbrella of "concept based" methods. United States Magistrate Judge Judge M. Facciola, writing in *Disability Rights Council of Greater Washington, et al. v. Washington Metropolitan Transit Authority,*[12] was the first to suggest in a reported case that parties consider discussing "concept searching" as a possible alternative to keyword searching. Chief Magistrate Judge Paul W. Grimm's *Victor Stanley* opinion goes on to list a number of such methods for parties to generally consider.[13] Although such tools and techniques are coming into greater use, there is as of this writing no reported case law where one or more parties have proposed an alternative search method be employed giving rise to a disagreement that needed court adjudication. Ideally, as such techniques come into their own, they will be subject to utilization and adoption in search protocols memorialized at meet and confers.[14]

A growing number of lawyers in particular are employing statistical clustering techniques that group together "like" documents (documents with similar terms or concepts) into categories. Doing so greatly decreases the amount of time of time needed in manual document review for purposes of responsiveness and privilege.[15] The extent to which a party's decision to employ this type of emerging technique is an appropriate (*e.g.,* nonprivileged) subject within the scope of a negotiated search protocol is an open issue.

C. Approaching the Search Negotiation: Strategies, Models, and Best Practices

While there has been much discussion about the efficacy of search terms, just as there has been with human review, one must begin somewhere, and search terms are commonly used and when used appropriately with a sound methodology, can yield good results. The traditional approach in this area has been for the responding party, after a review of the pleadings and in consultation with key custodians, to draft a list of terms to be searched. Increasingly parties are finding that a sample search can then be run to determine if the terms are likely to be under- or over-inclusive and can then be refined. In addition, a sample of information that has not been included as a result of the search can be reviewed to determine if other terms or alterations should be made.

11 *See Kay Beer Distributing, Inc. v. Energy Brands, Inc.,* 2009 WL 649592 (E.D. Wis. June 10, 2009) (holding that defendant must conduct a search for variants of plaintiff's name).

12 242 F.R.D. 139 (D.D.C. 2007).

13 250 F.R.D. 251 at 259 n.9 (D. Md. May 29, 2008).

14 *See generally United States v. O'Keefe,* 537 F. Supp. 2d 14 (D.D.C. Feb. 18, 2008) (in a criminal case, the court ordered further explanation of whether keyword searches were thorough, citing to authorities arising in civil case law, and suggested that in light of interplay of the sciences of computer technology, statistics, and linguistics, expert testimony may be needed in this complex area).

15 Ronni D. Solomon & Jason R. Baron, *Bake Offs, Demos & Kicking the Tires: A Practical Litigator's Brief Guide to Evaluating Early Case Assessment Software & Search & Review Tools,* Apr. 9, 2007, *available at* http://www.kslaw.com/Library/publication/BakeOffs_Solomon.pdf.

1. The Pros and Cons of Adopting the Requesting Party's Requests With Respect to Keyword Search Terms and Methodologies

A requesting party may have a legitimate, good faith belief that they are sufficiently informed regarding the causes of action at issue and underlying facts so as to be able to propose well-formed search queries, including through the use of keywords. As the propounder of the eventual discovery requests, the requesting party is in the best position to know what it believes are the most salient aspects of the case that are in need of discovery in the first place. To the extent the producing party is willing to allow the requesting party to control all or some of the keyword search, without raising a threshold objection, doing so holds out the possibility of significantly reducing the level of conflict and subsequent motion practice in discovery. An early agreement on search protocols also may be strategically valuable in diminishing the other side's ability to object to terms and methods that have been made subject to prior agreement (especially if they entail large resources or workloads). Cooperation in the form of reaching agreement on search terms ultimately reduces the legal risk in having to undertake new and different searches through large collections of data.

On the negative side, allowing a party with less information and less access to a data set to control the terms of a search often results in proposals for long lists of terms, simply due to unfamiliarity with the universe of "internal" terms that may be used, as well as the understandable desire to capture as much requested information as possible. This can result in the inclusion of considerable unresponsive information that is expensive and burdensome to both parties and slows advancement to reaching the true issues remaining in contention.[16]

2. The Pros and Cons of Allowing the Producing Party to Control the Keyword Search and Methodology

Given the nature of the discovery process, allowing the producing party to control the keyword search and other forms of search, at least in the first instance, makes eminent sense given that the producing party knows (or at least enjoys greater access to) the data, custodians, and internal terminology that will be made subject to the request. The producing party may hold out that upon receipt of properly framed requests for production of documents and ESI, they will undertake to first formulate their search strategy in light of a good faith interpretation of the content of the discovery requests received. If clarification from the requesting party is necessary to narrow the scope of the request, it will be forthcoming. Proceeding in this fashion does not necessarily mean defaulting to a traditional, unilateralist approach in responding to discovery, including the assertion that all efforts behind the scenes to construct a search strategy should be considered "privileged." Rather, the producing party would be representing that they are merely better positioned to begin the process while allowing for transparency with respect to search results at a later stage of the process. Indeed, a disadvantage can occur if there is not sufficient due diligence in the process to arrive at appropriate terms.

---

16 *See generally Wixon v. Wyndham Resort Development Corp.,* 2009 WL 3075649 (N.D. Cal. Sept. 21, 2009) (broader terms suggested by requesting party held reasonable); *In re Direct Southwest, Inc. Fair Labor Standards Act Litigation,* 2009 WL 2461716 (E.D. La. Aug. 7, 2009) (court accepts requesting parties' search terms over defendants' objections); *Flying J Inc. TCH LLC v. Pilot Travel Centers, LLC,* 2009 WL 1834998 (D. Utah June 25, 2009) (court to rule on search protocol after requiring justification from requesting party for 28 specific terms, subject to any objections lodged by defendant); *Capitol Records v. MP3 Tunes,* 261 F.R.D. 44 (S.D.N.Y. Aug. 13, 2009) (court steps in to adjudicate dispute after court "directed counsel to confer further in an attempt to agree on search terms," and where nine terms were agreed to with 30 remaining in dispute).

As stated in the *Sedona Achieving Quality Commentary*, the first step in constructing any automated search involves the ability "to effectively share and transfer knowledge among counsel and the managing team and those with knowledge of the corpus of ESI that is the subject of discovery. The knowledge gained in this process will be used in the development of one or more search strategies (*e.g.,* Boolean searches, concept searches, metadata filters, language-based approaches using taxonomies and ontologies, statistical clustering techniques, or other proprietary strategies)."

One leading commentator has written extensively on the tension inherent in allowing a requestor to control the keyword search protocol versus what he views as proper reliance on the producing party to re-assert traditional authority in this area, especially in light of the asymmetry in the position of the parties vis-à-vis knowledge of the ESI repository at issue.[17]

3. Adoption of a Phased or Iterative Approach

The rules require the parties to engage in at least one meet and confer; however, the parties have ample opportunity during the discovery process to engage in further discussions. One model for phased interactions is as follows:[18]

> *Step 1. The parties meet and confer on the nature of each others' computer hardware and software applications. Proposals are exchanged on the scope of search obligations, in terms of databases and applications to be searched, what active and possibly legacy media, key custodians, time periods. Additionally, keywords are proposed along with any other more sophisticated Boolean or concept search methods. A timetable for conducting searches after the propounding of discovery requests is agreed to.*
>
> *Step 2. In the interval between meet and confers, parties conduct searches in accordance with prior representations and the actual wording of discovery requests. In doing so they may utilize sampling techniques, and estimates are gathered on the volume of data or "hits" made subject to search.*
>
> *Step 3. The parties interact further in describing the result of initial searches and preliminary results. If the parties have agreed to a Rule 502 rubric, the parties may elect to share documents found to be potentially responsive. Search terms and protocols are adjusted and search methods are tuned or adjusted for the purpose of conducting more narrow, focused searches.*
>
> *Step 4. The parties may elect to continue iteratively until a mutually agreed time or cap on numbers of responsive documents is reached.*

As noted, a variation on this approach is to presume in Step 1 that the responding party or parties in control of the data to be searched take the lead in first conducting a search under unilaterally arrived at keywords and other search methods. Doing so does not in theory alter further following remaining steps 2 through 4 *supra*. The efficacy of

---

17 Ralph Losey, *Child's Game of Go Fish Is A Poor Model for E-discovery Search*, Oct. 4, 2009, *available at* http://e-discoveryteam.com/2009/10/04/childs-game-of-go-fish-is-a-poor-model-for-e-discovery-search/; *see also Spieker v. Quest Cherokee,* 2008 WL 4758604 (D. Kan. Oct. 30, 2008) ("Since the documents were created, stored, and/or maintained [by defendant], defendant is in the better position to develop the most appropriate list of search terms capable of producing the requested documents," and suggesting that defendants should modify plaintiff's proposed search terms if not specific enough).

18 George L. Paul & Jason R. Baron, *Information Inflation: Can the Legal System Adapt?* 13 RICH. J.L. & TECH. 10 (2007), *available at* http://law.richmond.edu/jolt/v13i3/article10.pdf.

engaging in multiple further "iterations" with opposing counsel to refine terms diminishes over time, and must be judged against available resources and discovery deadlines.[19]

4. Use of Sampling

The producing party should anticipate that regardless of what search protocol or method the parties agree upon, there will also be an expectation that the producing party has engaged in some degree of sampling to assure accurate and complete results. (The extent of the effort undertaken to sample should be proportional to the case's overall complexity, taking into account the degree of risk and what is at stake.) The *Sedona Achieving Quality Commentary* describes in greater detail how to conduct an automated search process utilizing statistical sampling:

> Once the responsive data set has been [arrived at or] characterized, a random sample of categorized material is chosen and reviewers will review this small, but statistically significant sample. This random sample will contain both responsive and unresponsive material, and reviewers classify these documents as they normally would under a manual review. The results of this classification are then compared to the results reached by the chosen categorization method(s). When there is a difference between the determination made by the human reviewer and the categorization method, the legal team reviews the document and decides which is correct. Adjustments are then made to the search strategy. Sometimes the differences require modifications so that a particular type of document is filtered in the future.[20]

5. Documentation and Defensibility

At every stage of the negotiation process, both the responding and requesting party should be well prepared in advance to explain the basis of the search strategies being proposed and the rationale as to why they are believed to yield or to have yielded quality results. At a basic level, counsel and/or any witness to be put forward must know the scope of the search conducted in terms of whose files and which files were searched, what people were told about how to go about conducting the search, and how the search was supervised. Additionally, rigorous documentation of the process is key to defensibility and to defining consistent repeatable additional searches, regardless of whether such documentation is shared with an opposing party. As noted in Practice Point 7 of the *Sedona Search Commentary*, "[t]his explanation may best come from a technical "IT" expert, a statistician, or an expert in search and retrieval technology. Counsel must be prepared to answer questions, and indeed, to prove the reasonableness and good faith of their methods." [21]

19 Feng C. Zhao, Douglas W. Oard & Jason R. Baron, *Improving Search Effectiveness in the Legal Discovery Process Using Relevance Feedback,* DESI III Conference, Barcelona, June 8, 2009, *available at* http://www.law.pitt.edu/DESI3_Workshop/DESI_III_papers.htm; *see generally Ameriwood Industries, Inc. v. Liberman,* 2007 WL 685623 (E.D. Mo. Feb. 23, 2007) (court orders expert report with number of "hits" based on negotiated search terms, with expectation that parties will continue to meet and confer to refine search based on false positives); *ClearOne Communications, Inc. v. Chiang*, 2008 WL 920336 (D. Utah Apr. 1, 2008) (court adjudicates dispute over conjunctive versus disjunctive operators between search terms, urging parties that further refinement of terms is possible subject to additional negotiations).

20 8 SEDONA CONF. J. 189 (2007).

21 *See generally, Victor Stanley,* 250 F.R.D. 251.

AN OVERVIEW OF ESI STORAGE & RETRIEVAL

John H. Jessen
Medina, WA

When considering the discovery of electronically stored information, or ESI, it is often useful to consider the various types of ESI that can be created, the physical ways in which such ESI can be stored, and the typical organizational schemas under which it can be organized. Based upon the needs of the matter at hand, these various metrics can be used to target that ESI most likely to provide useful information.

When one considers all of the ESI that a given organization may have in its possession, or its enterprise data set, the types and quantities can be staggering. In the three year period from 2004 to 2007, the average amount of data in a Fortune 1000 corporation grew from 190 terabytes to one thousand terabytes (one petabyte[1]). Over the same time period, the average data sets at 9,000 American, midsize companies grew from two terabytes to 100 terabytes.[2] Overall, the global data set grew from five exabytes (five billion gigabytes) in 2003 to 161 exabytes in 2006. It is estimated that in 2007 the amount of information created and replicated globally surpassed 255 exabytes.[3]

To place these numbers in some perspective, the Library of Congress, with 130 million items on approximately 530 miles of bookshelves—including 29 million books, 2.7 million recordings, 12 million photographs, 4.8 million maps and 58 million manuscripts—can be stored on ten terabytes. Accordingly, the entire collection of the Library of Congress could be stored more than ten times over in an average midsize company.

Given the amount of ESI that exists within the average organization, the ability to quickly and efficiently identify, locate, retrieve, and preserve the targeted set of ESI most likely to be responsive to the matter at hand becomes essential. Understanding what types of data are likely to play a role in the discovery, the possible storage locations of such data, and the likely ways in which the targeted data may be organized are all important factors in designing a discovery effort that will be focused and productive.

Types of ESI

There are potentially hundreds, if not thousands, of different types of data that could exist within an enterprise data set. Typically, however, most organizations have a limited set of potential data types as they have standardized on a limited number of applications that create ESI. Even a limited set of enterprise data types can run into the hundreds, however. This is why an important component of an ESI discovery plan is to identify the potential types of data that may play a role in providing responsive data.

---

1 A terabyte is 1,000 gigabytes, and a gigabyte is 1,000 megabytes. A petabyte is 1,000 terabytes, or 1,000,000 gigabytes.

2 From a 12-week study performed by research firm TheInfoPro, Inc., in New York in 2007. The study asked 400 data storage professionals how much storage capacity they have, how the capacity is allocated, and the capacity's effective utilization rate, or how much of its capacity is actually housing data.

3 Gantz, John F., *The Expanding Digital Universe—A Forecast of Worldwide Information Growth Through 2010*, *in* IDC WHITE PAPER, (Mar. 2007).

Interviews with key players and with the organization's computer staff are two ways to determine which data types are worth focusing on.

From the perspective of creating a discovery plan, there are two fundamental categories, or types, of ESI: (1) data created by individual custodians using local applications; and, (2) data created by individual custodians using an enterprise application and/or which is automatically created and/or captured by an enterprise application.

Custodian-Based ESI

Custodian-based ESI is familiar to anyone who uses a computer, as it is the data that is created when using application programs on personal computers or through the use of personal digital devices such as cell phones and personal digital assistants.

The key to custodian-based ESI from the discovery perspective is not necessarily what the application is or where the application is based, but rather that the custodian himself or herself controls the creation, content, storage, and disposition of the data file created.

The following are examples of common ESI created by individual custodians.

Application Data

An application program, often referred to simply as an application, is any program that is designed to perform a specific function directly for a custodian or, in some cases, for another application program. For the purposes of discovery, key features of application programs are that they are initiated by the custodian, the custodian creates the content (data) by directly interacting with application (whether personal-computer-based, a network application, or even an Internet-based application,) and the custodian determines where the resulting application data file (ESI) is going to be stored, what it is going to be named, its usage, and how long it will remain in existence.

Examples of application programs include:

- Word processors
- Spreadsheet programs
- Database programs
- Web browsers
- Software development tools
- Graphical presentation programs
- Document publishing programs
- Sales and personal contact management programs
- Document scanning and storage programs
- Voice-to-text conversion programs
- Printed-text-to-digital-text conversion programs
- Draw, paint, and image editing programs
- Financial management programs
- Music management programs
- Text and other instant messaging-type communication programs

Personal Digital Devices

A personal digital device is an electronic device operated by a custodian that is capable of creating ESI. These devices can be very specific in the types of data they hold, such as a photograph in a camera, or multi-purpose in the sense that they can hold specific types of ESI and act as a storage device for non-device-specific types of data. For example, an iPod is fundamentally a hard drive that has a music-playing application program (iTunes) on it. It can hold digitized music that is used by the application and/or it can be used to hold virtually any other type of data file.

Examples of common personal digital devices include:

- Cell phones
- Blackberry
- PDA's (Personal digital assistant)
- Cameras
- iPods or other similar device

Messaging Systems

Messaging systems are a special form of application in that they share characteristics of both custodian-based applications and enterprise applications. Most messaging systems, especially those within organizations, are enterprise-wide and enterprise-hosted applications, meaning that the messaging program itself is maintained in a central location and is available for use by all those with an authorized account. Furthermore, the messaging system typically stores some custodian-specific messaging data at this central location. Like a custodian-based application, however, most messaging systems also allow the individual custodian to maintain some portion of his or her messaging data locally on their personal computer or at some other location they may designate.

When targeting ESI in a messaging system during discovery, one must consider both the enterprise and the individual nature of the system and inquire accordingly. Inquiry must be made to both the enterprise staff charged with the housing and operation of the messaging system and the individual custodian using the system in order to determine the true nature of location, quantity, scope, and usage characteristics of the ESI sought.

Examples of common messaging systems include:

- Electronic Mail
    - Messages
    - Calendar entries
- Voice Mail
- Instant Messaging

Enterprise-based ESI

Enterprise-based ESI is data that has been created by individual custodians using an enterprise application and/or which has been automatically created and/or captured by an enterprise application. An enterprise application is typically a system where the application and its associated data reside in a central location within the organization. The application is generally one that is used by many custodians across business units within the organization, all of whom need access to all or part of the application data set.

For purposes of discovery, the key aspect of an enterprise application is that the custodian using the application does not have control over the application, its general interface, or where or how the associated data is stored or managed. Accordingly, when considering enterprise data, it is important to involve the organization's computer management staff responsible for the operation of the targeted enterprise application.

Common examples of enterprise applications include:

Organization-Specific Applications

Most large organizations have teams of software developers that write special-purpose, company-specific application programs designed to automate part of the company's business function. For example, an agricultural products company may develop an application designed specifically for tracking their crops. These applications are typically enterprise in nature and are managed by the company's information technology (IT) department.

For purposes of discovery, organization-specific applications often require more effort to identify, locate, assess, and review. Because these applications are unique, it is difficult to find information about them or their corresponding data sets in the vendor market. Focused effort must be made to identify the existence of these applications and the identity of those individuals who have knowledge about them.

Databases

Most organizations utilize database applications to organize their products and business workflow. Databases often serve as the "back-room" for other application programs, holding the information that is created in an organized fashion. Enterprise databases tend to be central stores of large volumes of structured data relating to a particular business activity or business function (i.e. product inventory.) As with organization-specific applications, databases require diligence to determine their existence, their structure, and their content.

Generic Enterprise Applications

In addition to customized organization-specific applications, many organizations employ standardized enterprise applications that have been designed and built by third-parties to solve a particular business need. Because these applications are generally available in the marketplace, it is relatively easy to find information about the application and about the data files that the application supports.

Common examples of generic enterprise applications include:

- Accounting

Automated accounting is the grandfather of all enterprise applications. Automated accounting systems record and process the accounting transactions of an organization. Most automated accounting systems are modular in nature, allowing the organization to choose those modules that it needs at the time, while allowing it to add additional functionality as needed.

Typical accounting modules include:

  - Accounts Receivable
  - Accounts Payable

  - General Ledger
  - Billing
  - Expense Entry
  - Purchase Order Management
  - Sales Order Management
  - Payroll—where the company tracks salary, wages, and related taxes
  - Employee Timesheet Management
  - Inventory Management
  - Reporting

- CRM — Customer relationship management
  CRM is a term applied to the systems and processes implemented by a company to facilitate their contact with their customers. CRM software is used to support these systems and processes, typically by storing information on current and past customers, prospective customers, and often sales leads. The information in the CRM application is typically accessed by employees in departments such as sales, marketing, product development, and customer service.

- EDRM — Electronic Document and Records Management
  The purpose of an EDRM system is to enable an organization to manage their documents throughout the document life cycle, from creation to destruction. EDRM applications typically follow a document from its inception as a work-in-progress until it has passed through a series of defined steps to become a formal record within the organization. EDRM applications are often used to associate a retention code with each record, thereby enabling the organization to destroy records once they have reached the end of their economic, regulatory, legal, or otherwise defined life cycle.

- ERP — Enterprise Resource Planning
  An ERP system is an organizational support system based on a common database that integrates the data needed for a variety of business functions such as Manufacturing, Supply Chain Management, Accounting, Human Resources, and Customer Relationship Management. Most ERP systems are modular in nature, allowing the organization to choose those modules that it needs at the time, while allowing it to add additional functionality as needed. The ultimate goal of the ERM system is to integrate all of the data in the organization into a single database that can then be used to optimize business workflows.

- PLM — Product Lifecycle Management
  A Product Lifecycle Management system provides an organization an automated platform to manage the entire lifecycle of a product, from its conception, through design and manufacture, to service and disposal. It provides the organization with a single source of all product-related information necessary for collaborating with business partners, for supporting product lines, and for developing new or enhanced product lines.

- SCM — Supply Chain Management
  A Supply Chain Management system provides an organization with an automated platform to plan, implement, and control all aspects of their supply chain by tracking the movement and storage of raw materials, work-in-process inventory, and finished goods from start to completion. A comprehensive SCM system encompasses all aspects of sourcing, procurement, logistics, and collaboration with channel partners, such as suppliers, intermediaries, third-party service providers, and customers.

- SDLC-Systems Development Life Cycle
  A Systems Development Life Cycle system provides an organization an automated platform to manage the models and methodologies that the organization uses to develop systems, generally computer systems. Most SDLC systems are modular in nature, allowing the organization to choose those modules that it needs at the time, while allowing it to add additional functionality as needed.

Typical SDLC modules include:

- ◦ Feasibility Planning
- ◦ Project planning
- ◦ Requirements Gathering
- ◦ Systems Analysis
- ◦ Systems design
- ◦ Build
- ◦ Testing
- ◦ Installation
- ◦ Deployment
- ◦ Maintenance
- ◦ Update

- SRM — Supplier Relationship Management
  A Supplier Relationship Management system provides an organization with an automated platform for managing their organizational buying processes, including the purchase of in-house supplies, raw materials for manufacturing, and goods for inventory. With the goal of reducing costs and ensuring that the organization has the materials it needs, a comprehensive SRM system measures and manages supplier performance, defines and enforces purchasing requirements, and coordinates the purchasing process with the real-time needs of the organization.

