Preliminary Report on the

Preservation Costs

Survey of Major Companies

prepared for

Civil Justice Reform Group

William H.J. Hubbard, J.D., Ph.D.
Assistant Professor of Law
University of Chicago Law School

Draft dated: September 8, 2011

Introduction

The Discovery Subcommittee of the Advisory Committee on Civil Rules is currently pursuing the possibility of proposing federal rules to address preservation and spoliation issues in civil litigation. At its most recent major conference, the May 2010 Conference on Civil Litigation at Duke Law School, there was considerable support for new rules in this area. The E-Discovery Panel led by Judges Scheindlin and Facciola issued a statement that the Panel "holds the consensus view that a rule addressing preservation (spoliation) would be a valuable addition to the Federal Rules of Civil Procedure."[1] Nonetheless, much work remains before specific rules can be proposed.

One consensus that emerged at the May 2010 Duke Conference was the need for further empirical research on the magnitude and nature of the costs