Internet

The Internet is a worldwide, publicly accessible series of interconnected computer networks that transmit data using a defined standard Internet Protocol. Functionally, the Internet is a "network of networks" comprised of millions of smaller academic, business, and government networks, which together carry information and services, such as electronic mail, text messaging, file transfer, and the Web pages and other resources of the World Wide Web.

From a discovery perspective, the information presented by an organization's Web pages, and the information gathered from visitors to those Web pages, comprises a set of ESI that can be investigated. Increasingly, organizations are connecting their Internet access points to databases and other application systems in an attempt to provide a low cost, single point of access to customers and prospective customers.

Intranet

An intranet is a private computer network established by an organization that uses Internet protocols and network connectivity to create a private, in-house version of the Internet. Intranets are typically used to provide a secure forum for the organization to share information with its employees. Utilizing a familiar Web browser interface, employees can access employee manuals, corporate calendars, updates on corporate events and milestones, records management policies, employee blogs, sales and marketing materials, stock quotations, and the like. Increasingly, intranets are being tied into corporate applications, legacy systems, and databases in an attempt to provide a single-source interface to the company.

Extranet

An extranet is a private network established by an organization that uses Internet protocols, and network connectivity to create a private, in-house version of the Internet that is then shared with selected extra-organizational parties, such as vendors, suppliers, clients, and business partners. Utilizing a familiar Web browser interface, those granted access to the organization's extranet can gain access to sales materials, catalogs, production updates, account information, electronic mail, instant messaging, blogs, and the like. Increasingly, extranets are being used to create virtual business communities where business partners come together to share information.

How ESI is Stored from a Technology Point-of-View

On-line Storage of ESI

When ESI is stored on-line, it means that the information is available to a user, on a computer system, in virtually real-time. The definition of on-line as established by the United States General Services Administration calls for an on-line system to be available for immediate use on demand without human intervention, in operation, functional and ready for service.[4]

ESI stored on-line is the most familiar form of data to users of computer systems. When a computer user sits at his or her computer or workstation, creates a data file using an application program, and then stores that file on the computer or on the corporate network, he or she has created ESI stored on-line. When a computer user sits at her or her computer or workstation and retrieves a file from the local hard drive or from a networked drive, he or she is retrieving ESI stored on-line.

On-line storage devices are primarily hard drives, whether singly in a personal computer or connected together in an array in a networked system. Hard drives allow fast access to data without any form of human intervention. As on-line storage provides the fasted retrieval time for ESI, it is typically used for those files that need to be immediately available at all times, which includes virtually all enterprise applications. Given the relatively low cost of on-line storage, most custodians choose to store their personal data files on-line as well.

From a discovery perspective, on-line data is relatively easy to identify, locate, search, retrieve and preserve, as electronic search and organization tools can be used in a real-time fashion to interrogate the data. Unlike near-line and off-line data, on-line data does not need to be "restored" before it can be utilized for discovery, thereby making on-line data a much cheaper, faster, and easier form of discovery data.

Near-Line Storage of ESI

Near-line storage is the storage of data on direct access removable media. When a near-line storage device is re-attached to a computer system, the ESI stored thereon

4 See Federal Standard 1037C, entitled Telecommunications: Glossary of Telecommunication Terms, which is a United States Federal Standard, issued by the General Services Administration pursuant to the Federal Property and Administrative Services Act of 1949, at http://www.its.bldrdoc.gov/fs-1037/fs-1037c.htm.

becomes available to the user in an on-line fashion. Near-line storage provides inexpensive, reliable, and virtually unlimited data storage, but with less accessibility than with on-line storage, as it requires the step of reintegrating the storage device with the computer system.

Near-line storage is often used for the portability of, and/or to make a backup copy of, ESI. Near-line storage is a convenient way to store ESI that is used periodically, such as music on a CD disk, or to transport ESI from one location to another.

The major categories of near-line storage include:

- Magnetic disks
    - 3.5-inch diskettes
    - Iomega Zip disk and Syquest-type removable disks
- Compact disks (CD)
    - CD recordable disks (CD-R)
    - CD rewriteable disks (CD-RW)
    - Digital versatile disk rewriteable disks (DVD-RW)
- Solid state storage (flash memory data storage device)
    - Memory card
    - Memory stick (USB flash drive)
- Removable DASD (Direct Access Storage Device) (Hard Drive) Devices
    - iPods
    - Portable hard drives

Other devices that can serve as near-line storage devices include:

- Remote on-line Backups
- Disk-based backups
- Printers with storage capability
- Fax Machines with storage capability
- Copy Machines with storage capability

From a discovery point of view, the portability of near-line storage can create identification and location problems. Additionally, while retrieval of ESI from a given near-line source is rarely an issue, retrieval from numerous near-line sources can create logistical and expense issues associated with the requirement for re-integrating the near-line storage device with the computer system before retrieval can be conducted.

Off-line Storage of ESI

As opposed to on-line storage, off-line storage is the storage of ESI on a medium or a device that is not under the control of a processing unit and which is not available for immediate use on demand by the system or custodian without human intervention.

Compared with on-line and near-line storage, sending data to off-line storage is slow. The advantage of off-line storage is that it is relatively inexpensive, easily transported, and protects the data from alteration and/or infection from computer viruses. Because of the benefits provided by off-line storage, it is often integral to an organization's backup, or disaster recovery, program.

The primary form of off-line storage is magnetic tape. So much so, in fact, that the term magnetic tape is virtually synonymous with off-line storage. When used as a backup medium, on-line ESI is written to (stored on) a magnetic tape. The recorded magnetic tape is typically then taken off-site from the organization and stored in a secure and environmentally controlled environment to protect it from natural disaster. If all or part of the ESI recorded on the magnetic tape is lost or damaged on the on-line system, the magnetic tape can be used to return a copy of the ESI to the on-line system. The time and cost associated with restoring ESI from a magnetic tape is substantial compared with the cost of on-line or near-line access, and backup tapes are therefore used as a last resort.

From a discovery perspective, magnetic tapes are a difficult and expensive environment in which to search for ESI. They must be retrieved, mounted and restored to the on-line system before any of the ESI contained thereon can be assessed.

Given that magnetic tapes are used for backup, however, means that the magnetic tape may be the only location that particular exists if it has been removed from all other on-line and near-line sources.

Backups

To fully understand off-line storage, one must understand the concept of a backup. In the computer environment, a backup refers to the making copies of data so that these additional copies may be used to restore the original after a data loss event.

Organizations typically make backups for three reasons:

First, a backup protects the organization from losing its valuable data in case of a disaster (natural or manmade) or in case of a computer system failure that results in data loss.

Second, a backup can be used to restore specific data files that have been accidentally deleted, modified, or corrupted.

Third, many organizations use backups as a generic form of long-term data archiving. In this capacity, backups are made and are held by the organization as a central repository of data over time.

Typical Categories of Backups

While a backup is technically any process that moves a file from its on-line storage location to another on-line, near-line or off-line storage location, there are some typical ways in backups are conducted by custodians and within organizations. In terms of discovery, it is important to understand the various ways in which both the client and the adversary conduct and organize their backup systems. This involves discussions with both individual custodians to determine how they may backup data as individuals, as well as with organization computer staff to determine how organizational backups are conducted.

Categorically, two different general types of backups exist, unstructured and structured:

Unstructured Backups

An unstructured backup is typically an ad-hoc copying of a small number of custodian-selected files to some form of on-line, near-line, or off-line repository. Unstructured backups are typically placed onto near-line stores like CD-R, DVD-R, or USB thumb-drive-type media.

Unstructured backups typically have little or no information about what was backed up or when the backup took place, and there is typically little consistency to the frequency and/or content of such backups. Unstructured backups are probably the easiest to implement by the custodian, but they are the least managed and are prone to dispersal and loss.

From a discovery perspective, unstructured backups are usually very difficult to deal with, as they require in-depth inquiry to identify, locate, and retrieve and, once retrieved, are costly to integrate into the discovery process due the resources required to identify the way in which the backup took place, the types and quantities of data, and the relative inefficiencies associated with loading a relatively small amount of data.

Structured Backups

A structured backup is a backup of a predictable target set of data that occurs on a predictable timetable. Structured backups are the types of backups that occur most frequently within organizations and they account for the vast majority of data stored within backups.

Structured backups, and especially those conducted systematically by an organization's computer services department, typically have detailed descriptions about what was backed up, when it was backed up, and how it was backed up. From a discovery perspective, structured backups are generally easier to identify, locate, and retrieve, and a greater level of analysis can generally be conducted as to the types and quantities of data contained thereon.

Local Backup

Local backups are typically backups of data files conducted by custodians through the use of devices contained within, or attached directly to, their personal computer workstation. From a discovery perspective, local backups are usually sporadic in nature, stored in various locations, inconsistent in terms of types and quantities of data stored, and difficult to restore.

Typical local backup schemas include:

- Backing up data files to magnetic disks such as floppy diskettes, Iomega Zip disks, or Syquest-type removable disks
- Backing up data files to compact disks (CD's) such as CD recordable disks (CD-R), CD rewriteable disks (CD-RW), or Digital versatile disk rewriteable disks (DVD-RW)
- Backing up data files to solid state storage (flash memory data storage devices) such as memory cards or memory sticks (USB flash drives)
- Backing up data files to removable DASD (Direct Access Storage Device) (Hard Drive) devices such as iPods or other portable hard drives

Internet Backup

As high-speed Internet service has become more widely available and more robust, backup methodologies utilizing the Internet to create remote backup stores is growing in popularity. These remote sites can simply be other personal or organizational sites that the custodian has access to, or they can be sites provided by third-party companies providing backup and storage services.

As remote Internet backup sites are organizationally and, typically, geographically removed, backing data up to the Internet can provide protection against geographically clustered disasters that could affect backup data stored in the same region as the host data. Even with high-speed Internet capability, Internet backups are substantially slower than backups conducted to local disk storage or to backup tape. This speed issue generally limits the amount of data that a custodian would choose to send to a remote Internet site. Some organizations also feel uncomfortable placing their data into the hands of third-parties to hold and manage, fearing that sensitive data may be compromised.

From a discovery perspective, it is important to understand that the custodian typically determines the frequency of, and the composition of, the backup set that is sent over the Internet. Care must be taken to fully understand the extent to which a given custodian uses Internet backup, the frequency of such backups, the manner in which data is selected for backup, and the details of the remote site at which the data is stored.

Enterprise Backup

Perhaps the most common form of backup in a corporation or other organizational entity is the enterprise backup. An enterprise backup is a backup conducted by an organization's computer services staff involving business unit-level or organization-wide computer systems. A backup of an organization's electronic mail system on a daily basis would be an example of an enterprise backup.

Because they are conducted by the organization's computer services staff for the purpose of providing a disaster recovery copy of the organization's data, enterprise backups tend to be the most structured in terms of the scope of the data targeted, the frequency of the backup, the consistency of the media onto which the backup is made, the recoverability of the backed up data, and the length of time the backup is maintained before disposal.

From a discovery perspective, enterprise backups are often the easiest to identify, locate, and retrieve, although the volumes of backup sets that often exist within an organization can make the logistics of the discovery very difficult. It is also important to keep in mind that magnetic tapes can fail, thereby compromising the entire backup set to which that tape belonged. There may also be difficulties associated with interpreting the many types of data files that are often co-mingled on enterprise backups.

Typical Types of Backup Schemas

Within categories of backups there are different backup schemas that can be employed. Understanding the schema chosen for a given backup is an important component in developing a proper model for restoring a backed up set of data, especially if one is restoring multiple backups to determine a set of data from a targeted time period.

Typical backup schemas include:

Full Backup

A full backup is a backup of every file on the targeted computer system, whether or not that file has changed since the previous backup.

Because a full backup copies every file on the targeted system to the backup media, a full backup takes the longest to accomplish of all the backup schemas and requires the most storage space on the backup media. In terms of restoration, however, a full backup provides the fastest restoration times when restoring the full data set.

Because of the time and tape space required, full backups are generally conducted on a periodic basis as part of a hybrid backup schema. For example, a full backup may be conducted every Sunday night, while an incremental backup is conducted on the days in between. Full backups are also typically performed on systems that are about to undergo hardware and/or software changes as a means to protect against data loss in case the changes do not work or damage the file storage systems.

If an organization chooses to save selected backups over a long period of time as a means of creating an ad-hoc data archive, full backups are usually the ones chosen. A typical example would be to save the last full backup of every month and to save that backup for one or more years.

Incremental Backup

An incremental backup is a backup of every file on the targeted computer system that has changed since the last backup took place, regardless of whether the last backup was a full backup or an incremental backup.

Because an incremental backup only targets those files that have changes since the last backup, which is typically a fraction of the total data set, it is typically the fastest type of backup and the one that requires the least storage space on the backup media. However, incremental backups also require the longest time and the most tapes to restore. When restoring a full system, however, an incremental backup schema may take the longest time to restore as the first incremental backup has to first be restored and then all of the subsequent incremental backups leading up to the targeted restoration date.

Because of the inefficiencies associated with restoring an incremental-only backup schema, one rarely sees an incremental-only backup schema in place. In most organizations, an incremental backup schema is used in conjunction with a full backup.

Differential Backup

A differential backup is a backup of every file on the targeted computer system that has changed since the last full backup.

While a differential backup is not as fast as an incremental backup, it is faster than a full backup as it does not have to copy every file. Correspondingly, a differential backup requires more storage space than an incremental backup, but less than a full backup.

When used in combination with a full backup, differential backups can provide an effective and efficient backup process. As with incremental backups, in discovery one must fully investigate the way in which backup schemas are utilized, in whole or in combination, to determine the appropriate restoration model.

Continuous Data Backup

A continuous backup is a real-time backup that immediately logs every change on the targeted computer system to a secondary system. This is often done by saving byte or block-level differences rather than file-level differences, which allows the real-time nature of the system to take place. Effectively, pieces of files are saved as they are changed. If a restoration needs to take place, the management system knows how to piece everything back together in proper form.

With a continuing decrease in hard disk storage costs, continuous backup, sometimes referred to as mirroring, may become more popular in the future.

Examples of the differing backup schemas

- ***Full Backup***
  If you perform a full backup every day of the week and the system crashes on Friday, you would need to restore the full backup set from Thursday to restore the data.

- ***Full plus Incremental Backup***
  If you perform a full backup each Sunday and incremental backups every night and the system crashes on Friday, you would need to restore the full backup from Sunday along with the incremental backups from Monday, Tuesday, Wednesday, and Thursday to restore the data.

- ***Full plus Differential Backup***
  If you perform a full backup each Sunday and differential backups every night and the system crashes on Friday, you would need to restore the full backup from Sunday and the differential backup from Thursday.

- ***Continuous Backup***
  If the system crashes on Friday, you simply restore the files from the secondary source.

- ***Backup Rotation***
  A backup rotation schema is the method chosen for managing backup sets when multiple media are used in the backup process. The rotation schema determines how and when each magnetic tape is used in a backup and for how long it is retained once it has backup data stored on it.

  The most common backup rotation schema is referred to as the Grandfather-Father-Son model. The Grandfather-Father-Son model defines three sets of backups—daily, weekly and monthly. The daily (Son) backups are rotated on a daily basis with one set graduating to Weekly (Father) status each week. The weekly backups are rotated on a weekly basis with one graduating to Monthly (Grandfather) status each month. Many organizations add to this model by removing one or more monthly tapes to an annual or multi-year storage.

  Another common rotation schema is to use a rolling set of magnetic tapes over and over again. This Incremental model defines a pool of backup media and, once the entire pool has been used, re-writes to the oldest set. For example, with a daily backup onto a set of 10 tape sets, you would have 10 days worth of individual daily backups. When all of the tape sets are used, the oldest one is inserted and re-used.

  Tape rotation schemas can get very complicated based upon the needs of the organization. In terms of discovery, it is important to determine what tape rotation model is used and how it is implemented. With any rotation model there will be gaps in the tape sets due to human, machine, or tape failures. There will likely also be extra-model tape sets in existence that have been created ad-hoc or for special purposes.

How ESI is Stored from a Custodian/Records Management Point-of-View

From a technology standpoint, ESI can be stored on a variety of magnetic, optical, and solid-state media. The manner in which ESI is stored by the custodian onto these media can vary greatly, however, and has to do with both the organization's records management plan and the custodian's own desires regarding the naming and storage location of his or her data.

When considering what ESI may relate to a given discovery matter, it is often useful to consider where such data may have been placed by a custodian or, indeed, whether such data was ever under the direct control of the custodian.

There are five typical ways in which ESI can be stored:

Custodian-Centric Data Storage

Much of the ESI used by a custodian on a day-to-day basis, especially application data, is under the direct control of the custodian. It is the custodian who creates the content associated with a given data file, names it, and determines where the file will be saved. The custodian is also the default "records manager" for his or her data in the sense that he or she determines how long data will survive before being deleted.

In terms of discovery, the custodian is often the best source of information about his or her data set, including:

- Types of data created (i.e. what applications were used, including enterprise applications)
- Quantities of data created
- File naming conventions used
- Data storage locations
- Whether custodian-based backups were created
- Others with whom the custodian corresponded and/or shared files
- Use of electronic mail and attachments[5]

Virtual Workgroup-Centric Data Storage

A virtual workgroup is group of individuals who work on a common project using digital technologies such as electronic mail, instant messaging, shared application programs and databases, calendaring, and file management. Many virtual workgroups share a common data file through the use of applications that support such use.

While the custodian creates some of the content for the application data file, he or she may have little or no say in how the data file is named, where it is stored, how it is

5 Electronic mail is a unique application in the sense that it has both enterprise and local characteristics. Technically electronic mail is an enterprise software platform, but users use and often store electronic mail messages and attachments locally. Accordingly, it makes sense in discovery to investigate electronic mail from both the enterprise level through discussions with organizational computer services personnel and locally through discussions with individual custodians as to how they use their electronic mail system.

ultimately used, or how log it remains in existence. Many times these issues are handled either by organization rules or by a custodian named as the workgroup leader.

As networking and the Internet become more pervasive, and as application software providers enable workgroup features into their software, the concept of virtual workgroups is likely to grow. Rather than sending a file around to a number of individuals and then trying to integrate their suggestions and changes, the data file remains in a central location and the users modify it directly, with each persons edits and/or notations identified with each such person.

In terms of discovery, the custodian is often the best source of information about his or her participation in a virtual workgroup, including:

- Virtual workgroups assigned to
- Other participants in the workgroup(s)
- Applications used by the workgroup, including enterprise applications
- The workgroup's data storage location
- The workgroup's computer services liaison
- File naming conventions used by the workgroup
- File management conventions related to the workgroup (i.e. data backup, data retention)

Business Unit-Centric Data Storage

Many organizations are structured like holding companies, made up of many stand-alone organizations (business units) that maintain their own computer operations but that share some overall application platforms, such as electronic mail. A single organization may also have different operating divisions that it treats as business units.

A custodian working in one business unit within a larger organization may spend most of their time working on the business unit's computer system, but at least part of their time on platforms owned and managed by the parent organization. From the custodians viewpoint he or she is working on a single system. Behind the scenes, however, many different operating and data storage environments may be involved.

While the custodian may create the content associated with a given data file and may name it, in some business-unit environments the custodian may have little or no control over where the data is saved. This is especially true when enterprise applications are used.

In terms of discovery, both the custodian and organizational computer services staff need to be considered as sources of information about the underlying computer system being used and the location(s) of related data stores, including:

- Custodian
  - Types of data created (i.e. what applications were used, including enterprise applications)
  - File naming conventions used
  - Data storage locations, if known

- Organizational Computer Services Staff
  - Desktop applications provided to the custodian(s)
  - Enterprise applications provided to the custodian(s)

  - Types and quantities of data created
  - Data storage locations for each application
  - Data management policies for each application (i.e. data backup, data retention)

Enterprise-Centric Data Storage

Virtually every organization utilizes enterprise applications in their business model. We have seen in a previous section the various types of enterprise applications that exist. If nothing else, electronic mail is pervasive and at its core it is an enterprise application.[6]

One of the key characteristics of an enterprise application is that the data file(s) associated with the application are stored and managed at a central location within the organization, typically by professional computer services staff. Custodians using the enterprise application may have desktop applications that belong to and/or interact with the enterprise application, or they may simply "log on" to the enterprise application and use it directly at its central location.

While the custodian may create new data using the enterprise application and/or modify existing information, the custodian typically has no say in how the data file is named, where it is stored, or how it is managed.

In terms of discovery, both the custodian and organizational computer services staff need to be considered as sources of information about the enterprise applications being used and the location(s) of related data stores, including:

- Custodian
  - Types enterprise applications used
  - The custodian's typical usage of such applications
  - Data storage locations, if known
  - Identity of the custodian's computer services liaison for each such application

- Organizational Computer Services Staff
  - Enterprise applications provided to the custodian(s)
  - Types and quantities of data created
  - Data storage locations for each application
  - Data management policies for each application (i.e. data backup, data retention)

3rd Party-Centric Data Storage

With the increased use of outsourced computer operations and the use of Internet-based applications, more and more organizational data is being stored and managed by third-parties. In an outsourced situation, a third-party manages the hardware and software infrastructure for an organization for a fee. In effect, the third-party is serving as the computer services department for the organization. An Internet-based application is one in which a user using the Internet goes to a third-party site and logs onto an application program provided by the third-party. The user then uses the application just as they would if it resided on their desktop or on the enterprise computers. In both situations, the data created by the user remains with the third-party provider.

6 See footnote 5.

In terms of discovery, the custodian, the organizational computer services staff, and the third-party's computer staff may need to be considered as sources of information about the applications being used and the location(s) of related data stores, including:

- Custodian
  - Types of third-party applications used
  - The custodian's typical usage of such applications
  - Data storage locations, if known
  - Identity of the custodian's computer services (both in-house and third-party) liaison for each such application

- In-house Organizational Computer Services Staff
  - Enterprise applications provided to the custodian(s)
  - Identity of third-party providers
  - Types and quantities of data created
  - Data storage locations for each application
  - Data management policies for each application (i.e. data backup, data retention)

- Third-party Organizational Computer Services Staff
  - Enterprise applications provided to the custodian(s)
  - Types and quantities of data created
  - Data storage locations for each application
  - Data management policies for each application (i.e. data backup, data retention)

Fundamental Computer Forensic Issues

Forensic Disk Images

When used in conjunction with discovery, the term forensics relates to the use of specialized techniques for the recovery, authentication, and analysis of specific ESI.

Forensic examinations are typically used when a matter involves issues that require the reconstruction of computer usage patterns; the examination of residual data left after deletion; technical analysis of computer usage patterns; and/or other testing of the data that may de destructive in nature.

In order for a forensic examination to occur, the ESI, and the storage device on which the ESI resides, must be collected in a manner that requires specialized expertise that typically goes beyond normal data collection and preservation techniques that are generally available to users and even system support personnel.

The most common form of forensic collection is to make an image of the storage media on which the targeted ESI resides.

This image, sometimes called a bit image, is an exact copy of the storage device—such as a hard drive, a CD, or any other disk format—including all areas that contain data and all areas that appear to be empty (but which may actually contain remnants of data.)

The image is a single file containing the complete contents and structure of the storage device. A disk image file is created by making a sector-by-sector copy of the source media, thereby completely copying the entire structure and contents of the storage media. Forensic images are acquired with the use of specialized software tools. When used properly, these images contain a copy of everything that is on the target media, including live and deleted data. Forensic images are also sometimes referred to as a bitstream image, a bit image, or a cloned image.

This image can be used to re-create an exact copy of the storage device on which a forensic examination can be conducted. This examination can then be conducted on the re-created drive in exactly the same way in which it could have been done on the original device. Because forensic examinations often involve destructive testing, and because they require the ability to replicate their findings, this ability to work on re-created drives is critical.

The primary question when considering a forensic collection is whether or not the facts surrounding the matter at hand suggest that a forensic examination is going to be needed.

Was unique, important data deleted? Is it likely that deleted data can be recovered? Is it important to show usage activity and usage patterns? Is it important to authenticate a particular file in order to show that the represented data and/or time that of creation is accurate? Do you need to confirm that all of the text in a document is original or that a critical email was really sent when it appears to have been?

If the matter is one where a forensic examination may be important, then a forensic collection is required. If not, then a forensic collection is not required and is ultimately a waste of time and resources and which often sets the stage for needless battles over additional forensic examinations that could be conducted on the collected data.

Because imaging software is commonly available, and because the vast majority of training programs in the field of electronic discovery revolve around forensics, there is a growing tendency to want to "image everything." Unless an argument can be made that the matter at hand will benefit from a forensic collection and additional examination, there is no reason to do a forensic collection just because the technology exists to do it.

If the matter allows a non-forensic acquisition and analysis of ESI, then a data collection is what is required. A data collection, as opposed to a forensic collection, collects files at the file level, not at the disk level, basically by copying the desired information and processing it into a review system. Data collection is faster and cheaper than a forensic collection and is the type of collection that is warranted if a forensic collection is not required.

File Deletion

A computer's file system determines how the computer stores and manages files on its attached storage media. There are several file systems in use today, and all offer some form of file recovery once a file is deleted.

For illustration purposes, we will discuss the FAT file system, one of the most commonly used file systems today. When a file is deleted on a File Allocation Table (FAT) file system, its directory entry[7] in the FAT remains stored on the disk, although the file name is altered in a way that lets the system know that the storage space occupied by the

7 A file's directory entry is much like a person's listing in a telephone book. It holds the file's name and it's storage location on a piece of storage media, such as a hard drive, a CD, or a DVD. The directory entry basically tells the computer where to go to find the data file.

(now deleted) file is again available for use by a new file or by an expanded version of an existing file. The majority of the deleted file's information, such as it's name, time stamp, file length and location on the disk, remain unchanged in its directory entry in the FAT.

The deleted file's content will remain on the storage media until it is overwritten by another file. The more file activity there is on a particular computer system, the more unlikely it is that a file can be recovered, as the likelihood that the storage areas where the file had resided will be overwritten is greater.

Specialized software utilities, some provided with, or built into, the operating system, allow for the recovery of a deleted file provided that a new file or data set has not overwritten the areas of the storage device holding the deleted file in question. At the simplest level, these tools allow the modified file name of the deleted file to be changed back into a name format that does not indicate a deleted file. The file then becomes a "live" file again, and available for use by an application program. In some cases a greater level of reconstruction is required to retrieve some or all of a deleted file. If the directory entry for the deleted file has been overwritten, or if some of the data storage areas for the deleted file have been overwritten, it will be more difficult to perform the file recovery.

Some computer operating systems provide a layered approach to data deletion. Microsoft's Windows platform, for example, does not really delete a file when a normal deletion request is made. The file is places in a "recycle bin" where it awaits final deletion. Until the file is removed form the recycle bin, it can be easily recovered as it had not really been deleted in a technical sense. When the file is "dumped" from the recycle bin for deletion, it can often be recovered along the lines described above in the previous section.

As with forensic collection, the key question in discovery regarding the recovery of deleted data is whether or not the facts surrounding the matter at hand suggest that a data recovery is going to be needed.

Was unique, important data deleted? Is it likely that deleted data can be recovered? Was the file located on a system where file activity was such that recovery is likely to be effected? Is the matter-at-hand one where file deletion is suspected or traditionally part of the pattern of activity for such matters, such as in trade secret theft?

If the matter is one where deleted data recovery may be important, then attempts should be made to identify and recover appropriate files. If not, then deleted data recovery is not warranted and is ultimately a waste of time and resources.

As with imaging, data recovery software is commonly available, and because the many of the training programs in the field of electronic discovery revolve around forensics (which is often targeted towards data recovery), there is often a bias to target deleted data. Unless an argument can be made that the matter at hand will benefit from the recovery of deleted data, there is no reason to attempt such recovery just because the technology exists to do it.

Metadata

Generally, metadata is information about a particular data set which describes "how, when, and by whom it was collected, created, accessed or modified, and how it was formatted."[8]

8 THE SEDONA CONFERENCE®, THE SEDONA CONFERENCE® GLOSSARY: EDISCOVERY & DIGITAL INFORAMTION MANAGEMENT 33 (2d ed. 2007).

Metadata provides context for data files and is used to facilitate the understanding, characteristics, and management usage of such data. The metadata required for effective data management varies with the type of data in use and the potential use of such data. In a DVD video collection, for example, where the data is the content of the videos, metadata about a given video title would typically include a description of the content, the title, the producer, the director, the actors, the release data, and the physical location of the video. From this example, one can see how the metadata helps identify and organize the data content and make the collection more useful

In terms of discovery, one should consider various categories of metadata and determine which of these categories may play a role in the matter at hand. In addition, one should always consider metadata from the perspective of a given data file, and whether or not such metadata will provide useful information regarding that file. The common categories of metadata include systems, application, and embedded metadata.

E-DISCOVERY IN CRIMINAL MATTERS – EMERGING TRENDS & THE INFLUENCE OF CIVIL LITIGATION PRINCIPLES

Post-Indictment E-Discovery Jurisprudence

Justin P. Murphy
Crowell & Moring LLP
Washington, DC

Although considerable attention has been paid to e-discovery issues and decisions in civil litigation, little focus has been presented to similar issues arising in the criminal post-indictment context. This article will address the influence of established civil litigation principles on criminal matters, how courts manage what can often be government intransigence toward defendants' discovery rights and requests related to Electronically Stored Information ("ESI"), and suggests civil litigation principles that may eventually impact criminal defendants.

There are relatively few court decisions relating to post-indictment discovery of ESI. Because of that small universe, the influence of civil litigation principles on these decisions is magnified. From these decisions, several guiding principles emerge: First and most generally, the government should comply with certain civil procedural principles governing e-discovery, particularly Rule 34, as well as court rulings enforcing those rules. Second, government "data dumps" may be impermissible, absent producing the materials in a searchable, indexed and reasonably organized format. Third, criminal defendants may insist on expansive ESI discovery based on their own theories of the case; the government may not limit discovery based on its own assumptions of what evidence may or may not be relevant to the case.

Importantly, although trends and guiding principles can be drawn from court decisions, the decisions themselves are very fact-specific, and different courts have decided similar e-discovery issues in very different ways. Some courts take a permissive approach and allow defendants expansive discovery, or punish the government for non-compliance. Other courts take a more pragmatic approach and seek to strike a balance between the needs of the parties. Finally, some courts are restrictive and limit defendants' access to electronic evidence. Examples of these approaches are in the pages that follow.

The *O'Keefe* Framework – The Government Must Comply with Rule 34 of the Federal Rules of Civil Procedure

United States v. O'Keefe[1] represents the first systematic discussion of how civil discovery rules and principles – specifically those related to ESI – may apply to criminal matters. Given that the discovery provisions of the Federal Rules of Criminal Procedure do

1 *United States v. O'Keefe*, 537 F. Supp. 2d 14 (D.D.C. 2008).

not specifically address e-discovery, and given the paucity of post-indictment e-discovery precedent in criminal matters, *O'Keefe* represents fresh thinking that civil standards can be imported into criminal matters in order to clarify defendants' discovery rights.

In *O'Keefe*, the court held that a document production by the government in a criminal matter must adhere to standards similar to those set forth in Rule 34 of the Federal Rules of Civil Procedure.[2] The defendants argued that the government produced documents such that it was impossible to identify the source or custodian of the document.[3] Noting that there was no analogous criminal rule to guide judges in determining whether a government production had been tendered in an appropriate form or format, and acknowledging that the "big paper case" would be the exception rather than the rule in criminal cases, the court observed that Rule 34 of the Federal Rules of Civil Procedure speaks directly to form of production and should be looked to for guidance.[4] "The Federal Rules of Civil Procedure in their present form are the product of nearly 70 years of use and have been consistently amended by advisory committees consisting of judges, practitioners, and distinguished academics to meet perceived deficiencies. It is foolish to disregard them merely because this is a criminal case, particularly where . . . it is far better to use these rules than to reinvent the wheel when the production of documents in criminal and civil cases raises the same problems."[5] Following this pragmatic approach, the court determined that the government's production must, at a minimum, be labeled and ordered as they were maintained in the ordinary course of business.

The *O'Keefe* court is not the only body to advance the influence of civil principles in criminal matters. The Advisory Notes of the 2009 amendments to the Federal Rules of Criminal Procedure state that its use of the term ESI is "drawn" from the Federal Rules of Civil Procedure. The Advisory Committee Notes to Rule 41(e)(2) state that the term "electronically stored information" is drawn from Rule 34(a) of the Federal Rules of Civil Procedure, which states that ESI includes "writings, drawings, graphs, charts, photographs, sound recordings, images, and other data or data compilations stored in any medium from which information can be obtained. The 2006 Committee Note to Rule 34(a) explains that the description is intended to cover all current types of computer-based information and to encompass future changes and developments. The same broad and flexible description is intended under Rule 41."[6]

O'Keefe's reliance on the civil rules in a criminal case has already been employed by other criminal defendants. For example, in *United States v. Stevens*, the defense objected that the government produced thousands of pages of documents in an unusable format that "appeared to be an undifferentiated mass, with no discernible beginning or end of any given document."[7] Citing *O'Keefe* and highlighting the unnecessary and increased burden to the defendant, the defense argued that "even civil litigants must either produce documents as they are kept in the course of business or label the documents in response to requested subject areas. Where the government produces documents in an undifferentiated mass in a large box without file folders or labels, then these documents have not been produced in the manner in which they were ordinarily maintained as [Fed. R. Civ. P. 34] requires and thus the government has equally failed to meet its obligations under Fed. R. Crim. P. 16."[8] The defense also requested metadata and logs related to the government's forensic

2 *Id.* at 18-19.
3 *Id.* at 18.
4 *Id.* at 18-19.
5 *Id.* at 19.
6 Fed. R. Crim. P. 41(e)(2), *Advisory Committee Notes, 2009 Amendments*.
7 Defendant's Motion to Compel Discovery, *United States v. Stevens*, No. 08-231, at 14 (D.D.C. Sept. 2, 2008).
8 *Id.* at 15.

photography, adding that "[t]here is nothing remarkable about asking the government to produce metadata. Courts routinely permit the discovery of metadata in the civil context . . . and there is no principled reason why it ought not be produced in a criminal case."[9] Other criminal defendants have also relied on *O'Keefe* in challenging disorganized government productions.[10]

In sum, the *O'Keefe* decision provides a vehicle for parties and courts in criminal matters – where no criminal rule applies – to allow civil rules to guide their practice relating to ESI. The guiding principles discussed below emerge from these circumstances.

Government "Data Dumps" Must Be Organized, Searchable and Indexed

E-discovery in criminal matters should not be more burdensome for criminal defendants who enjoy more enhanced constitutional protections of the Fifth and Sixth Amendments, for example, than their civil counterparts who enjoy the benefits of procedural standards established by the Federal Rules of Civil Procedure and the growing body of caselaw interpreting those rules. One such civil standard – as discussed in detail in *O'Keefe* – is Rule 34 of the Federal Rules of Civil Procedure, providing guidance relating to form of production in discovery.[11] Both civil and criminal courts have addressed issues relating to Rule 34, and, specifically, a party's obligation when it backs up the proverbial electronic dump truck to the requesting party's door.

One such example is *United States v. Skilling,*[12] where the defendant asserted that the government's several hundred-million page production violated the government's *Brady* obligations because the voluminous "open file" suppressed exculpatory evidence.[13] Skilling claimed that no amount of diligence could have successfully identified exculpatory materials within the morass of information produced by the government. The Fifth Circuit, adopting a pragmatic approach, rejected this argument, and instead focused on the reasonable steps that the government had taken in their "open file" production. Importantly, the files in the "data dump" were searchable, the government provided an index of the documents along with access to databases from other related cases, and produced a set of "hot documents" that the government believed were important to its case and/or potentially relevant to the defendant's case. Highlighting that the government was in no better position than the defendant to identify potentially exculpatory material, the court identified several factors that in its view would render such "data dumps" unacceptable: (1) open file productions padded with "pointless or superfluous" materials to frustrate a defendant's review of those materials; (2) producing files so voluminous that access is "unduly onerous;" and (3) the placement of exculpatory evidence within voluminous files in order to conceal it.[14] The Fifth Circuit, finding that

9 Defendant's Reply in Support of Motion to Compel Discovery, *United States v. Stevens*, No. 08-231, at 9 (D.D.C. Sept. 6, 2008). Ultimately, the issues were resolved without a written opinion by the court.

10 *See* Defendant's Reply to Government's Opposition to Discovery Motion, *United States v. Nozette*, No. 09-276 (D.D.C. Jan. 4, 2010) (noting that the vast bulk of discovery produced by the government lacked "any indication of its provenance or relevance" and that the government is required to produce the documents as they are maintained in the ordinary course of business or it must organize and label the materials to correspond to the discovery requests).

11 Fed. R. Civ. P. 34(b)(e) provides that, *inter alia*, "Unless otherwise stipulated or ordered by the court, these procedures apply to producing documents or electronically stored information: (i) A party must produce documents as they are kept in the usual course of business or must organize and label them to correspond to the categories in the request; (ii) If a request does not specify a form for producing electronically stored information, a party must produce it in a form or forms in which it is ordinarily maintained or in a reasonably usable form or forms. . . ."

12 *United States v. Skilling*, 554 F.3d 529 (5th Cir. 2009).

13 *Id. at* 576.

14 *Id*. at 577.

these factors were not present, determined that government had not acted in bad faith or violated its *Brady* obligations.[15]

The Fifth Circuit, in considering the additional steps that the government took beyond merely producing an "open file," settled on a rule of reasonableness in its decision in *Skilling*. But what if the government had not taken such step? Such production deficiencies, especially in civil enforcement cases where the government's discovery failures have been the subject of judicial opinions, might bring the *O'Keefe* paradigm and civil rules and cases interpreting them into play.

In *SEC v. Collins & Aikman Corp,*[16] a defendant argued that the size of the SEC's production – 1.7 million documents spread across six databases – impermissibly hampered its defense because the SEC failed to search for and identify documents relevant to the factual allegations in the complaint. The SEC argued that it did not conduct such searches in the normal course of business and was not obligated to do so for the defendant, and while acknowledging that it had internally organized the produced documents in folders corresponding to the specific allegations of the complaint, stated that its organization constituted attorney work product.

The court disagreed, stating "[i]t is patently inequitable to require a party to search ten million pages to find documents already identified by its adversary as supporting the allegations of a complaint."[17] Noting that even if the document compilation folders constituted attorney work product, the defendant had shown a substantial need for the material because of the "undue hardship" and considerable expense he would face if forced to search through the government's document dump, page by page.[18] The court also criticized the "SEC's blanket refusal to negotiate a workable search protocol responsive to [the defendant's] requests" as "patently unreasonable"[19] and ordered the parties to meet and develop a search protocol that would reveal "*at least some* of the information defendant seeks."[20] Following a pragmatic approach, the court noted that the SEC's documents, collected through its investigation, were not part of the government's "ordinary course of business" (as opposed to government contracting and other business-like activities) and held that such investigations are never within the "ordinary scope of business" under Rule 34. The court's conclusion – that "[w]hen a government agency initiates litigation, it must be prepared to follow the same discovery rules that govern private parties"[21] – may have relevance to criminal matters as well.

If one subscribes to the *O'Keefe* construct, *Collins* provides intriguing applications to criminal matters. *Collins* sets boundaries for an impermissible "data dump" and implies that a government agency cannot unilaterally attempt to restrict the scope of discovery (by rejecting proposals to establish search protocols, for example), particularly where it initiated the action. Moreover, there is no "usual course of business" where government agency investigative files are at issue – as the *Collins* court found – so such materials must be produced in a format that organizes or indexes the materials according to the subjects of the

15 An example of a court taking a slightly more restrictive approach is *United States v. Ferguson*, 478 F. Supp. 2d 220 (D. Conn. 2007). In *Ferguson*, the defendant argued that the government failed to provide sufficient guidance to navigate a 3.5 million-page database provided during discovery. The court said the government provided sufficient guidance by making the database searchable and handing over a list of the most relevant documents. The court held that those factors, combined with a sufficiently detailed indictment, was enough guidance for the defendant to prepare a defense, avoid prejudicial surprise at trial, and protect against future double jeopardy. *Id.* at 227.
16 *SEC v. Collins & Aikman Corp*., 256 F.R.D. 403 (S.D.N.Y. 2009).
17 *Id.* at 411.
18 *Id.*
19 *Id.* at 414.
20 *Id.* at 415.
21 *Id.* at 418.

discovery request. The *Collins* holding leaves one to wonder whether in a criminal case a defendant could request ESI from the government in a similarly "usable" format. In addition, if the ESI is housed and/or organized in a database, a defendant could also request some or all of the folder structure and/or coding that accompanies the data – which *Collins* determined was not attorney work product – which would be the equivalent of the investigative files in *Collins* that the SEC was ordered to turn over.

However, some courts have taken a more restrictive approach to criminal e-discovery and found, in circumstances similar to *Collins*, that a defendant cannot gain access to the government's internal search and coding information. For example, in *United States v. Schmidt*,[22] the government provided summaries of certain bank records in a tax fraud case that were part of a discovery file that both sides acknowledged was too voluminous to be examined in court. The defendant requested access to the IRS computer programs used to compile and generate a summary spreadsheet of the records. The court refused, holding that "access to the computer database program would reveal what queries [the Government] ran in order to prepare the spreadsheets. Such access . . . would clearly invade the province of the agent's work product by giving defendants insight into the agent's thought processes as he analyzed and compiled the underlying documents. The database, therefore, is not discoverable."[23]

Ultimately, government "data dumps" or failures to properly manage the production of ESI may lead to the dismissal of its case. In *United States v. Graham*, the government was slow to produce millions of documents and significant amounts of other media, productions were often incomplete or tainted with computer viruses, and the defendants had great difficulty in coping with the large volume.[24] The court, taking a more permissive approach for the defendant, dismissed the indictment for Speedy Trial Act violations but noted that discovery was at the heart of the matter: "In this case, the problem . . . is and has been discovery . . . . One, the volume of discovery in this case quite simply has been unmanageable for defense counsel. Two, like a restless volcano, the government periodically spews forth new discovery, which adds to defense counsels' already monumental due diligence responsibilities. Three, the discovery itself has often been tainted or incomplete."[25] In dismissing the case, the court noted that although the government did not act in bad faith, "discovery could have and should have been handled differently."[26]

The Government's Production of ESI Cannot be Unduly Limited By Its Own Assumptions About What is Relevant.

In civil discovery, parties may seek discovery about items both relevant and/or likely to lead to the discovery of admissible evidence. The standard for relevance is expansive,[27] in contrast with the more limited discovery afforded by Rule 16 of the Federal Rules of Criminal Procedure. However, in some instances, courts have permitted discovery requests related to ESI which resemble the expansive view of discovery followed in civil matters.

22 *United States v. Schmidt*, 04-cr-00103-REB, 2007 WL 1232180 (D. Colo. Apr. 25, 2007).
23 *Id.* at *1.
24 *United States v. Graham*, No. 1:05-CR-45, 2008 WL 2098044, at *2-3 (S.D. Ohio May 16, 2008). *See also State v. Dingman*, 202 P.3d 388 (Wash. Ct. App. 2009) (court reversed conviction and remanded for new trial after finding that trial court erred by denying defendant meaningful access to hard drives seized from his house).
25 *Graham*, at *5.
26 *Id.* at *8. *But see United States v. Qadri*, Crim. No. 06-00469 DAE, 2010 WL 933752, at *5 (D. Haw. Mar. 9, 2010) (despite considerable and lengthy delays in Government's discovery production due "in part to the nature of electronic discovery [and] the complex nature of the alleged crimes," court declined to dismiss indictment because defendant, after several years of delay, was provided with evidence seven months before trial).
27 *See, e.g.*, FED. R. CIV. P. 26(b) and 34(a) (Advisory notes to Fed. R. Civ. P. 34(a) indicate that federal rules take an "expansive approach toward discovery of ESI and that discovery of [ESI] stands on equal footing with discovery of paper documents").

In *United States v. Safavian*,[28] the defendant served discovery requests on the government seeking emails between government officials and third parties. The government objected, arguing that the emails had no bearing on the defendant's state of mind at the time he allegedly made false statements because the defendant had never seen the emails. The government also argued that the term "government" did not encompass more than the Justice Department, FBI, the GSA-OIG and other investigative agencies. The court rejected the government's arguments, noting that "simply because the e-mails themselves were not sent to or received by Mr. Safavian and therefore do not reflect on his state of mind, and may or may not be admissible in evidence at trial, does not mean they are not material to the preparation of a defense or that they will be unlikely to lead to admissible evidence."[29] Further, the court took a broad view of "government," determining that it was irrelevant whether documents were in the possession of Justice Department prosecutors or the FBI; instead, the documents were to be produced if they were in the possession, custody, or control of any agency of the executive branch of the government and either material to the defense or *Brady* material.[30]

Other Civil Litigation Principles that May Influence Criminal Matters

As the prevalence of e-discovery in criminal matters expands, the need to rely on established civil litigation rules and principles to make the criminal discovery process more efficient and predictable may increase, especially at early stages in criminal proceedings where various issues – including evidentiary disputes – can be resolved. Two such areas are Rules 26(f) and 26(b)(2)(B) and (C) of the Federal Rules of Civil Procedure.

Rule 26(f)

Rule 26(f) requires parties in civil litigation to meet and confer to develop a proposed discovery plan to be submitted to the court in writing, addressing various discovery issues and discovery scheduling. Issues relating to the disclosure and/or production of ESI, including the sources of such data, its form of production and the costs of such production, among others, are required to be discussed at the conference.

Although there is no criminal rule equivalent, it is just as important to identify and address potential ESI issues in criminal matters given, among other things, the severe consequences that can result from spoliation.[31] In conversations similar to a Rule 26(f) conference, when ESI issues are present, the parties should try to reach agreement on the scope, form and potential limits of production, the potential use of search terms, the use of non-waiver agreements for privileged materials and whether any unique ESI issues need to be addressed (such as third party repositories or dynamic and/or proprietary databases), among other items. *United States v. Graham*, discussed *supra*, provides an excellent example of the potential consequences that can occur when parties skip this step in criminal matters.

Interestingly, one judicial district has recently adopted "Best Practices for Electronic Discovery of Documentary Materials in Criminal Cases," which contains a specific requirement that the prosecutor and defense counsel meet and confer, post-indictment,

---

28 *United States v. Safavian*, 233 F.R.D. 12 (D.D.C. 2005).
29 *Safavian*, 233 F.R.D. at 18.
30 *Id.* at 18-19.
31 *See generally*, 18 U.S.C. § 1519 and 18 U.S.C. § 1512(c) (Obstruction of justice provisions of Sarbanes-Oxley Act of 2002).

about "electronic discovery of documentary materials."[32] At the discovery conference, the prosecutor and defense counsel must specifically address certain issues, including the nature and volume of discovery; the litigation capabilities of counsel; the time frame and process by which defense counsel will review the evidence and select documents it wishes to be produced in electronic or other format; that electronic discovery will be produced in a standard .pdf format; and that if access issues arise involving electronic discovery of audio and/or video files from a third party, the government and defense counsel will either provide the proprietary software needed to view the material or work together to devise a viewing format.[33] These "Best Practices" certainly recognize the lack of guidance in Fed. R. Crim. P. 16 relating to e-discovery, and may signal a shift towards a model based on civil principles. Moreover, in the absence of changes to the criminal rules to address ESI, local rules– such as from the Western District of Oklahoma – may attempt to fill this void for criminal matters.

Rules 26(b)(2)(B) and (C)

In addition to Rule 26(f), one of the hallmarks of the civil rules is proportionality of discovery, as discussed in Rule 26(b)(2)(B). In the civil realm, if a party from whom discovery is sought can show that the ESI is not "reasonably accessible because of undue burden or cost," the party may not have to provide that discovery, its responses may be limited by the court, or the court may order cost shifting to the requesting party.[34] Again, there is no criminal rule analog, but with the marked increase in ESI in criminal matters – particularly the rising volumes of ESI – proportionality will likely come into play, as it already has, for example, in the Western District of Oklahoma. The "Best Practices" for ESI in criminal cases instruct that "[o]pen communications between the government and defense counsel is critical to ensure that discovery is handled and completed in a manner agreeable to all parties. *Cost-sharing arrangements should also be discussed when appropriate*."[35] Further, the "Best Practices" mandate that at the discovery conference previously described, the "parties *will* discuss and consider in good faith possible cost-sharing measures in handling voluminous discovery . . . provided, however, that an ability to enter into cost-sharing agreements may be limited by the government's budget constraints and/or the Department of Justice requirement that we allocate litigation expenditures only for mandatory obligations."[36]

32 *See* General Order Regarding Best Practices for Electronic Discovery of Documentary Materials in Criminal Cases (W.D. Okla. Aug. 20, 2009) (ordering that the "Best Practices for Electronic Discovery of Documentary Materials in Criminal Cases" has been adopted). The Western District of Oklahoma is not the only judicial district to adopt protocols relating to e-discovery in criminal matters. *See also* Northern District of California Suggested Practices Regarding Discovery in Complex Cases and Northern District of California Protocol Regarding Discovery in Complex Cases (suggesting practices regarding discovery in wiretap and other complex, document-intensive cases but noting that the Practices and Protocol are not intended to expand the parties' discovery obligations under Federal Rule of Criminal Procedure 16, the Jencks Act, or other federal statutes or rules); and U.S. Attorney's Office Best Practices for Electronic Discovery of Documentary Materials in Large Cases (W.D. Wash. Sept. 2005) (requiring the U.S. Attorney's Office to discuss with all defense counsel whether electronic discovery of materials is appropriate and identifying other best practices that the U.S. Attorney's Office will pursue).

33 *See* Best Practices for Electronic Discovery of Documentary Materials in Criminal Cases ¶ 1. *See also* U.S. Attorney's Office Best Practices for Electronic Discovery of Documentary Materials in Large Cases (W.D. Wash. Sept. 2005) ("no later than the Local Rule 16 discovery conference the U.S. Attorney's Office will discuss with all defense counsel whether electronic discovery of documentary materials is appropriate in the case and, if so, what arrangements should be made").

34 *See generally*, *Mancia v. Mayflower Textile Servs. Co.*, 253 F.R.D. 354 (D. Md. 2008). FED. R. CIV. P. 26(b)(2)(C) states that: "On motion or on its own, the court must limit the frequency or extent of discovery otherwise allowed by these rules or by local rule if it determines that:
(i) the discovery sought is unreasonably cumulative or duplicative, or can be obtained from some other source that is more convenient, less burdensome, or less expensive;
(ii) the party seeking discovery has had ample opportunity to obtain the information by discovery in the action; or
(iii) the burden or expense of the proposed discovery outweighs its likely benefit, considering the needs of the case, the amount in controversy, the parties' resources, the importance of the issues at stake in the action, and the importance of the discovery in resolving the issues."

35 *See* Best Practices for Electronic Discovery of Documentary Materials in Criminal Cases (emphasis added).

36 *Id.* ¶ 4 (emphasis added). *See also* U.S. Attorney's Office Best Practices for Electronic Discovery of Documentary Materials in Large Cases ¶ 7 (W.D. Wash. Sept. 2005) ("At the Local Rule 16 discovery conference, we will discuss and consider in good faith possible cost-sharing measures in handling voluminous discovery, such as jointly-commissioned Bates numbering, scanning, and/or 'objective coding' of documentary materials by outside vendors; provided, however, that our ability to enter into cost-sharing agreements may be limited by our budget constraints and/or the Department of Justice requirement that we allocate litigation expenditures only for mandatory obligations").

While the development of proportionality principles related to e-discovery in criminal matters could produce a more efficient and effective discovery process in some instances, it may also create unique issues in the application of the Rule 26(b)(2)(C) "factors" to determine whether or not discovery of certain ESI should be limited – issues quite distinct from those confronted in civil matters. For example, a factor such as "the importance of the issues at stake in the action" may be disproportionately significant in a criminal case where the "issue at stake" may be an individual's liberty. Additional questions arise: How does one measure proportionality? By the seriousness of the alleged crime? How does one measure the relative value of the case – potentially one's liberty – against the burden or expense of the proposed discovery? And, from the government's point of view, at what point does the preservation, collection, searching, and production of ESI become so expensive that it is unable to prosecute potential crimes?

Conclusion

It is ironic that e-discovery in criminal matters, where defendants enjoy enhanced constitutional protections, is relatively ungoverned by clear procedural rights and remedies – at least on the face of the Federal Rules of Criminal Procedure – despite the complexities and challenges attendant to discovery involving ESI. However, certain civil litigation principles relating to the discovery of ESI are now beginning to influence criminal matters. Defense counsel and prosecutors would be prudent to monitor these developments and consider pro-actively applying the e-discovery principles and practices that civil litigators live with every day.

THE SEDONA CONFERENCE® COMMENTARY ON LEGAL HOLDS: THE TRIGGER & THE PROCESS

*A Project of The Sedona Conference® Working Group on Electronic Document Retention & Production (WG1)**

Author:
The Sedona Conference®

*2010 Editors:**
Thomas Y. Allman
Conor R. Crowley
Jonathan M. Redgrave

2007 Editors:
Conor R. Crowley
Eric Schwarz
Gregory J. Wood

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference® Working Group 1. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference.®

REPRINT REQUESTS:
Requests for reprints or reprint information should be directed to Richard Braman, Executive Director of The Sedona Conference,® at tsc@sedonaconference.org or 1-866-860-6600.

* The Sedona Conference® also wishes to acknowledge the contributions of all the Working Group 1 members who contributed to the dialogue while this Commentary was a work-in-progress. Finally, we wish to thank all of our Working Group Series℠ Sponsors, whose support is essential to The Sedona Conference's® ability to develop Working Group Series℠ publications such as this. For a complete listing of our Sponsors, click on the "Sponsors" navigation bar on the home page of our website at http://www.thesedonaconference.org.

Preface

Welcome to the 2010 final post-public comment version of *The Sedona Conference® Commentary on Legal Holds: The Trigger & The Process.* This document contains numerous changes from the 2007 public comment version. The changes reflect the informal and formal suggestions and comments we received in the past 2+ years since initial publication.

In a nutshell, the edits take into consideration the continued evolution of law and best practices in the area over the past few years. Just as the awareness and consideration of issues involved in implementing legal holds evolved significantly since the founding of the Working Group in 2002 to the initial publication of this document, so too has the legal world evolved since 2007. The guidelines and accompanying text have been revised to harmonize the enhanced understanding of the technical, process, and legal issues that have emerged since we first issued this Commentary. Notably, our treatment of the issues in this revised edition expressly addresses the recent *Pension Committee* and *Rimkus* cases that have been widely disseminated and discussed in 2010, as well as recent cases addressing the application of Rule 37(e) of the Federal Rules of Civil Procedure that was added in 2006.

While we have no doubt this area will continue to evolve, we believe this document represents an accurate view of reasonable and defensible practices that organizations should consider in 2010 and going forward when addressing the issue of legal hold triggers and process.

While all Working Group members, as well as public commentators, played a role in the revisions and enhancements to this document, I would like to especially thank Thomas Y. Allman, Conor R. Crowley, Jonathan M. Redgrave, and Kenneth J. Withers for their efforts in shepherding the final review and editing process for this document.

Finally, although this document is now in its post-public comment version, we welcome additional input and involvement in this and other publications of The Sedona Conference.® Please reach out to us at our website at www.thesedonaconference.org or email me at rgb@sedonaconference.org.

Richard G. Braman
Executive Director
The Sedona Conference®
August 2010

Introduction

PRESERVATION OBLIGATIONS AND LEGAL HOLDS

Information is the lifeblood of the modern world, a fact that is at the core of our civil discovery system. Accordingly, the law has developed rules regarding the manner in which information is to be treated in connection with litigation. One of the principal rules is that whenever litigation[1] is reasonably anticipated, threatened, or pending against an organization,[2] that organization has a duty to undertake reasonable and good faith actions to preserve relevant and discoverable information and tangible evidence. This duty arises at the point in time when litigation is reasonably anticipated whether the organization is the initiator or the target of litigation.

The duty to preserve requires a party to identify, locate, and maintain information and tangible evidence that is relevant to specific and identifiable litigation. It typically arises from the common law duty to avoid spoliation of relevant evidence for use at trial and is not explicitly defined in the Federal Rules of Civil Procedure. *See, e.g., Silvestri v. General Motors,* 271 F.3d 583 (4th Cir. 2001) (applying the "federal common law of spoliation"); *Chambers v. NASCO, Inc.,* 501 U.S. 32 (1991).

The concept of "legal holds" or "litigation holds"[3] has gained momentum in the last 10 years as part of common process by which organizations can begin to meet their preservation obligations. The use of a "litigation hold" as a means to satisfy preservation obligations was popularized by the 2003 decision in *Zubulake v. UBS Warburg* ("*Zubulake IV*").[4] The court suggested that "[o]nce a party reasonably anticipates litigation, it must suspend its routine document retention/destruction policy and put in place a 'litigation hold.'"[5] In a subsequent decision in 2010 in *Pension Committee v. Bank of America Securities, LLC,*[6] the same court held that "the failure to issue a written litigation hold constitutes gross negligence." However, not all courts require a written legal hold. For example, in *Kinnally v. Rogers Corporation,*[7] a district court held that sanctions do not lie merely because of the "absence of a *written* litigation hold"[8] when a party has taken "the appropriate actions to preserve evidence."[9]

The 2006 Amendments

The 2006 Amendments to the Federal Rules of Civil Procedure ("the 2006 Amendments" or "the Amendments") did not define preservation obligations given the difficulty in drafting an appropriate Rule.[10] The use of a "litigation hold" as a method of implementation was referenced, however, in the Note to Rule 37(f), which was subsequently renumbered as Rule 37(e).[11]

1 Throughout this Commentary, the term "litigation" is used to refer primarily to civil litigation. However, the principles apply with equal force to regulatory investigations and proceedings.
2 Where appropriate, the term "organization" should be understood to include natural persons.
3 Throughout this Commentary we use the term "legal hold" rather than "litigation hold" to reflect that the duties and processes may apply in circumstances where there is no litigation. (e.g., pre-litigation or investigation).
4 220 F.R.D. 212 (S.D.N.Y. 2003).
5 *Id.* at 218.
6 2010 WL 184312 (S.D.N.Y. Jan. 15, 2010). The opinion was subtitled "*Zubulake* Revisited: Six Years Later."
7 2008 WL 4850116, at *7 (D. Ariz. Nov. 7, 2008).
8 *Id.* *6 ("the absence of a written litigation hold . . . does not in itself establish [a violation]" (emphasis in original).
9 *Id.* at *7 (noting use of a verbal litigation hold).
10 ADVISORY COMMITTEE MINUTES, Apr. 14-15, 2005, at p. 39-40 ("the Committee has concluded that the difficulties of drafting a good rule would be so great that there is no occasion even to consider the question whether a preservation rule would be an authorized or wise exercise of Enabling Act authority."); copy available at http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/Minutes/CRAC0405.pdf.
11 FED R. CIV. P. 37(f), Committee Note observes that "[w]hen a party is under a duty to preserve information because of pending or reasonably anticipated litigation, intervention in the routine operation of an information system is one aspect of what is often called a 'litigation hold.'" (2006).

The primary focus in the Amendments was on process improvements designed to encourage early party agreements while providing guidance for courts facing losses due to preservation failures. Thus, Rule 26(f) now requires discussion of "issues about preserving discoverable information" at the "meet and confer" held prior to the Scheduling Conference required by Rule 16(b). The Committee hoped that by encouraging early discussion, parties would reach agreement on "reasonable preservation steps." However, the requirement to discuss preservation "does not imply" that courts should routinely enter preservation orders.[12] In addition, Rule 37(e) provides that "[a]bsent exceptional circumstances, a court may not impose sanctions under these rules on a party for failing to provide electronically stored information lost as a result of the routine, good-faith operation of an electronic information system."

The Duty to Preserve

In enforcing the duty to preserve through spoliation sanctions, courts primarily rely upon their inherent powers, although Rule 37 also plays a limited role where a court order has been violated. A party that violates a preservation order or an order to compel production, or otherwise fails to preserve and produce information, may be exposed to a range of sanctions.[13]

Preservation obligations may also be acknowledged and enforced because of statutes or regulations that are deemed to apply under the circumstances at issue.[14] *See Byrnie v. Town of Cromwell Board of Education*[15] ("Several courts have held that the destruction of evidence in violation of a regulation that requires its retention can give rise to an inference of spoliation."). Criminal penalties at the federal and state level may also be invoked in specific cases within the coverage of those laws. *See, e.g.,* 18 U.S.C. § 1519 (Sarbanes-Oxley Act § 802).

A duty to preserve may arise or be "triggered" before commencement of litigation. The duty "arise[s] not only during litigation but also extends to that period before the litigation when a party reasonably should know that the evidence may be relevant to anticipated litigation."[16] Once it arises, a party must take reasonable steps to preserve "what it knows, or reasonably should know is relevant in the action, is reasonably calculated to lead to the discovery of admissible evidence, is reasonably likely to be requested during discovery and/or is the subject of a pending discovery request."[17] In some states, including federal courts sitting in diversity, an independent action in tort may lie for violation of a duty to preserve.[18]

The basic principle that an organization has a duty to preserve relevant information in anticipation of litigation is easy to articulate. However, the precise application of that duty can be elusive. Every day, organizations apply the basic principle to real-world circumstances, confronting the issue of when the obligation is triggered and,

12 FED. R. CIV. P. 26, Committee Note, Subdivision (f) (2006).
13 The preservation obligation typically applies only to parties, although service of a subpoena can trigger a duty as to non-parties. *See, e.g., Caston v. Hoaglin*, 2009 WL 1687927 (S.D. Ohio June 12, 2009) (subpoenas issued for the purposes of requiring preservation of relevant information).
14 Some record retention regulations that create preservation obligations are not necessarily enforceable for the benefit of private parties. *See* 17 C.F.R. § 240.17a-4 (SEC rule mandating retention of communications by members, brokers, or dealers), as discussed in *Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309, 322 n.70 (S.D.N.Y. 2003) (plaintiff was not an intended beneficiary of the records retention regulation at issue).
15 243 F.3d 93, 108-09 (2d Cir. 2001).
16 *Silvestri v. General Motors,* 271 F.3d 583, 591 (4th Cir. 2001).
17 *Wm T. Thompson Co. v. General Nutrition Corp.,* 593 F. Supp. 1443 (C.D. Cal. 1984).
18 *See, e.g., Kearney v. Foley & Lardner,* 582 F.3d 396 (9th Cir. 2009) (dismissing state claim because of other viable remedies).

once triggered, what is the scope of the obligation. Principle 5 of *The Sedona Principles: Best Practices, Recommendations & Principles for Addressing Electronic Document Production* ("*The Sedona Principles*") (2d ed. 2007), suggests that preservation obligations require "reasonable and good faith efforts," but that it is "unreasonable to expect parties to take every conceivable step to preserve potentially relevant data."

This Commentary is intended to provide pragmatic suggestions and guidance in carrying out preservation obligations. For ease of analysis, the Commentary is divided into two parts: The "trigger" and the "legal hold." Part I addresses the "trigger" issue and provides practical guidelines for determining when the duty to preserve relevant information arises. What should be preserved and how the preservation process should be undertaken, including the implementation of "legal holds," are addressed in Part II.

The keys to addressing these issues, as with all discovery issues, are reasonableness and good faith. Where ESI is involved, there are also practical limitations due to the inaccessibility of sources[19] as well as the volume, complexity and nature of electronic information, which necessarily implicates the proportionality principles, found in Rule 26(b)(2)(C)(iii).[20]

The Guidelines in this Commentary are intended to facilitate compliance by providing a framework an organization can use to create its own preservation procedures. *The Guidelines are not intended and should not be used as an all-encompassing "checklist" or set of rules that are followed mechanically.* Instead, they should guide organizations in articulating a policy for implementing legal holds that is tailored to their individual needs. In addition to the Guidelines, suggestions and illustrations are included under hypothetical factual situations. These illustrations are not to be taken as "right answers" for the circumstances posed. Indeed, there may be other circumstances or facts that could well result in a different analysis and result. As such, the illustrations are intended to impart understanding of the analytical framework to be applied and not to be considered as reasons for reaching a particular result.

Guideline 1
A reasonable anticipation of litigation arises when an organization is on notice of a credible probability that it will become involved in litigation, seriously contemplates initiating litigation, or when it takes specific actions to commence litigation.

Guideline 2
Adopting and consistently following a policy or practice governing an organization's preservation obligations are factors that may demonstrate reasonableness and good faith.

Guideline 3
Adopting a process for reporting information relating to a probable threat of litigation to a responsible decision maker may assist in demonstrating reasonableness and good faith.

---

19 FED. R. CIV. P. 26(b)(2)(B) provides that information stored in sources that are not reasonable accessible because of undue burden or cost are not initially discoverable, but the fact that they may become so if "good cause" is shown prompts them to be a subject of consideration for possible preservation.

20 The Hon. Paul W. Grimm et al., *Proportionality in the Post-Hoc Analysis of Pre-Litigation Preservation Decisions,* 37 U. BALT. L. REV 381, 388 (2008) (it would be "anomalous to sanction a party" for failure to preserve information that is later determined by the court not to be discoverable under FED. R. CIV. P. 26(b)(2)(C)).

Guideline 4
Determining whether litigation is or should be reasonably anticipated should be based on a good faith and reasonable evaluation of relevant facts and circumstances.

Guideline 5
Evaluating an organization's preservation decisions should be based on the good faith and reasonableness of the decisions undertaken (including whether a legal hold is necessary and how it should be executed) at the time they are made.

Guideline 6
The duty to preserve involves reasonable and good faith efforts, taken as soon as is practicable and applied proportionately, to identify and, as necessary, notify persons likely to have relevant information to preserve the information.

Guideline 7
Factors that may be considered in determining the scope of information that should be preserved include the nature of the issues raised in the matter, the accessibility of the information, the probative value of the information, and the relative burdens and costs of the preservation effort.

Guideline 8
In circumstances where issuing a legal hold notice is appropriate, such a notice is most effective when the organization identifies the custodians and data stewards most likely to have relevant information, and when the notice:

(a) Communicates in a manner that assists persons in taking actions that are, in good faith, intended to be effective
(b) Is in an appropriate form, which may be written
(c) Provides information on how preservation is to be undertaken
(d) Is periodically reviewed and, when necessary, reissued in either its original or an amended form, and
(e) Addresses features of relevant information systems that may prevent retention of potentially discoverable information.

Guideline 9
An organization should consider documenting the legal hold policy, and, when appropriate, the process of implementing the hold in a specific case, considering that both the policy and the process may be subject to scrutiny by opposing parties and review by the court.

Guideline 10
Compliance with a legal hold should be regularly monitored.

Guideline 11
Any legal hold policy, procedure, or practice should include provisions for releasing the hold upon the termination of the matter at issue so that the organization can adhere to policies for managing information through its useful lifecycle in the absence of a legal hold.

PART 1: TRIGGERING THE DUTY OF PRESERVATION

The duty to preserve relevant information arises when litigation is "reasonably anticipated." The duty to preserve relevant information is certainly triggered when a complaint is served, a governmental proceeding is initiated, or a subpoena is received. However, the duty to preserve could well arise before a complaint is served or a subpoena is received and regardless of whether the organization is bringing the action, is the target of the action, or is a third party possessing relevant evidence. The touchstone is "reasonable anticipation."

Determining whether a duty to preserve is triggered is fact-intensive and is not amenable to a one-size-fits-all or a checklist approach. An organization will likely not be able to resolve the question the same way each time it arises. In general, determining whether the duty to preserve attaches will require an approach that considers a number of factors, including the level of knowledge within the organization about the claim, and the risk to the organization of the claim. Weighing these factors will enable an organization to decide when litigation is reasonably anticipated and when a duty to take affirmative steps to preserve relevant information has arisen.

Guideline 1. A reasonable anticipation of litigation arises when an organization is on notice of a credible probability that it will become involved in litigation, seriously contemplates initiating litigation, or when it takes specific actions to commence litigation.

When the duty to preserve arises is often unambiguous. For example, the receipt of a summons or complaint, receipt of a subpoena, or formal notice that an organization is the target of a governmental investigation puts an organization on notice that it has a duty to preserve relevant information. However, other events may trigger a duty to preserve only when considered in the context of the entity's history and experience, or the particular facts of the case. For instance, an insurer's receipt of a claim from an insured often will not indicate the probability of litigation, as the insurer is in the business of paying claims often without litigation. On the other hand, the filing of an EEOC charge by a current or former employee may or may not, in the experience of the employer, indicate a probability of litigation. Similarly, the receipt of a preservation notice letter from an opposing party may or may not give rise to a credible probability of litigation, depending on the circumstances.

On the plaintiff's side, seeking advice of counsel, sending a cease and desist letter or taking specific steps to commence litigation may trigger the duty to preserve. In both *Pension Committee*[21] and *Rimkus Consulting v. Cammarata,*[22] the activities of the plaintiffs prior to litigation came under close examination. The test of the timing of the trigger is often based on when the party "determine[d] [that] legal action is appropriate."[23]

On the defendant's side, credible information that it is the target of legal action may be sufficient to trigger the duty to preserve. The degree to which anticipated litigation must be clear and certain is debatable. In *Goodman v. Praxair Services*,[24] the court refused to require an unequivocal notice of impending litigation. In *Phillip M. Adams & Associates v.*

21 2010 WL 184312 (S.D.N.Y. Jan. 15, 2010).
22 2010 WL 645353 (S.D. Tex. Feb. 19, 2010).
23 *Milenkamp v. Davisco Foods Int'l*, 562 F.3d 971, 981 (9th Cir. 2009) (no duty to preserve since destruction of evidence occurred "by the time" that plaintiffs determined legal action was appropriate).
24 632 F. Supp. 2d 494, 494 at n. 7 (D. Md. 2009).

Dell, Inc.,[25] the duty to preserve was held to have been triggered many years before suit was filed because of mere awareness of the dispute by others in the industry. However, there are circumstances when the threat of litigation is not credible and it would be unreasonable to anticipate litigation based on that threat. In *Cache LaPoudre Fees v. Land O'Lakes*,[26] for example, a letter referencing potential "exposure" did not trigger the obligation to preserve since a mere possibility of litigation does not necessarily make it likely and the letter referred to the possibility of amicable resolution.

This Guideline suggests that a duty to preserve is triggered *only* when an organization concludes (or should have concluded), based on credible facts and circumstances, that litigation or a government inquiry is probable. Whether litigation can be reasonably anticipated should be based on a good faith and reasonable evaluation of the facts and circumstances as they are known at the time. Of course, later information may require an organization to reevaluate its determination and may result in a conclusion that litigation that previously had not been reasonably anticipated (and consequently did not trigger a preservation obligation) is then reasonably anticipated. Conversely, new information may enable an organization to determine that it should no longer reasonably anticipate a particular litigation, and that it is consequently no longer subject to a preservation obligation. A party that obtains new information, after the initial decision is made, should reevaluate the situation as soon as practicable. *See, e.g., Stevenson v. Union Pacific RR Co.,* 354 F.3d 739 (8th Cir. 2004).

Consequently, to help understand when the duty to preserve arises, one should consider when the duty does *not* arise. For example, a vague rumor or indefinite threat of litigation does not trigger the duty; nor does a threat of litigation that is not credible or not made in good faith. A lack of credibility may arise from the nature of the threat itself or from past experience regarding the type of threat, the person who made the threat, the legal bases upon which the threat is purportedly founded, or any of a number of similar facts. In addition, the trigger point for a small dispute, where the stakes are minor, might occur at a later point than for a dispute that is significant in terms of business risk or financial consequences.[27]

A reasoned analysis of the available facts and circumstances is necessary to conclude whether litigation or a government inquiry is or is not "reasonably anticipated." That determination is fact-intensive and should be made by an experienced person who can make a reasoned judgment.

Another issue to be considered is what constitutes notice to the organization. For corporations this can be a complicated issue. If one employee or agent of the organization learns of facts that might lead one to reasonably believe litigation will be forthcoming, should that knowledge be imputed to the organization as a whole, thereby triggering its preservation obligations? Often, the answer will depend on the nature of the knowledge, the potential litigation, and the agent. Generally, "[a]n agent's knowledge is imputed to the corporation where the agent is acting within the scope of his authority and where the knowledge relates to matters within the scope of that authority."[28]

---

25 621 F. Supp. 2d 1173 (D. Utah 2009).

26 244 F.R.D. 614, 623, 623 (D. Colo. 2007).

27 A dispute that may have significant consequences will be more likely to result in litigation because the entity potentially asserting the claim is more likely to be willing to bear the costs of litigation. Thus, such a dispute is more likely to result in a reasonable anticipation of litigation.

28 *In re Hellenic, Inc.*, 252 F.3d 391, 395 (5th Cir. 2002).

Organizations that become aware of a credible threat from which litigation might arise may have a duty to make reasonable inquiry or possibly undertake a more detailed investigation regarding the facts related to that "threat." Whether an inquiry or detailed investigation is warranted will be fact-driven and based on reasonableness and good faith. Thus, while there may be no duty to affirmatively disprove allegations associated with a threat before concluding that a threat lacks credibility, the facts and circumstances may suggest the prudence of making an inquiry before reaching such a conclusion.

The case law as to when an organization should reasonably anticipate litigation varies widely from jurisdiction to jurisdiction. In *Zubulake IV*,[29] the court stated that UBS should have reasonably anticipated litigation *at the latest* when Ms. Zubulake filed a charge with the EEOC. However, the court found that UBS reasonably anticipated litigation—thereby triggering the duty to preserve—five months before the filing of the EEOC charge, based on the emails of several employees revealing that they knew that plaintiff intended to sue.[30]

In *Willard v. Caterpillar, Inc.*,[31] the court rejected a claim that the defendant tractor manufacturer should have preserved documents related to the design of the tractor, where the model at issue had been out of production for 20 years. The court noted that:

> There is a tendency to impose greater responsibility on the defendant when spoliation will clearly interfere with the plaintiff's prospective lawsuit and to impose less responsibility when the interference is less predictable. Therefore, if Caterpillar destroyed documents which were routinely requested in ongoing or clearly foreseeable products liability lawsuits involving the D7-C tractor and claims similar to Willard's, its conduct might be characterized as unfair to foreseeable future plaintiffs. However, the document destruction at issue began more than ten years before Willard was injured, and the evidence disclosed only one other accident involving on-track starting and none involving the wet clutch. In our opinion, such remote pre-litigation document destruction would not be commonly understood by society as unfair or immoral.

ILLUSTRATIONS

Illustration i: An organization receives a letter that contains a vague threat of a trade secret misappropriation claim. The letter does not specifically identify the trade secret. Based on readily available information, it appears that the information claimed to be the misappropriated trade secret had actually been publicly known for many years. Furthermore, the person making the threat had made previous threats without initiating litigation. Given these facts, the recipient of the threat could reasonably conclude that there was no credible threat of litigation, and the entity had no duty to initiate preservation efforts.

Illustration ii: An organization receives a demand letter from an attorney that contains a specific threat of a trade secret misappropriation claim. Furthermore, the organization is aware that others have been sued by this same plaintiff on similar claims. Given these facts,

29 220 F.R.D. 212, 216-17 (S.D.N.Y. 2003).
30 The scope of that duty to preserve seems to have been quite limited encompassing a small number of emails over a limited period of time suggesting that even though the duty to preserve had arisen, the scope of the preservation obligations may have been quite modest.
31 40 Cal. App. 4th 892 (1995).

there is a credible threat of litigation, and the organization has a duty to preserve relevant information. The duty to preserve on the part of the potential plaintiff arises no later than the date of the decision to send the letter, and, in some circumstances, may arise earlier.

Illustration iii: An organization learns of a report in a reputable news media source that includes sufficient facts, consistent with information known to the organization, of an impending government investigation of a possible violation of law by the organization stemming from the backdating of stock options given to executives. Under these circumstances, a government investigation (and possibly litigation) can reasonably be anticipated and a preservation obligation has arisen.

Illustration iv: An event occurs which, in the experience of the organization, typically results in litigation. Examples of such events may include a plant explosion with severe injuries, an airplane crash, or an employment discrimination claim. The experience of the organization when these claims arose in the past would be sufficient to give rise to a reasonable anticipation of litigation.

Illustration v: A cease-and-desist letter for misuse of a trademark is received by a business. The recipient replies with an agreement to comply with the demand and, in fact, does comply with the demand. The recipient does not have a reasonable basis to anticipate litigation and does not have an obligation to preserve relevant information. However, the duty to preserve on the part of the sender arises no later than the date of the decision to send the letter.

Guideline 2. Adopting and consistently following a policy or practice governing an organization's preservation obligations is one factor that may demonstrate reasonableness and good faith.

A policy or practice setting forth a process for determining whether the duty to preserve information has attached can help ensure that the decision is made in a defensible manner. As stated in *The Sedona Principles*,[32] "[b]y following an objective, preexisting policy, an organization can formulate its responses to electronic discovery not by expediency, but by reasoned consideration."[33] Thus, any policies that provide for management of ESI should include provisions for implementing procedures to preserve documents and electronically stored information related to ongoing or reasonably anticipated litigation, government investigations or audits. *Id*. However, "[t]he nomenclature (e.g., 'litigation hold') is not important; the important factor is that the organization has a means to comply with its legal obligations to preserve relevant information in the event of actual or reasonably anticipated litigation or investigation." [34]

While the particulars of the policy or practice will necessarily be driven by the structure and culture of the organization, the key is to have a process that is followed. In cases where the preservation efforts are likely to be challenged, it can be helpful to memorialize the steps taken to follow that process so the organization can demonstrate its compliance with the process. A defined policy and memorialized evidence of compliance should provide strong support if the organization is called upon to prove the reasonableness of the decision-making process.

32 The Sedona Conference,® THE SEDONA PRINCIPLES (2d ed. 2007) at Comment 1.b.
33 Principle 1 of *The Sedona Principles* provides, in relevant part, that "[o]rganizations must properly preserve electronically stored information that can reasonably be anticipated to be relevant to litigation."
34 *Id.* at n. 36.

ILLUSTRATION

Illustration i: Upon receipt of an anonymous threat sent to a corporation's ombudsman, the ombudsman consults the legal hold policy. That policy provides criteria for an assessment of the threat and whether the issues raised by information, including the circumstances surrounding its receipt, indicate the potential for litigation or governmental investigation. It also provides for a preliminary evaluation of the allegations before determining whether a hold should be implemented. Based on the policy, the ombudsman concludes that the corporation does not reasonably anticipate litigation and memorializes that decision in a memorandum to the file. In a subsequent challenge, the corporation is able to demonstrate that it exercised reasonableness and good faith.

Guideline 3. Adopting a process for reporting information relating to a probable threat of litigation to a responsible decision maker may assist in demonstrating reasonableness and good faith.

In any organization—but particularly in large organizations—individuals within the organization may have information that indicates a threat of litigation that the decision makers for the organization do not have. An organization formulating a legal hold policy should consider how to enable that information to be communicated to persons charged with evaluating the threat and, if warranted, instituting legal holds. The particulars of how this process is implemented will vary from organization to organization, based on the way the business is conducted and the culture of the organization. However, to be effective, the procedure should be simple and practical, and individuals within the organization should be trained on how to follow the procedure.

ILLUSTRATIONS

Illustration i: Westerberg Products is a large corporation with tens of thousands of employees and offices throughout the United States. Westerberg Products establishes an internal compliance "help line" or Web site that allows employees to submit information they have regarding matters of concern, including potential claims against the company. The information received is forwarded to the legal department of Westerberg Products, which is charged with determining whether and when to implement a legal hold. Each employee is trained on how to use the help line or Web site and instructed that they should use it to report any relevant information. Westerberg Products can use these procedures to demonstrate its good faith efforts to ensure it is aware of information indicating a threat of litigation.

Illustration ii: Stinson Software is a small software developer with eight employees. Every month, all eight employees attend a staff meeting and a regular topic of discussion is whether any employee is aware of any ongoing threats to the company, including possible claims or demands that might result in litigation against the company. Stinson Software's Chief Operations Officer follows up on any tips with Stinson Software's outside counsel. Stinson Software can use these practices to demonstrate its good faith effort to ensure it is aware of information indicating a threat of litigation.

Guideline 4. Determining whether litigation is or should be reasonably anticipated should be based on a good faith and reasonable evaluation of relevant facts and circumstances.

Determining whether litigation is or should be reasonably anticipated requires considering many different factors. Depending on the nature of the organization and the nature of the litigation, factors that might be pertinent to consider could include:

- The nature and specificity of the complaint or threat;
- The party making the claim
- The business relationship between the accused and accusing parties;
- Whether the threat is direct, implied, or inferred
- Whether the party making the claim is known to be aggressive or litigious
- Whether a party who could assert a claim is aware of the claim
- The strength, scope, or value of a known or reasonably anticipated claim
- Whether the company has learned of similar claims
- The experience of the industry, and
- Reputable press and/or industry coverage of the issue either directly pertaining to the client or of complaints brought against someone similarly situated in the industry.

These factors are not exhaustive. They and other considerations must be weighed reasonably and in good faith in the context of what steps are reasonable and practicable.

ILLUSTRATIONS

Illustration i: A musician writes a song that sounds very similar to a famous song. Immediately there are critical reviews and radio DJs calling the song a "blatant rip-off." Although the copyright owners of the original song have not yet made any claim, the high profile nature of the criticism is a consideration that may lead a determination by the music publisher that a preservation obligation has arisen.

Illustration ii: A restaurant chain's central management office receives a series of anonymous emails purporting to be from customers claiming food poisoning after the much-publicized introduction of a new dish. In the absence of any corroborating reports from the restaurants and with no specific details on which to act, the chain's counsel reasonably concludes that litigation is not reasonably anticipated.

Guideline 5. Evaluating an organization's preservation decisions should be based on the good faith and reasonableness of the decisions (including whether a legal hold is necessary and how it should be executed) at the time they were made.[35]

The reasonableness of an organization's preservation decisions, such as whether to implement a legal hold, can only be made in light of the facts and circumstances reasonably known to it at the time of its decision, and not on the basis of hindsight or information acquired after the decisions are made. An organization seeking to determine whether a preservation obligation has arisen has no choice but to rely on the information available to it; consequently, whether decisions made were reasonable should turn on that knowledge, and not other circumstances of which the organization was unaware.

35 Similarly, judicial evaluation of an organization's legal hold implementation should be based on the good faith and reasonableness of the implementation at the time the hold was implemented. In doing so, proportionality considerations may become relevant.

ILLUSTRATION

Illustration i: Joey Music Co. manufactures music compact discs using a state-of-the-art process it licenses from D.D. Electronics. D.D. Electronics also licenses the process to Johnny Computing, which uses the process to manufacture CD-ROMs. In January, Johnny Computing receives reports that many of the compact discs it has sold are defective. After investigating, Johnny Computing determines that the defect is caused by the process it licenses from D.D. Electronics. The news of this discovery is kept out of the media, and the class action case brought by Johnny Computing's customers is quickly settled out of court by March. In April, Joey Music, who had no knowledge of the suit against Johnny Computing or the subsequent settlements, disposes of certain documents relating to its use of the D.D. Electronics process. In May, Joey Music begins receiving complaints from its customers. Because Joey Music had no knowledge of the concerns with the process it licenses from D.D. Electronics, its decision to dispose of documents in April was reasonable, particularly if done in compliance with an existing records and information management policy.

PART 2: IMPLEMENTING THE LEGAL HOLD

Once the duty to preserve information arises, an organization must decide what to preserve and how to do it. In some circumstances, the duty to preserve requires only locating and preserving a limited number of documents. In other circumstances, the scope of the information is larger and the sources of the information may not be known to counsel.

The typical legal hold process focuses on key custodians and data stewards,[36] who are asked to take steps to preserve relevant information and help prevent losses due to routine business operations. The effort involves discoverable material, i.e., usually that "relevant to a claim or defense.[37] As noted by one court, there is no broad requirement to preserve information that is not relevant: "[m]ust a corporation, upon recognizing the threat of litigation, preserve every shred of paper, every email or electronic document, and every backup tape? The answer is clearly, 'no.' Such a rule would cripple large corporations."[38]

Identifying and preserving potentially relevant information can be complex and may require trained people, processes, and technology, particularly when ESI is at issue. This may include creating teams to identify the sources, custodians, and data stewards of potentially relevant information within the organization, and to define what needs to be preserved and to coordinate with outside counsel. It is often advisable to maintain sources of ESI in their native formats with metadata[39] to preserve the ability to make production in some variant of a native file format, if necessary. In the case of *In re Priceline.Com Inc. Securities Litigation*,[40] a court approved an agreement that the original data would be maintained in its original file format for the duration of the litigation. The need to produce metadata is also recognized in Principle

36 Of course, while the focus is on key custodians and data stewards, sometimes referred to as "key players," there may be other individuals who are asked to take preservation steps. Notably, the efforts undertaken for key custodians may be different from other custodians.

37 FED. R. CIV. P. 26(b)(1). In some cases, the rule states, "for good cause" the court may order discovery of any matter relevant to the subject matter involved in the action. Logically, the duty to preserve information relevant to the broader scope would not attach until at least the motion or order to expand the scope: before that, discovery under the broader scope would not be reasonably likely.

38 *Zubulake v. UBS Warburg*, 220 F.R.D. 212, 217 (S.D.N.Y. 2003).

39 *Compare U.S. v. O'Keefe*, 2008 WL 449729 (D.D.C. Feb. 18, 2008) (court applying amended Rule 34(b) as persuasive authority in the criminal discovery context ordered preservation of ESI in its native format with metadata until ruling regarding production).

40 233 F.R.D. 88 (D. Conn. 2005).

12 of *The Sedona Principles* which recognizes "the need to produce reasonably accessible metadata that will enable the receiving party to have the have the same ability to access, search, and display the information as the producing party where appropriate or necessary in light of the nature of the information and the needs of the case."

For large preservation efforts, a process that is planned, systemized, and scalable is useful, although *ad hoc* manual processes are often appropriate for cases involving relatively small numbers of key custodians and identifiable issues. It is usually inefficient to collect information from every custodian, server, or other source of active data without making any initial effort to identify relevant information. With no means to triage the information and to filter out irrelevant ESI, the collection may be overbroad, with a great deal of irrelevant information aggregated into a central repository where it is then further processed and searched.

There is a growing consensus that the proportionality principle must be applied in assessing preservation issues. In *Rimkus Consulting*, the court noted that "[w]hether preservation or discovery conduct is acceptable in a case depends on what is *reasonable,* and that in turn depends on whether what was done—or not done—was *proportional* to that case and consistent with clearly established applicable standards."[41] (emphasis in original). Similarly, the Seventh Circuit Pilot Program on E-Discovery (2009)[42] provides, in Principle 2.04 (Scope of Preservation), that "every party to litigation and its counsel are responsible for taking reasonable and proportionate steps to preserve relevant and discoverable ESI within its possession, custody or control."

To develop an appropriate process for a large organization, the responsible business and functional units, including legal, IT, and records management personnel, should be trained on the organization's legal hold policies and practices and their responsibilities. Plans and protocols appropriate to the type of data and the manner in which it is maintained should be developed. Consultants and vendors can also play a valuable role by helping to design efficient and systemized processes that are executed by IT personnel and/or consultants. For smaller cases, or for entities without internal resources, outside counsel may provide the services on a case-by -case basis and may be deeply involved in drafting the initial preservation notices and in collecting documents and ESI.

While the traditional role of counsel is to "inform the client of its duty to preserve potentially relevant documents in the client's custody or control and of the possible consequences of failing to do so,"[43] some decisions hold that counsel also owes an independent duty to actively supervise a party's compliance with the duty to preserve.[44] In *Zubulake V*,[45] the court went further and suggested that "counsel must issue a 'litigation hold' at the outset of litigation or whenever litigation is reasonably anticipated, communicate directly with the 'key players,' instruct all employees to produce electronic copies of their relevant active files," and secure unique backup media that should be retained.[46]

---

41 2010 WL 645353 at *6 (S.D. Tex. Feb. 19, 2010).

42 SEVENTH CIRCUIT ELECTRONIC DISCOVERY PILOT PROGRAM (Oct. 2009) at 21, *available at* http://www.ilcd.uscourts.gov/Statement%20-%20Phase%20One.pdf.

43 Standard 10, ABA CIVIL DISCOVERY STANDARDS (Aug. 2004) ("This Standard is . . . an admonition to counsel that it is counsel's responsibility to advise the client as to whatever duty exists, to avoid spoliation issues.") *See Turner v. Hudson Transit Lines*, 142 F.R.D. 68, 73 (S.D.N.Y. 1991) (the preservation obligation runs first to counsel, who has a duty to advise, with "corporate managers" having the responsibility to convey that information to the relevant employees).

44 *Phoenix Four, Inc. v. Strategic Resources Corp.,* 2006 WL 1409413 (S.D.N.Y. May 23, 2006).

45 229 F.R.D. 422 (S.D.N.Y. 2004).

46 Counsel was not sanctioned for their preservation failures in *Zubulake* since a party on notice of its obligations "acts at its own peril." *Id.* at 434; *cf. Green v. McClendon,* 262 F.R.D. 284, 290 (S.D.N.Y. 2009) (monetary sanctions imposed on *both* counsel and client for failure to institute a timely legal hold).

The following Guidelines are intended to help organizations create legal hold procedures that are effective in preserving necessary information in a manner consistent with the requirements of the situation at hand. As with the triggers of the legal hold, there is no one-size-fits-all answer to implementing a legal hold. Rather, organizations must approach implementing a legal hold in light of the particular documents and information in their possession, the nature of the matter, and the culture of the organization.

Guideline 6. The duty to preserve involves reasonable and good faith efforts, taken as soon as is practicable and applied proportionately, to identify and, as necessary, notify persons likely to have relevant information to preserve the information.

After determining it has a duty to preserve, the organization should begin to identify information to be preserved. The obligation to preserve ESI requires reasonableness and good faith efforts, but it is "unreasonable to expect parties to take every conceivable step to preserve all potentially relevant data."[47] The organization should consider the sources of information within its "possession, custody, and control"[48] that are likely to include relevant, unique information. The most obvious of these sources are those that the organization physically has in its possession or custody—for example, the file cabinets of documents in its office, the emails that reside on its servers located in its corporate headquarters — but also may include sources such as thumb drives, company furnished laptops, and PDAs used by employees for business purposes.

Some sources of information under the control of third parties may also be deemed to be within the control of the organization because of contractual or other relationships. Examples include information held by outsourced service providers, storage facilities operators, and application service providers (ASPs).[49] With respect to those sources, the organization should consider providing appropriate notice concerning the need to preserve material that is likely to be relevant.

In executing preservation obligations, special attention should be paid, where necessary, to information that is held outside of the United States. Many such locations have laws that potentially conflict with United Sates discovery requirements. Such laws include those that limit the retention of certain types of information and those that limit the processing or transfer of information to the United States for discovery purposes.[50]

It must be noted that a mere delay in implementing a legal hold is not necessarily fatal. In *Rahman v. The Smith & Wollensky Restaurant Group*,[51] the court concluded that "even assuming there was, in fact, no litigation hold" until late in the litigation, the plaintiff had failed to establish that there was "any gap" in production which was "attributable to the

47 *See* Principle 5, THE SEDONA PRINCIPLES (2d ed. 2007).
48 *See* FED. R. CIV. P. 34 and its state equivalents; *see also In re NTL, Inc., Sec. Litig.,* 244 F.R.D. 179 (S.D.N.Y. 2007) (party had access to records held by third party).
49 Notably, the advent of "cloud computing" will, over time, likely increase the number of organizations using third parties to host, manage, store, and retrieve electronic information in the course of business.
50 *See, e.g.*, The European Directive 95/46/EC (the "Directive"), effective October 1998. The Directive governs the processing and use of personal data for all EU Member States, and identifies eight data protection principles. These include the principle that personal data shall not be kept for longer than is necessary for the purposes for which it is processed and the principle that personal data shall not be transferred to a country or territory outside the EU, unless that country or territory ensures an "adequate" level of protection for the rights and freedom of data subjects in relation to the processing of personal data. At this time, the United States is not considered by the EU to ensure an "adequate" level of protection and data may be transferred only if the transfer meets a particular exception found in the Directive or if certain steps are taken to qualify for the European Commission's Safe Harbor status or to adopt the European Commission's model contractual clauses for Data Transfer and Data Processing or Binding Corporate Rules. One exception to the Directive that may apply in certain cases is when the transfer is required for the exercise or defense of legal claims.
51 2009 WL 773344 (S.D.N.Y. Mar. 18, 2009).

failure to institute [a] litigation hold at an earlier date."[52] The test is what was reasonable under the circumstances, with an eye towards the ultimate end goal (e.g., whether relevant information was preserved). Thus, there is no per se negligence rule and if the organization otherwise preserved the information then there is no violation of the duty to preserve.[53]

ILLUSTRATION

Illustration i: Strummer Holdings is a large corporation that sends many of its historic documents to an offsite storage facility managed by Jones Storage. Typically, documents older than five years are sent to Jones Storage. At all times, Strummer Holdings retains all legal rights with respect to the documents, and has the right to require their return from Jones Storage at any time. Jones Storage has standing instructions from Strummer Holdings to automatically destroy certain documents when they are 10 years old. Strummer Holdings reasonably anticipates litigation relating to events that occurred nine years ago such that its preservation obligations are triggered. If Strummer Holdings does not take steps to ensure that the relevant documents it has stored at Jones Storage (if any) are preserved, Strummer Holdings may be subject to sanctions for spoliation if any relevant documents are destroyed.

Guideline 7. Factors that may be considered in determining the scope of information that should be preserved include the nature of the issues raised in the matter, the accessibility of the information, the probative value of the information, and the relative burdens and costs of the preservation effort.

Executing preservation obligations typically involves an initial focus on documents and ESI available in accessible or "active" sources. Rule 26(f) provides parties in litigation with the opportunity at the "meet and confer" stage to discuss and evaluate potential discovery and agree on a reasonable preservation scope. The emphasis in the Rules is on cooperative action, as promoted by The Sedona Conference® *Cooperation Proclamation*.[54] Parties are admonished to pay particular attention "to [maintaining] the balance between the competing needs to preserve relevant evidence and to continue routine operations critical to ongoing activities."[55]

Unfortunately, it is not always feasible to secure prior agreement on preservation steps to be undertaken.[56] This is particularly true when preservation decisions must be made in the pre-litigation context, but it also is a problem after commencement of litigation. Parties are often in the position of having to make unilateral preservation decisions based on their best judgment.

There are numerous factors to be weighed in determining the scope of a particular hold. Some factors include the cost to preserve and potentially restore information; the number of individual custodians involved in the matter; the type of information involved; and whether the hold is on active data, historical data, or future data because the litigation

---

52 *Id.* at *6 & n. 9 (emphasizing that the proof is directed at the destruction of relevant evidence, not, per se, institution of a legal hold).

53 *Rimkus Consulting Grp. v. Cammarata*, 2010 U.S. Dist. LEXIS 14573 (S.D. Tex. Feb. 19, 2010); *cf. Pension Comm. v. Banc of Am. Sec., LLC,* 2010 WL 184312 (S.D.N.Y. Jan. 15, 2010).

54 *See* The Sedona Conference,® THE SEDONA CONFERENCE JOURNAL® SUPPLEMENT, 10 SEDONA CONF. J. 331 (Fall 2009) (calling for cooperative action by participants in relation to the discovery process).

55 FED. R. CIV. P. 26, Committee Note, Subdivision (f) (2006).

56 Kenneth J. Withers, *"Ephemeral Data" and the Duty To Preserve Discoverable Electronically Stored Information*, 37 U. BALT. L. REV. 349, 377 (Spring 2008) ("By the time the parties sit down at the Rule 26(f) conference, the preservation issues surrounding ephemeral data may be moot and the fate of the responding party may already be sealed, if sanctions are later found to be warranted.")

involves future or ongoing business activities. The court in *Zubulake IV* indicated that a "party or anticipated party must retain all relevant documents (but not multiple identical copies) in existence at the time the duty to preserve attaches, and any relevant documents created thereafter."[57] The court also explained that "[i]n recognition of the fact that there are many ways to manage electronic data, litigants are free to choose how this task is accomplished."[58]

Another key factor involves the accessibility of the information, especially when ESI is involved. While "[a] party's identification of sources of ESI as not reasonably accessible does not relieve the party of its common-law or statutory duties to preserve evidence,"[59] this observation should be read in conjunction with Rule 37(e), which provides that where data is lost as a result of good-faith, routine operations of electronic systems, no sanctions under the Federal Rules may be levied.[60] The Sedona Conference® *Commentary on Preservation, Management and Identification of Sources of Information That Are Not Reasonably Accessible*[61] suggests that in the absence of agreement, it is often "reasonable to decline to preserve" inaccessible sources if the party concludes that the "burdens and costs of preservation are disproportionate to the potential value of the source of data."[62]

For example, *Zubulake IV,* also concluded that "as a general rule,"[63] a "litigation hold does not apply to inaccessible backup tapes" which "may continue to be recycled." It also established an exception when the producing party could identify "the tapes storing the documents of 'key players'."[64] *The Sedona Principles* are in accord with this view.[65] Principle 8[66] also cautions against the assumption that there is an automatic need to preserve backup media. Thus, in *Escobar v. City of Houston*,[67] the fact that other relevant information had been preserved and was available mitigated concern about the failure to preserve audio tapes. Notably, the reasoning behind the general rule excluding inaccessible data (such as back up tapes) from preservation is not based simply on costs as the expense of saving a tape in isolation is relatively slight, but instead it is based upon a broader view of the need for preservation in the context of other sources of evidence and also balanced against the ultimate cost of later restoring data sources and culling them for particular content.

Likewise, transient or ephemeral data that is not kept in the ordinary course of business and that the organization may have no means to preserve may not need to be preserved. In *Columbia Pictures v. Bunnell*,[68] a court refused to find a duty to preserve information temporarily stored in RAM where the producing party had no reason to

57 220 F.R.D. 212, 218 (S.D.N.Y. 2003).
58 The court gave as an example, the retention of "all then-existing backup tapes for the relevant personnel (if such tapes store data by individual or the contents can be identified in good faith and through reasonable effort)," and noted that the party could "catalog any later-created documents in a separate electronic file." *Id.* at 218.
59 FED. R. CIV. P. 26, Committee Note, Subdivision (b)(2) (2006).
60 *Olson v. Sax,* 2010 WL 2639853 at *2 (E.D. Wis. June 25, 2010).
61 10 SEDONA CONF. J. 281 (2009) (in determining accessibility, a combination of "media based factors" and "data complexity factors" should be used); copy also available at http://www.thesedonaconference.org.
62 *Id.* (proposing a "decision tree" form of analysis under which the burdens and costs of accessing and preserving are balanced against the "reasonably anticipated need and significance of the information").
63 220 F.R.D. 212 at 217, n. 22.
64 220 F.R.D. 212 at 218, 220, n. 17 ("Litigants are now on notice, at least in this court, that [key player] backup tapes "must be preserved"). *See also Pension Comm. v. Bank of Am. Sec., LLC*, 2010 WL 184312 at *27 (S.D.N.Y. Jan. 15, 2010) ("I am not requiring that *all* backup tapes must be preserved. Rather, if such tapes are the *sole* source of relevant information (e.g., the active files of key players are no longer available), then such backup tapes should be segregated and preserved. When accessible data satisfies the requirement to search for and produce relevant information, there is no need to save or search backup tapes."); *Forest v. Caraco***,** 2009 WL 998402 (E.D. Mich. Apr. 14, 2009) (announcing proceedings limited to assessing *Zubulake* exception on delayed decision to cease recycling backup media).
65 Comment 5.h., THE SEDONA PRINCIPLES (2d ed. 2007) ("[a]bsent specific circumstances, preservation obligations should not extend to disaster recovery backup tapes created in the ordinary course of business.")
66 Comment 8.a., THE SEDONA PRINCIPLES (2d ed. 2007) ("The mere suspicion that a source may contain potentially relevant information is not sufficient to demand [its] preservation.")
67 2007 WL 2900581 (S.D. Tex. Sept. 29, 2007).
68 2007 WL 2080419 at *3-6 (C.D. Cal. May 29, 2007), *motion to review denied*, 245 F.R.D. 443 (2007).

anticipate that it would be sought and the requesting party first asserted a duty to preserve in a motion for sanctions.[69] Absent a special showing of need, Principle 9 of *The Sedona Principles* suggests that it is not ordinarily required to "preserve, review, or produce deleted, shadowed, fragmented, or residual [ESI]." Similarly, many organizations have made a good faith decision to not retain information such as instant messaging, chats, or voicemail messages in the ordinary course of business so that, absent compelling circumstance or an order of the court, there should be no expectation of preserving and producing information from such sources.

Parties sometimes seek to compel creation of a "mirror image" of hard drives to preserve data pending forensic examinations.[70] As part of the 2006 Amendments, the right to "test or sample" with reference to ESI was added to Rule 34(a). That amendment does not, however, create a "routine right of direct access" for such purposes.[71] Instead, such access is granted on a proper showing and perhaps with certain defined conditions.[72]

In some cases, parties may wish to affirmatively create "snapshots" of data as a defensive measure.[73] For example, the ability to access the hard drives of laptops issued to key employees upon their departure may be useful if it is the sole source of deleted information.[74]

If there are many custodians or if there is ongoing business information subject to the legal hold, *collecting* data at the outset of the legal hold may not be feasible. Sequestering the data can be disruptive to the business or technically unworkable in such circumstances. As a result, it is important to distinguish between preserving information and collecting and sequestering it.

If collecting data at an initial stage is not warranted, reasonable, or feasible, communications and monitoring processes become more important. It is critical that recipients of hold notices understand their duty to preserve information and how to meet that duty. Training sessions on legal hold compliance can be a useful tool to foster the effectiveness of legal holds.

Guideline 8. In circumstances where issuing a legal hold notice is appropriate, such a notice is most effective when the organization identifies the custodians and data stewards most likely to have relevant information, and when the notice:

- **(a) Communicates in a manner that assists persons in taking actions that are, in good faith, intended to be effective**
- **(b) Is in an appropriate form, which may be written**

---

69 The magistrate judge held that "the defendants failure to retain the server log data in RAM was based on a good faith belief that preservation of data temporarily stored only in RAM was not legally required" because, *inter alia*, there had been "no specific request by the defendants to preserve server log data present solely in RAM." *Id.* at *14.

70 *Bank of Mongolia v. M&P Global Fin. Servs.,* 258 F.R.D. 514 (S.D. Fla. 2009) (expert appointed to "retrieve any deleted responsive files" in light of production of responsive documents from third party sources).

71 FED. R. CIV. P. 34, Committee Note, Subdivision (a) (2006).

72 *See id., and Covad Communications v. Revonet, Inc.,* 258 F.R.D. 5 (D.D.C. 2009) where the court ordered forensic imaging of email servers for purposes of "preserving information as it currently exists."

73 It should be noted that forensic collection is not, nor should it be, the default method of collection and preservation. Instead, the duty to collect and preserve forensically only arises if: 1) the facts known to the preserving party or which the party should reasonably know would establish the need; or, 2) the requesting party has specifically requested it and the producing party has either agreed or notified the requesting party upon receiving the request that it will not comply, at which point the requesting party seeks judicial intervention an obtains an order compelling such preservation and collection. *See* Comment 8.c., THE SEDONA PRINCIPLES (2d ed. 2007) ("While [forensic data acquisition] is clearly appropriate in some circumstances, it should not be required unless exceptional circumstances warrant the extraordinary cost and burden;" also noting the need for careful protocols to address such collections).

74 *See, e.g., Cache La Poudre Feeds v. Land O'Lakes, supra,* 244 F.R.D. 614 (D. Colo. 2007) (failure to refrain from "expunging" hard drives of former key employees sanctioned where backup tapes were no longer available for use in seeking deleted email).

(c) Provides information on how preservation is to be undertaken
(d) Is periodically reviewed and, when necessary, reissued in either its original or an amended form, and
(e) Addresses features of relevant information systems that may prevent retention of potentially discoverable information.

When designing a legal hold it is particularly important that it be understandable by different groups within an organization. Counsel should review relevant pleadings or other documents and then describe the litigation in a way that will be understood by those with responsibility for preserving documents.

The initial and subsequent hold notices and reminders should describe the matter at issue, provide specific examples of the types of information at issue, identify potential sources of information, and inform recipients of their legal obligations to preserve information, and include reference to the potential consequences to the individual and the organization of noncompliance.[75] It should be in a form, which may include email, written hard copy or, in some cases, oral notice, which is appropriate to the circumstances. The notice should also inform recipients whom they should contact if they have questions or need additional information. Again, each case must be evaluated based on its own individual facts and a preservation notice adapted to conform to the facts and circumstances unique to that case.

Because of the distributed nature of ESI, it may be appropriate to communicate a legal hold notice not only to relevant data-generating or –receiving custodians, but also to appropriate data stewards, records management personnel, information technology (IT) personnel, and other potentially knowledgeable personnel.

Organizations should consider requiring confirmations of compliance with such hold notices as a means of verifying that recipients understand their preservation duties and obligations. *See* Guideline 10. Appropriate responses to hold notices and the organization's expectations for compliance with them should be included in organization's compliance programs.

Importantly, while the use of a written legal hold is often appropriate, it is simply one method of executing preservation obligations, not the only one. An organization should consider whether a written notice is necessary to effectively implement the hold and preserve the requisite information. In many instances, a written notice may not be necessary and, in fact, may be an encumbrance or source of confusion. Examples include situations in which sources of likely relevant information are subject to retention for sufficiently long periods pursuant to the organization's information management or record retention policy such that they will be held without a formal legal hold for the duration of the litigation. In addition, there may be situations in which sources of relevant information can be immediately secured without requiring preservation actions by employees. A read-only system of record for all pertinent research-and-development and product-quality information harnessed by a document management system would be one such example. There are other circumstances where the collection of information prior to any notice may be prudent in light of the risk that a custodian is the subject of the investigation or litigation and there is reason to believe that he or she might take steps to delete or destroy relevant information if aware of the circumstances.

75 The organization "must inform its officers and employees of the actual or anticipated litigation, and identify for them the kinds of documents that are thought to be relevant to it." *Samsung Electronics Co., Ltd., v. Rambus, Inc.*, 439 F. Supp. 2d 524, 565 (E.D. Va. 2006).

ILLUSTRATIONS

Illustration i: Lydon Enterprises obtains information that makes it reasonably anticipate litigation. Lydon Enterprises issues a written legal hold notice to certain of its employees. The document clearly identifies the recipients of the notice and explains in easily understandable terms which documents fall within the scope of the employees' preservation duties. The notice also explains how employees are expected to gather and preserve relevant documents. Whenever new information is obtained regarding the litigation that could affect the scope of the legal hold, in-house counsel for Lydon Enterprises reviews the notice. The notice is revised and reissued as necessary, and a periodic reminder is issued to all employees with preservation obligations. Compliance with the notice is regularly evaluated. This legal hold is likely to be considered effective or reasonable.

Illustration ii: Jones, Inc., obtains information that makes it reasonably anticipate litigation. In-house counsel for Jones identifies 40 people who she thinks might have relevant documents and instructs her secretary to call them and tell them to hold onto any documents relevant to the potential litigation, which she describes in general terms. The secretary calls the employees, but is unable to answer many of their questions. In-house counsel does not follow up on any of the employee questions. No written hold notice is ever issued. Litigation does not actually occur until 18 months later; at that point, in-house counsel begins collecting the relevant documents. This approach may or may not be effective, depending upon the circumstances, including the prejudice, if any, caused by the failure to issue a legal hold.

Illustration iii: Qualum Industries owns various properties, completes its financial accounting for 2008, and files its tax returns. Under its record retention policy and supporting schedules, tax-related papers are held for five years or until that tax year's audit is complete (whichever occurs later), and documentation supporting its financial reports are held for eight years. In 2010, Qualum was audited by the IRS, and questions were raised about Qualum's valuation of certain of its properties, but no litigation was filed. If Qualum reasonably concludes that the information needed to respond to questions during the audit are being retained pursuant to the company's information management and retention policy, Qualum need not issue a formal legal hold. If, however, litigation is later filed – either by the government or by Qualum for a refund after an adverse agency determination, and it is reasonably likely that information beyond the parameters of the retained records may be necessary to address claims or defenses in the action, Qualum would then be well-advised to issue a legal hold.

Guideline 9. An organization should consider documenting the legal hold policy and, when appropriate, the process of implementing the hold in a specific case, considering that both the policy and the process may be subject to scrutiny by opposing parties and review by the court.

An organization should consider documenting both the legal hold policy and, when appropriate, the steps taken to ensure the effective implementation of specific holds. Considering issues regarding work product and attorney-client privilege, the documentation need not disclose strategy or legal analysis. However, sufficient documentation should be included to demonstrate to opposing parties and the court that the legal hold was implemented in a reasonable, consistent, and good faith manner should there be a need to defend the process. In most cases, the process of issuing and implementing the legal hold and following up to preserve the data will provide sufficient documentation. If documentation of the legal hold process is deemed appropriate, it may include:

- The date and by whom the hold was initiated and possibly the triggering event
- The initial scope of information, custodians, sources, and systems involved
- Subsequent scope changes as new custodians or data are identified or initial sources are eliminated, and
- Notices and reminders sent, confirmations of compliance received (if any), and handling of exceptions.

In addition, in certain cases it may be appropriate to further document the process of how a specific legal hold was implemented. Examples may include:

- Description as to the collection protocol, persons contacted, and the date information was collected
- Notes (at least as to procedural matters) from any interviews conducted with employees to determine additional sources of information, and
- Master list of custodians, data stewards, and systems involved in the preservation effort.

While it may never be necessary to disclose this information, or disclosure may be made only to the court *in camera* to preserve privileged legal advice and work product information, the availability of documentation will preserve the option of the party to disclose the information in the event a challenge to the preservation efforts is raised and may provide a valuable resource when responding to discovery requests.

One reason to document the legal hold process and the implementation of it is to help avoid possible sanctions for the loss of relevant information. It can be very difficult for organizations to implement the legal hold and suspend or terminate routine operations of their large information systems to preserve relevant information before that information is deleted or overwritten in the normal course of operations.

Sanctions may be avoided under the Federal Rules if an organization can show that the information was lost by the routine operation of the information systems before a legal hold was instituted. Rule 37(e) provides that "absent exceptional circumstances, a court may not impose sanctions under these rules on a party for failing to provide ESI lost as a result of routine, good-faith operation of an electronic information system." Thus, while the Rule "does not set preservation obligations," it does tell judges that a spoliation claim involving ESI "cannot be analyzed in the same way as similar claims involving static information."[76]

Effective invocation of Rule 37(e) will require parties, as part of their legal hold implementation, to take good faith steps to suspend ordinary destruction processes or auto-

76 Lee H. Rosenthal, *A Few Thoughts on Electronic Discovery After December 1, 2006,* 116 YALE L. J. 167, 174 (Supp. 2006).

delete functionality of systems.[77] There is a split in authority, however, on the issue of whether the existence of a preservation obligation per se excludes the application of Rule 37(e).

Accordingly, effective use of Rule 37(e) places a premium on the use of the legal hold process, which may include the ability to communicate such holds promptly and repeatedly and to monitor compliance with them. Without a defined process, the safe harbor will be difficult to invoke and may offer little safety at all.

Guideline 10. Compliance with a legal hold should be regularly monitored.

Organizations should develop ways to regularly monitor a legal hold to ensure compliance. Some tools to accomplish this may include requiring ongoing certifications from custodians and data stewards, negative consequences for noncompliance, and audit and sampling procedures. Organizations may also consider employing technological tools, such as automated solutions and dedicated "legal hold" servers to facilitate and track employee compliance.

Organizations could also consider designating one or more individuals within the legal department to be responsible for issuing the legal hold notice, answering employee questions, and ensuring ongoing compliance with the notice. For smaller companies, outside counsel may be retained to perform this oversight function.

The effort to ensure compliance by affected employees is an ongoing process throughout litigation. This may include distributing periodic reminders of the legal hold and requiring employee confirmations, as well as issuing updated legal hold notices reflecting developments in the litigation itself or changes in the scope of the legal hold. As the number of custodians or other recommended recipients of the legal hold notice changes, it is important that the organization ensure that the expanded list of recommended recipients receives proper notification. Additional or revised notices should be promptly issued to persons who are added to the distribution.[78]

The argument is sometimes made that reliance on individuals to comply with preservation notices is unreasonable.[79] For example, a special master in a case involving a massive legal hold questioned the efficacy of preservation requirements that relied on recipients to move emails to avoid automatic deletion.[80] Another court expressed the view that "it is *not* sufficient to notify all employees of legal hold and expect that the party will then retain and produce all relevant information."[81] In *Pension Committee,*[82] the same court noted that "not every employee will require hands-on supervision from an attorney [but] attorney oversight of the process, including the ability to review, sample, or spot-check the collection efforts is important."[83]

---

77 *Compare, e.g., KCH Servs. v. Vanaire*, 2009 WL 2216601 at *1 (W.D. Ky. July 22, 2009) (oral instruction to delete software that might evidence violation of law "falls beyond the scope of 'routine, good faith operation'" of Rule 37(e)) *with Southeastern Mechanical Servs. v. Brody,* 2009 WL 2242395 at *3 (M.D. Fla. July 24, 2009) (declining to impose sanctions where losses covered were not intentionally caused in bad faith). *Cf.* remarks of Judge Shira Scheindlin, Panel Discussion, *Sanctions in Electronic Discovery Cases: Views from the Judges,* 78 FORDHAM L. REV.1, 30-31 (Oct. 2009) ("[Rule 37(e)] says if you don't put in a litigation hold when you should there's going to be no excuse if you lose information.").

78 This parallels Guideline 8, Illustration i, on communicating changes in the scope of the legal hold.

79 *Treppel v. Biovail ("Treppel V"),* 249 F.R.D. 111, 115-118 (S.D.N.Y. 2008) (noting inadequacies of mere notification to employees of a legal hold).

80 *In re Intel,* 258 F.R.D. 280 (D. Del. 2008).

81 *Zubulake V, supra,* 229 F.R.D. 422 at 432 (S.D.N.Y. 2004).

82 2010 WL 184312 (S.D.N.Y. Jan. 15, 2010).

83 *Id*. at n. 68.

However, in most cases, a careful combination of notification, collection, and individual action should enable parties to rely on the good faith actions of their employees. In an analogous context in *Concord Boat v. Brunswick*,[84] the court held that "[t]he fact that Defendant allowed individual employees to use discretion whether to retain e-mail is simply not indicative of bad faith."

If the legal hold applies to information created on a going-forward basis and pertains to a matter that represents substantial benefits or risks to an organization, the organization may wish to consider alternative means of auditing compliance. For example, the process could include a certification requirement that must be signed by the person responding to the legal hold. For holds involving ongoing business activities and future data, organizations may consider a periodic certification program to ensure ongoing compliance.

Guideline 11. Any legal hold policy, procedure, or practice should include provisions for releasing the hold upon the termination of the matter at issue so that the organization can adhere to policies for managing information through its useful lifecycle in the absence of a legal hold.

An organization creating a legal hold process should include procedures for releasing the holds once that organization is no longer obligated to preserve the information that was subject to a legal hold. These release procedures should include a process for conducting a custodian and data cross check so the organization can determine whether the information to be released is subject to any other ongoing preservation obligations. Organizations may consider using automation software that can perform custodian, system, and data cross checking and provide for efficient legal hold management.

When the organization is satisfied that the information is not subject to other preservation obligations, notice that the hold has been terminated should be provided to the recipients of the original notice (and any modifications or updated notices), and to records management, IT, and other relevant personnel, as well as any third parties notified of their obligation to preserve. Organizations may wish to conduct periodic audits to ensure that information no longer subject to preservation obligations is not unnecessarily retained and is being appropriately disposed of in accordance with the organization's records and information management policy.

84 1997 WL 33352759 at *6 (E.D. Ark. Aug. 29, 1997).

THE SEDONA CONFERENCE® COMMENTARY ON PROPORTIONALITY IN ELECTRONIC DISCOVERY

*A Project of The Sedona Conference® Working Group on Electronic Document Retention & Production (WG1)**

Author:
The Sedona Conference®

Editor-in-Chief:
Conor R. Crowley

Senior Editors:
Sean R. Gallagher
Gina Trimarco
Jennifer L. Young

*Contributing Editors:**
Kevin F. Brady
Peg Duncan
Ronald J. Hedges
Peter Pepiton, II

The opinions expressed in this publication, unless otherwise attributed, represent consensus views of the members of The Sedona Conference® Working Group 1. They do not necessarily represent the views of any of the individual participants or their employers, clients, or any other organizations to which any of the participants belong, nor do they necessarily represent official positions of The Sedona Conference®.

REPRINT REQUESTS:
Requests for reprints or reprint information should be directed to Richard Braman, Executive Director of The Sedona Conference®, at tsc@sedonaconference.org or 1-866-860-6600.

* The Sedona Conference® also wishes to acknowledge the contributions of Gil S. Keteltas, David Walton, and all the Working Group 1 members who contributed to the dialogue while this Commentary was a work-in-progress. Finally, we wish to thank all of our Working Group Series℠ Sponsors, whose support is essential to The Sedona Conference®'s ability to develop Working Group Series℠ publications such as this. For a complete listing of our Sponsors, click on the "Sponsors" navigation bar on the home page of our website at www.thesedonaconference.org.

Preface

Welcome to another major publication in The Sedona Conference® Working Group Series (WGS℠): *The Sedona Conference® Commentary on Proportionality in Electronic Discovery.*

This effort is a product of our Working Group in Electronic Document Retention and Production (WG1) and represents the collective expertise of a diverse group of lawyers and representatives of firms providing consulting and legal services. Not only has a diverse drafting team been responsible for the writing, it has been the subject of dialogue at three meetings of WG1, and numerous WG1 members have contributed comments and edits outside of the context of our meetings.

On behalf of The Sedona Conference®, I want to thank the editorial team and all WG1 members whose dialogue and comments contributed to this Commentary for all of their efforts to draft a commentary of immediate benefit to the bench and bar.

As with all of our WGS℠ publications, this Commentary is first being published as a "public comment version." After sufficient time for public comment has passed, the editors will review the public comments, and to the extent appropriate, edit the current version. The Commentary will then be re-published in a "final" version, as always, subject to future developments in the law that may warrant a second edition.

We hope our efforts will assist lawyers, judges, and others involved in the legal system work with the concept of proportionality in discovery. If you wish to submit any suggested edits or comments, please utilize the "public comment form" on the download page of our website at www.thesedonaconference.org. You can also submit feedback by emailing us at rgb@sedonaconference.org.

Richard G. Braman
Executive Director
The Sedona Conference®
August 2010

The Sedona Conference® Principles of Proportionality

1. The burdens and costs of preservation of potentially relevant information should be weighed against the potential value and uniqueness of the information when determining the appropriate scope of preservation.

2. Discovery should generally be obtained from the most convenient, least burdensome, and least expensive sources.

3. Undue burden, expense, or delay resulting from a party's action or inaction should be weighed against that party.

4. Extrinsic information and sampling may assist in the analysis of whether requested discovery is sufficiently important to warrant the potential burden or expense of its production.

5. Nonmonetary factors should be considered when evaluating the burdens and benefits of discovery.

6. Technologies to reduce cost and burden should be considered in the proportionality analysis.

Introduction

This Commentary discusses the origins of the doctrine of proportionality, provides examples of its application, and proposes principles to guide courts, attorneys, and parties. The principles do not merely recite existing rules and case law but rather provide a framework for the application of the doctrine of proportionality to all aspects of electronic discovery. Although the Commentary cites primarily federal case law and rules, the principles are equally applicable to electronic discovery in state courts.

In 1938, the Federal Rules of Civil Procedure ("Federal Rules") were adopted, providing for "the just, speedy, and inexpensive determination of every action." FED. R. CIV. P. 1 (1938). Over the years, the Federal Rules have witnessed various technological revolutions, such as the "modern miracle of photographic reproduction,"[1] which one court noted "lessen[s] what might otherwise be burdensome transportation of records and documents."[2]

Since their enactment in 1938, the Federal Rules have been amended several times to keep pace with the changing demands of courts and parties. In 1983, Rule 26(b)(1) was amended to grant courts the authority to limit discovery where it was found to be redundant or duplicative. The Advisory Committee noted that the amendments were intended to "guard against redundant or disproportionate discovery by giving the court authority to reduce the amount of discovery that may be directed to matters that are otherwise proper subjects of inquiry."[3] This was important, because "[e]xcessive discovery and evasion or resistance to reasonable discovery requests pose significant problems."[4] As explained by the Advisory Committee, the amendments were designed to address the problem of disproportionate discovery, listing factors to be considered in determining whether discovery is proportional: the nature and complexity of the lawsuit, the importance of the issues at stake, the parties' resources, the significance of the substantive issues, and public policy concerns.[5]

The Federal Rules were amended again in 1993 to address the tremendous increase in the amount of potentially discoverable information caused by the "information explosion of recent decades" and the corresponding increase in discovery costs with the addition of a new paragraph, Rule 26(b)(2).[6] Rule 26(b)(2) provided courts with even greater discretion to limit the scope and extent of discovery. The Advisory Committee Notes explain that "[t]he revisions in Rule 26(b)(2) are intended to provide the court with broader discretion to impose additional restrictions on the scope and extent of discovery . . . ."[7]

In 2006, Rule 26(b)(2) was amended to limit the discovery of ESI deemed not reasonably accessible by reason of the costs and burdens associated with retrieving such information. The Advisory Committee Notes to this amendment state that the costs and burdens of retrieving not reasonably accessible information are properly considered as part of the proportionality analysis, and that discovery of such information may be limited or the costs of such discovery shifted from the responding to the requesting party.[8]

1 *Herbst v. Able*, 278 F. Supp. 664, 667 (S.D.N.Y. 1967).
2 *Goldberg v. Taylor Wine Co.*, No. 77-1548, 1978 U.S. Dist. LEXIS 19891, at *14 (E.D.N.Y. Jan. 27, 1978).
3 Advisory Committee Notes to 1983 Amendments to FED. R. CIV. P. 26(b).
4 Advisory Committee Notes to 1983 Amendments to FED. R. CIV. P. 26.
5 Advisory Committee Notes to 1983 Amendments to FED. R. CIV. P. 26(b).
6 Advisory Committee Notes to 1993 Amendments to FED. R. CIV. P. 26(b).
7 *Id.*
8 Advisory Committee Notes to 2006 Amendments to FED. R. CIV. P. 26(b)(2).

Notwithstanding the foregoing amendments, courts have not always applied proportionality in circumstances when its application was warranted.[9] When courts have applied proportionality, they have not always described it as such.[10] In the electronic era, it has become increasingly important for courts *and* parties to apply the proportionality doctrine to manage the large volume of ESI and associated expenses now typical in litigation. Below is a discussion of the key issues to consider when conducting a proportionality analysis pursuant to Rule 26(b)(2)(C), followed by a discussion of recommended principles.

The Availability of the Information From Other Sources

Rule 26(b)(2)(C)(i) provides that courts must impose limitations where "the discovery sought is unreasonably cumulative or duplicative, or can be obtained from some other source that is more convenient, less burdensome, or less expensive." FED. R. CIV. P. 26(b)(2)(c)(i). Where relevant information is available from multiple sources, this rule allows courts to limit discovery to the least expensive source.[11]

Waiver and Undue Delay

Rule 26(b)(2)(C)(ii) provides that courts must limit discovery where "the party seeking discovery has had ample opportunity to obtain the information by discovery in the action." FED. R. CIV. P. 26(b)(2)(c)(ii). This provision allows courts to prevent protracted discovery by applying the concept of waiver to reject untimely discovery requests and objections.[12]

The rule allows courts to ensure proportionality in the discovery process by denying as untimely discovery requests that should have been propounded earlier in the litigation or objections interposed too late. Pursuant to this provision, parties' discovery requests, and any corresponding objections, must be reasonably prompt or they may be deemed waived.

---

9 *See* Advisory Committee Notes to 2000 Amendments to FED. R. CIV. P. 26(b)(1) ("The Committee has been told repeatedly that courts have not implemented these limitations with the vigor that was contemplated."); CHARLES A. WRIGHT & ARTHUR R. MILLER, FEDERAL PRACTICE & PROCEDURE § 2008.1, at 121 (2d ed. 1994) (describing the "paucity of reported cases" applying the proportionality rule and concluding "that no radical shift has occurred").

10 *See, e.g., Waldron v. Cities Serv. Co.*, 361 F.2d 671, 673 (2d Cir. 1966), *aff'd*, 391 U.S. 253 (1968) ("The plaintiff . . . may not seek indefinitely . . . to use the [discovery] process to find evidence in support of a mere 'hunch' or 'suspicion' of a cause of action."); *Jones v. Metzger Dairies, Inc.*, 334 F.2d 919, 925 (5th Cir. 1964) ("Full and complete discovery should be practiced and allowed, but its processes must be kept within workable bounds on a proper and logical basis for the determination of the relevancy of that which is sought to be discovered."); *Dolgow v. Anderson*, 53 F.R.D. 661, 664 (E.D.N.Y. 1971) ("A trial court has a duty, of special significance in lengthy and complex cases where the possibility of abuse is always present, to supervise and limit discovery to protect parties and witnesses from annoyance and excessive expense."); *Welty v. Clute*, 1 F.R.D. 446 (W.D.N.Y. 1940) (finding a second deposition of a plaintiff unnecessary given the availability of other discovery).

11 For example, in *Young v. Pleasant Valley School District*, the court rejected the plaintiffs' request for production of emails located on back-up tapes, citing Rule 26(b)(2)(C)(i), and noting that "[t]he burden and expense of rebuilding the district's email system in order to provide the discovery requested by the plaintiffs, along with the additional and less expensive means available for plaintiffs to get this material[,] makes the plaintiffs' discovery request impractical." *Young v. Pleasant Valley Sch. Dist.*, No. 07-854, 2008 U.S. Dist. LEXIS 55585, at *6-7 (M.D. Pa. July 21, 2008).

12 *Ford Motor Co. v. Edgewood Props.*, 257 F.R.D. 418, 426 (D.N.J. 2009) ("One may reasonably expect that if document production is proceeding on a rolling basis where the temporal gap in production is almost half a year apart, a receiving party will have reviewed the first production for adequacy and compliance issues for a reason as obvious as to ensure that the next production of documents will be in conformity with the first production or need to be altered. It was incumbent on Edgewood to review the adequacy of the first production so as to preserve any objections. The Court is not dictating a rigid formulation as to when a party must object to a document production. Reasonableness is the touchstone principle, as it is with most discovery obligations. The simple holding here is that it was unreasonable to wait eight months after which production was virtually complete.").

Burden Versus Benefit

In assessing whether to limit discovery, courts consider whether "the burden or expense of the proposed discovery outweighs its likely benefit, considering the needs of the case, the amount in controversy, the parties' resources, the importance of the issues at stake in the action, and the importance of the discovery in resolving the issues." FED. R. CIV. P. 26(b)(2)(C)(iii). In analyzing these factors, courts weigh the burdens of discovery against the potential benefit of the information to be produced, in light of the specific circumstances of the case. For example, a court may order a party to engage in a burdensome and costly production if the information sought is extremely relevant to the case and is unavailable elsewhere. But what if the cost of producing the information exceeds the total value of the case? Or, what if expensive production is warranted given the value of the case, but the producing party lacks the resources to pay for the production? What if the amount in controversy is low, but the case raises important societal issues? These are all issues a court may consider in assessing whether to limit discovery.[13]

Although Rule 26(b)(2)(C)(iii) discusses a number of monetary considerations, courts may likewise analyze whether nonmonetary concerns, such as the societal benefit of the case or the nonmonetary burden on the producing party, weigh in favor of limiting discovery. The "metrics" set forth in Rule 26(b)(2)(C)(iii) provide courts significant flexibility and discretion to assess the circumstances of the case and limit discovery accordingly to ensure that the scope and duration of discovery is reasonably proportional to the value of the requested information, the needs of the case, and the parties' resources.

Most courts that have addressed proportionality have focused on Rules 26(b) and (c). However, Rule 26(g) also requires parties propounding or responding to discovery requests to conduct a proportionality analysis. Described by one court as "[o]ne of the most important, but apparently least understood or followed, of the discovery rules,"[14] Rule 26(g)(1) provides that

> [e]very discovery request, response, or objection must be signed by at least one attorney of record in the attorney's own name . . . . By signing, an attorney or party certifies that to the best of the person's knowledge, information, and belief formed after a reasonable inquiry . . . with respect to a discovery request, response, or objection, it is . . . neither unreasonable nor unduly burdensome or expensive, considering the needs of the case, prior discovery in the case, the amount in controversy, and the importance of the issues at stake in the action.

13 *See, e.g.*, *Young v. Pleasant Valley Sch. Dist.*, 2008 U.S. Dist. LEXIS 55585, at *6-7; *see also Spieker v. Quest Cherokee, LLC*, No. 07-1225, 2008 U.S. Dist. LEXIS 88103, at *10-11 (D. Kan. Oct. 30, 2008) (assessing a request to limit discovery in a class action, rejecting "defendant's argument that the 'amount in controversy' is limited to the named plaintiffs' claims" and stating that "defendant's simplistic formula for comparing the named plaintiffs' claims with the cost of production is rejected"); *Southern Capitol Enters., Inc. v. Conseco Servs., L.L.C.*, No. 04-705, 2008 U.S. Dist. LEXIS 87618, at *7 (M.D. La. Oct. 24, 2008) ("Perfection in document production is not required. . . . In these circumstances Rule 26(b)(2)(C)(iii) comes into play. At this point in the litigation, the likely benefit that could be obtained from [further discovery] is outweighed by the burden and expense of requiring the defendants to renew their attempts to retrieve the electronic data."); *Metavante Corp. v. Emigrant Sav. Bank*, No. 05-1221, 2008 U.S. Dist. LEXIS 89584, at *10 (E.D. Wis. Oct. 24, 2008) ("In viewing the totality of the circumstances, including the amount in controversy in this case, the parties' resources, and the issues at stake, the court concludes that the burden [of production] does not outweigh the value of the material sought."); *Mancia v. Mayflower Textile Servs. Co.*, 253 F.R.D. 354, 364 (D. Md. 2008) ("I noted during the hearing that I had concerns that the discovery sought by the Plaintiffs might be excessive or overly burdensome, given the nature of this FLSA and wage and hour case, the few number of named Plaintiffs and the relatively modest amounts of wages claimed for each."); *Cenveo Corp. v. Slater*, No. 06-2632, 2007 U.S. Dist. LEXIS 8281, at *4 (E.D. Pa. Jan. 31, 2007) ("The dispute before the Court requires a weighing of defendants' burden in producing the information sought against plaintiff's interest in access to that information. Because of the close relationship between plaintiff's claims and defendants' computer equipment, the Court will allow plaintiff to select an expert to oversee the imaging of all of defendants' computer equipment.").

14 *Mancia v. Mayflower Textile Servs. Co.*, 253 F.R.D. 354, 357 (D. Md. 2008).

FED. R. CIV. P. 26(g)(1). The Advisory Committee acknowledged that the rule is intended to impose "an affirmative duty to engage in pretrial discovery in a responsible manner that is consistent with the spirit and purposes of Rules 26 through 37."[15] To that end, the rule "provides a deterrent to both excessive discovery and evasion by imposing a certification requirement that obliges each attorney to stop and think about the legitimacy of a discovery request, a response thereto, or an objection."[16] Indeed, the Advisory Committee noted that "the premise of Rule 26(g) is that imposing sanctions on attorneys who fail to meet the rule's standards will significantly reduce abuse by imposing disadvantages therefor."[17]

In sum, courts applying the doctrine of proportionality may consider a variety of factors, including the benefit of the proposed discovery (including nonmonetary considerations), the burden or expense of the proposed discovery, the availability of the information from other sources, and undue delay on the part of the party seeking discovery.

We recognize that some parties may inappropriately raise proportionality arguments, either as a sword to increase the burden on the producing party,[18] or as a shield to avoid legitimate discovery obligations.[19] Courts must be wary of such abuses. In any event, the burden or expense associated is simply one factor to be considered in a proportionality analysis and may not be dispositive or even determinative in specific cases.

---

15 Advisory Committee Notes to 1983 Amendments to FED. R. CIV. P. 26(g).

16 *Id.*

17 *Id.*

18 *See, e.g.*, *Kay Beer Distrib., Inc. v. Energy Brands, Inc.*, No. 07-1068, 2009 WL 1649592 (E.D. Wis. June 10, 2009). In this matter involving the terms of a distribution agreement, the plaintiff moved to compel a native production of five DVDs containing the defendant's emails and other ESI for a five-year period. The court denied the motion, holding that "[Defendant] has no obligation to turn over to an opposing party in a lawsuit non-discoverable and privileged information. . . . The mere possibility of locating some needle in the haystack of ESI . . . does not warrant the expense [defendant] would incur in reviewing it." *Id.* at *4.

19 *See Kipperman v. Onex Corp.,* 260 F.R.D. 682 (N.D. Ga. 2009). The court noted that "rather than seeking a protective order [the defendants] determined themselves that it would be overly burdensome" to produce the discovery in the court-ordered format. *Id.* at 693. The court sanctioned the defendants under Rules 26 and 37, for their failure to follow discovery orders, their lack of diligence in discovery, and "making blatant misrepresentations about the value of email discovery in this case in an effort to influence the court's ruling." *Id.* at 692.

The Sedona Conference® Principles of Proportionality

Principle 1: The burdens and costs of preservation of potentially relevant information should be weighed against the potential value and uniqueness of the information when determining the appropriate scope of preservation.

The Federal Rules do not apply until litigation has commenced. However, courts may invoke their inherent authority to sanction parties for pre-litigation preservation failures. The proportionality principles set forth in the Federal Rules, while not directly applicable, may serve as a guide to parties considering pre-litigation preservation obligations.[20]

Thus, parties for whom an obligation to preserve potentially relevant information has arisen should weigh the burdens and costs of preservation against the potential value and uniqueness of the information when determining the appropriate scope and manner of such preservation.[21]

Courts conducting a post hoc analysis of a party's pre-litigation preservation decisions should evaluate the decisions based on the proportionality factors set forth in Rule 26(b)(2)(C) and the preserving party's good faith and reasonableness considering the knowledge available at the time of preservation.[22] Although there is no case law that applies the proportionality factors set forth in the Federal Rules in the pre-litigation context, parties who demonstrate that they acted thoughtfully, reasonably, and in good faith in preserving or attempting to preserve information prior to litigation should generally be entitled to a presumption of adequate preservation. However, parties must be prepared to make this demonstration and cannot rely on "pure heart, empty head." Courts should not allow hindsight bias to color the analysis of a party's deliberate, reasonable, and good faith preservation efforts.[23]

Principle 2: Discovery should generally be obtained from the most convenient, least burdensome, and least expensive sources.

Rule 26(b)(2)(C)(i) provides that courts must limit discovery when the requested material can be obtained from sources that are "more convenient, less burdensome, or less expensive." FED. R. CIV. P. 26(b)(2)(C)(i). In other words, where relevant information is available from multiple sources, this provision grants courts the authority to limit discovery

20 *See, e.g.*, THE SEDONA CONFERENCE®, THE SEDONA PRINCIPLES (2d ed. 2007), Principle 5 ("The obligation to preserve electronically stored information requires reasonable and good faith efforts to retain information that may be relevant to pending or threatened litigation. However, it is unreasonable to expect parties to take every conceivable step to preserve all potentially relevant electronically stored information."); THE SEDONA CONFERENCE®, THE SEDONA CONFERENCE® COMMENTARY ON PRESERVATION, MANAGEMENT AND IDENTIFICATION OF SOURCES OF INFORMATION THAT ARE NOT REASONABLY ACCESSIBLE 14 (July 2008) ("If the burdens and costs of preservation are disproportionate to the potential value of the source of data at issue, it is reasonable to decline to preserve the source."); THE SEDONA CONFERENCE®, THE SEDONA CONFERENCE® COMMENTARY ON INACTIVE INFORMATION SOURCES 11 (July 2009) ("A reasonableness/proportionality analysis should be conducted to determine whether it would be reasonable under the circumstances to take steps to preserve a specific inactive information store . . . ."); *see also* The Hon. Paul W. Grimm, et al., *Proportionality in the Post-Hoc Analysis of Pre-Litigation Preservation Decisions*, 37 U. BALT. L. REV. 381 (2008) (urging for "the application, by analogy, of Federal Rules of Civil Procedure 26(b)(2)(C) and 37(e) to the pre-litigation duty to preserve").

21 The determination of whether a preservation obligation has arisen is addressed in THE SEDONA CONFERENCE® COMMENTARY ON LEGAL HOLDS: THE TRIGGER & THE PROCESS (2007). This principle addresses the appropriate scope and manner of preservation after the determination has been made that a preservation obligation exists.

22 *Rimkus Consulting Group, Inc. v. Cammarata*, 688 F. Supp. 2d 598, 613 (S.D. Tex. 2010) ("Whether preservation or discovery conduct is acceptable in a case depends on what is *reasonable*, and that in turn depends on whether what was done – or not done – was *proportional* to that case and consistent with clearly established applicable standards.").

23 *Cf. Rhoads Indus., Inc. v. Bldg. Materials Corp. of Am.*, 254 F.R.D. 216, 226 (E.D. Pa. 2008) ("[H]indsight . . . should not carry much weight, if any, because no matter what methods an attorney employed, an after-the-fact critique can always conclude that a better job could have been done.").

to the least burdensome source, thus empowering courts to control litigation costs and promote efficiency in accordance with Rule 1. *See* FED. R. CIV. P. 1. Likewise, this provision enables courts to protect parties from abusive discovery requests. Although any one source is unlikely to meet the requirement of being the most convenient, least burdensome, and least expensive source, parties should carefully weigh these factors when determining which source is optimal.

For example, a court may consider limiting discovery of not reasonably accessible[24] back-up tapes if the information stored on the tapes can be obtained from other more accessible sources, such as through hard copy records, testimony, or nonparty discovery. If the producing party can easily produce hard copies of the emails, why should that party incur the costs of restoring back-up tapes containing the same emails?[25]

In determining whether to limit purportedly burdensome or expensive discovery pursuant to Rule 26(b)(2)(C)(i), a court's analysis should be tailored to the specific facts at issue, taking into account the various sources in which the requested information is located, the burden and expense of production from those sources, and whether limiting discovery to less burdensome or expensive sources will result in a reduction in the utility of the information sought. For example, hard copy emails may be more accessible and easily produced but because they are not in electronic form, the requesting party must incur the costs of scanning the hard copies and loading them onto a search platform or conducting a manual search. In this situation, it may be appropriate for a court to consider the totality of litigation costs, and who should bear certain of those costs, in assessing a request to limit production.

Depending on the facts available and the stage of the litigation, the parties and the court may be unable to assess whether limiting discovery is appropriate. If the litigation is in its early stages, the parties may not yet be fully aware of all of the viable claims and defenses or factual or legal issues that will ultimately be important to the outcome of the litigation. Similarly, if a requesting party seeks production of ESI archived several years ago, the responding party may not have a full understanding of the content of the ESI or its potential value to the litigation.[26]

Under these circumstances, the court, or the parties on their own initiative, may find it appropriate to conduct discovery in phases, starting with discovery of clearly relevant information located in the most accessible and least expensive sources. Phasing discovery in this manner may allow the parties to develop the facts of the case sufficiently to determine whether, at a later date, further potentially more burdensome and expensive discovery is necessary or warranted. In addition, given that the vast majority of cases settle, phasing discovery may allow the parties to develop a sufficient factual record upon which to base settlement negotiations without incurring the costs of more burdensome discovery that may only be necessary if the case goes to trial.

---

24 Rule 26(b)(2)(B) provides that "[a] party need not provide discovery of electronically stored information from sources that the party identifies as not reasonably accessible because of undue burden or cost. On motion to compel discovery or for a protective order, the party from whom discovery is sought must show that the information is not reasonably accessible because of undue burden or cost. If that showing is made, the court may nonetheless order discovery from such sources if the requesting party shows good cause, considering the limitations of Rule 26(b)(2)(C). The court may specify conditions for the discovery." FED. R. CIV. P. 26(b)(2)(B).

25 *Cf. U.S. v. O'Keefe*, 2008 U.S. Dist. LEXIS 12220 at *22 (D.D.C. Feb. 18, 2008).

26 *See* Advisory Committee Notes to 2006 Amendments to Rule 26(b)(2) (noting that, "because the court and parties may know little about what information the sources identified as not reasonably accessible might contain, whether it is relevant, or how valuable it may be to the litigation," it may be appropriate for the parties to engage in "focused discovery . . . to learn more about what burdens and costs are involved in accessing the information, what the information consists of, and how valuable it is for the litigation in light of information that can be obtained by exhausting other opportunities for discovery.").

Parties who wish to conduct phased discovery must communicate with one another about the issues relevant to the litigation and the potential repositories – both accessible and inaccessible – of relevant information. Moreover, the parties must cooperate with one another to prepare and propose to the court a phased discovery plan.[27]

Principle 3: Undue burden, expense, or delay resulting from a party's action or inaction should be weighed against that party.

Although the Federal Rules do not set forth specific deadlines by which parties must propound or otherwise sequence discovery, courts will often set discovery deadlines in scheduling orders or based on local rules. Courts may also sequence fact and expert discovery, set specific dates for completion of document production, or limit the time period in which parties can raise discovery disputes. From a proportionality perspective, propounding discovery requests at the early stages of the litigation allows parties sufficient time to explore compliance with the discovery requests and bring any disputes before the court for resolution. Accordingly, parties should raise any discovery disputes as soon as is reasonably possible, but only after engaging in good faith attempts to resolve the dispute without the court's involvement.

Where a dispute cannot be resolved, it should be raised with the court promptly. In determining whether the requested relief is appropriate, the court may consider the time at which the issue arose and whether the moving party could have raised the issue earlier. The resolution of such disputes can be fact-intensive, requiring the court to assess whether the producing party complied with its discovery obligations, the degree of culpability involved, and the prejudice to the requesting party.

Traditionally, a party must bear the costs of responding to discovery requests, including the costs of production. In assessing whether a particular discovery request or requirement is unduly burdensome or expensive, a court should consider the extent to which the claimed burden and expense grow out of the responding party's action or inaction. In practice, this principle typically focuses on actions taken, or not taken, by the responding party with regard to the duties to preserve, search, and produce relevant information.[28]

A failure to preserve relevant information in an accessible format at the outset of litigation should be weighed against a party seeking to avoid the resultant burden of restoring the information. The Advisory Committee noted that an "appropriate consideration" in assessing burden and expense in the context of claims that information is not reasonably accessible is "the failure to produce relevant information that seems likely to have existed but is no longer available on more easily accessed sources."[29]

This proportionality principle also applies to a party's failure to engage in early, meaningful discussions with an opponent to develop a discovery plan and avoid potential

27 These issues may be considered at the Rule 26(f) conference, at which the parties must "discuss any issues about preserving discoverable information; and develop a proposed discovery plan." FED. R. CIV. P. 26(f)(2).

28 *Quinby v. WestLB AG*, 245 F.R.D. 94, 104-05 (S.D.N.Y. 2006) ("[I]f a party creates its own burden or expense by converting into an inaccessible format data that it should have reasonably foreseen would be discoverable material at a time when it should have anticipated litigation, then it should not be entitled to shift the costs of restoring and searching the data. This would permit parties to maintain data in whatever format they choose, but discourage them from converting evidence to inaccessible formats because the party responsible for the conversion will bear the cost of restoring the data. Furthermore, it would prevent parties from taking unfair advantage of a self-inflicted burden by shifting part of the costs of undoing the burden to an adversary. If, on the other hand, it is not reasonably foreseeable that the particular evidence in issue will have to be produced, the responding party who converts the evidence into an inaccessible format after the duty to preserve evidence arose may still seek to shift the costs associated with restoring and searching for relevant evidence.").

29 Advisory Committee Notes to 2006 Amendments to Rule 26(b)(2).

disputes. Application of the principle in this context is appropriate, given that a party can be sanctioned for failing "to participate in good faith in developing and submitting a proposed discovery plan as required by Rule 26(f)."[30] Examples of how a party's failure to engage in an early and meaningful meet and confer may weigh against the party in a subsequent proportionality analysis include cases in which: (1) a party refuses to consult with an opponent in the development of a keyword search protocol and a second search is necessary given the inadequacy of keywords initially applied; (2) a second production of material already produced is necessary because a party fails to confer on form of production and produces ESI in a form that is not reasonably usable.

Principle 4: Extrinsic information and sampling may assist in the analysis of whether requested discovery is sufficiently important to warrant the potential burden or expense of its production.

Rule 26(b)(2)(C)(iii) provides that in considering whether to limit potentially burdensome or expensive discovery, courts should consider "the importance of the discovery in resolving the issues." FED. R. CIV. P. 26(b)(2)(C)(iii). In other words, the court may limit discovery if the information sought, while relevant, is not sufficiently important to warrant the burden and expense of its production.[31] This issue often arises in the context of requests for discovery of information that is duplicative, cumulative, or not reasonably accessible.[32] *See* FED. R. CIV. P. 26(b)(2)(B) (incorporating factors set forth in Federal Rule 26(b)(2)(C)).

When asked to limit discovery on the basis of burden or expense, courts must make an assessment of the importance of the information sought. Discovery should be limited if the burden or expense of producing the requested information is disproportionate to its importance to the litigation. Performing such an assessment can be challenging, given that it may be impossible to review the content of the requested information until it is produced.[33]

In some cases, it may be clear that the information requested is important—perhaps even outcome-determinative.[34] In other cases, courts order sampling of the requested information, consider extrinsic evidence, or both, to determine whether the requested information is sufficiently important to warrant potentially burdensome or expensive discovery.[35]

In *Kipperman v. Onex Corporation*, the court required the defendants to produce two "sample" backup tapes so that the court could gauge the volume and importance of the

30 FED. R. CIV. P. 37(f).

31 An alternative to limiting burdensome or expensive discovery is to shift its cost to the requesting party. *See* FED. R. CIV. P. 26(c); *see also Rowe Entm't, Inc. v. William Morris Agency, Inc.*, 205 F.R.D. 421, 428 (S.D.N.Y. 2002) ("[T]here is no justification for a blanket order precluding discovery of the defendants' e-mails on the ground that such discovery is unlikely to provide relevant information… The more difficult issue is the extent to which each party should pay the costs of production."); *McPeek v. Ashcroft*, 202 F.R.D. 31, 34 (D.D.C. 2001) ("The converse solution is to make the party seeking the restoration of the backup tapes pay for them, so that the requesting party literally gets what it pays for.").

32 Courts may also employ sampling for the purpose of evaluating a request to shift costs. *See Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309, 324 (S.D.N.Y. 2003) ("Requiring the responding party to restore and produce responsive documents from a small sample of backup tapes will inform the cost-shifting analysis.").

33 Advisory Committee Notes to 2006 Amendments to Rule 26(b)(2) ("The good-cause determination, however, may be complicated because the court and parties may know little about what information the sources identified as not reasonably accessible might contain, whether it is relevant, or how valuable it may be to the litigation."); *see also Peskoff v. Faber*, 244 F.R.D. 54, 59 (D.D.C. 2007) ("Application of this factor can be challenging because the importance of the results of the forensic examination can only be assessed after it is done.").

34 *See Covad Communs. Co. v. Revonet, Inc.*, 258 F.R.D. 5, 13 (D.D.C. 2009) (permitting discovery that "should establish once and for all" a key issue in the case).

35 Advisory Committee Notes to 2006 Amendments to Rule 26(b)(2) ("[T]he parties may need some focused discovery, which may include sampling of the sources, to learn more about what burdens and costs are involved in accessing the information, what the information consists of, and how valuable it is for the litigation in light of information that can be obtained by exhausting other opportunities for discovery.").

information located on the tapes against the costs of their restoration and production before deciding whether to permit additional discovery.[36] After reviewing the results of the sample, the court determined that the information contained on the backup tapes was sufficiently important to warrant further discovery: "I don't . . . declare these to be smoking guns but they certainly are hot and they certainly do smell like they have been discharged lately."[37]

In addition to sampling, courts generally consider extrinsic information submitted by the parties to determine whether requested discovery is sufficiently important to warrant potentially burdensome or expensive discovery. Such evidence may include: the parties' opinions regarding the likely importance of the requested information,[38] whether the requested information was created by "key players,"[39] whether discovery already produced permits an inference that the requested information is likely to be important,[40] whether the creation of the information requested was contemporaneous with key facts in the case,[41] and whether the information requested is unique.[42] The analysis of the importance of the requested information is fact-specific and thus will vary from case to case.

The party opposing discovery may put forth evidence that the burden or expense of producing the requested information outweighs its potential importance.

Principle 5: Nonmonetary factors should be considered when evaluating the burdens and benefits of discovery.

The Federal Rules recognize that proportionality encompasses nonmonetary considerations. Rule 26(g)(1)(B)(iii) requires that an attorney (or *pro se* party) who promulgates discovery must consider "the needs of the case, prior discovery in the case, the amount in controversy, and the importance of the issues at stake in the action." Rule 26(b)(2)(C)(iii) similarly requires that a court consider "the importance of the issues at stake in the action" in assessing whether to limit discovery. FED. R. CIV. P. 26(b)(2)(C)(iii).

The Advisory Committee Note to Section 26(b)(2)(C)(iii) states:

> The elements of Rule 26(b)(1)(iii) address the problem of discovery that is disproportionate to the individual lawsuit as measured by such matters as its nature and complexity, the importance of the issues at stake in a case seeking damages, the limitations on a financially weak litigant to withstand extensive opposition to a discovery program or to respond to discovery requests, and the significance of the substantive issues, as measured in philosophic, social, or institutional terms. Thus the rule recognizes that many cases in public policy spheres, such as employment practices, free speech, and other matters, may have importance far beyond the monetary amount involved.[43]

36 *Kipperman v. Onex Corp.*, 260 F.R.D. 682 (N.D. Ga. 2009).
37 *Id.* at 691 ("The court believes that some of the most interesting evidence in this matter has come from e-mail production.").
38 *Id. at* 689 ("Defendants argued that . . . the value of the information on the tapes was dubious at best.").
39 *Zubulake*, 217 F.R.D. at 317 ("[E]mail constituted a substantial means of communication among UBS employees.").
40 *Peskoff*, 244 F.R.D. at 60 ("[I]t can be said that the information that has been produced thus far in this case permits the court to infer the possible existence of additional similar information that warrants further judicial action."); *Ameriwood Indus., Inc. v. Liberman*, No. 06-524, 2006 U.S. Dist. LEXIS 93380, at *11 (E.D. Mo. Dec. 27, 2006) ("In light of the Samsung email, the Court finds that other deleted or active versions of emails may yet exist on defendants' computers.").
41 *Ameriwood Indus., Inc.*, 2006 U.S. Dist. LEXIS 93380, at *14-15 ("In the instant action, defendants are alleged to have used the computers, which are the subject of the discovery request, to secrete and distribute plaintiff's confidential information.").
42 *See* FED. R. CIV. P. 26(b)(2)(C)(i) (providing that courts must limit discovery that is "unreasonably cumulative or duplicative").
43 Advisory Committee Notes to 1983 Amendments to Rule 26(b).

What role should nonmonetary factors such as "the importance of the issues at stake" play in a proportionality analysis? Plainly, in the many civil actions that are essentially private disputes (such as breach of contract or traditional tort actions), nonmonetary factors may be irrelevant. However, in civil actions deriving from constitutional or statutorily created rights (such as those brought under 42 U.S.C. § 1983 or Title VII), nonmonetary factors may favor broader discovery. Any proportionality analysis should consider the nature of the right at issue and any other relevant public interest or public policy considerations, and whether, under the particular circumstances of the case, there should be restrictions on discovery.

For example, in *Disability Rights Council of Greater Washington v. Washington Metropolitan Transit Authority*, an action for injunctive relief under the Americans with Disabilities Act, the court denied the defendant's request to limit discovery of backup tapes, given "the importance of the issue at stake and the parties' resources."[44] Other courts have considered nonmonetary issues such as "the strong public policy in favor of disclosure of relevant materials,"[45] and even the health concerns and family obligations of the producing party.[46]

Principle 6: Technologies to reduce cost and burden should be considered in the proportionality analysis.

It is well documented that the volume of ESI is exploding in every corner of the digital world, increasing the volume of potentially discoverable information. Where appropriate, the application of technology to quickly isolate essential information serves the goal of proportionality by creating efficiencies and cost savings. Parties should meet and confer regarding technological approaches to preservation, selection, review, and disclosure that reduce overall costs, better target discovery, protect privacy and confidentiality, and reduce burdens.

Parties and/or their counsel should have a general understanding of the technology available to reduce the cost and burden of electronic discovery in accordance with the proportionality doctrine.[47] These tools and techniques change rapidly and keeping abreast of these changes can present a challenge. Counsel should remain current in the advancements or engage experts as needed to ensure they take advantage of best practices. The Sedona Conference® has published a number of commentaries recently that discuss the application of technology to contain costs and reduce expense and burden.[48]

---

44 *Disability Rights Council of Greater Washington v. Washington Metro. Transit Auth.*, 242 F.R.D. 139, 148 (D.D.C. 2007). The court noted that: "Plaintiffs are physically challenged citizens of this community who need the access to public transportation that WMATA is supposed to provide. That persons who suffer from physical disabilities have equal transportation resources to work and to enjoy their lives with their fellow citizens is a crucial concern of this community. Plaintiffs have no substantial financial resources of which I am aware and the law firm representing them is proceeding pro bono. . . . I will therefore order the search of the backup tapes Plaintiffs seek."*Id.* at 148.

45 *Patterson v. Avery Dennison Corp.*, 281 F.3d 676, 681 (7th Cir. 2002).

46 *Hunter v. Ohio Indem. Co.*, No. 06-3524, 2007 WL 2769805, at *1 (N.D. Cal. Sept. 21, 2007) ("[T]he burden of a deposition on Ms. Jansen, who has virtually no knowledge of any [relevant] issues . . . , and is caring for a spouse with a life-threatening illness, would be inhumane as well as unproductive.").

47 Principle 11 of The Sedona Principles notes that parties may use technological tools for preservation and production: "A responding party may satisfy its good faith obligation to preserve and produce relevant electronically stored information by using electronic tools and processes, such as data sampling, searching, or the use of selection criteria, to identify data reasonably likely to contain relevant information." THE SEDONA CONFERENCE®, THE SEDONA PRINCIPLES (2d ed. 2007), Principle 11.

48 *See* The Sedona Conference®, THE SEDONA CONFERENCE® COMMENTARY ON ACHIEVING QUALITY IN THE E-DISCOVERY PROCESS (May 2009); The Sedona Conference®, THE SEDONA CONFERENCE® COMMENTARY ON PRESERVATION, MANAGEMENT AND IDENTIFICATION OF SOURCES OF INFORMATION THAT ARE NOT REASONABLY ACCESSIBLE (Aug. 2008); The Sedona Conference®, THE SEDONA CONFERENCE® BEST PRACTICES COMMENTARY ON SEARCH & RETRIEVAL METHODS (Aug. 2007) and others available on the Publications page at http://www.thesedonaconference.org. In addition, the ongoing TREC Legal Track has yielded insights into best practices. *See* DOUGLAS W. OARD, JASON R. BARON, AND DAVID D. LEWIS, SOME LESSONS LEARNED TO DATE FROM THE TREC LEGAL TRACK (2006-2009), Feb. 24, 2010, http://treclegal.umiacs.umd.edu/LessonsLearned.pdf.

In considering arguments related to cost and burden, courts may request that the parties provide detailed information regarding the retrieval of electronic information, the use of review tools, and key word searches.[49] Parties familiar with the available technological tools and their costs will have an edge in asserting, or responding to, arguments as to cost and burden.

Parties and law firms that are involved in a significant amount of electronic discovery may choose a standard tool that meets their overall needs. The fact that the standard tool is not the best fit for an individual case should not be held against the firm or the party unless it is conspicuously inadequate for the case, as might happen where the volume of information is unusually high. Parties and law firms may have to consider other tools for cases that exceed the capacity of the standard tool.

While technology may create efficiencies and cost savings, it is not a panacea and there may be circumstances where the costs of technological tools outweigh the benefits of their use.

49 *Apsley v. Boeing Co.*, No. 05-1368, 2007 U.S. Dist. LEXIS 5144, at *19 (D. Kan. Jan. 17, 2007) (scheduling a hearing to consider arguments related to the burden of producing emails, including technological issues).

TSC

ADVANCED LEARNING IN A PANORAMIC SETTING[SM]

Visit www.thesedonaconference.org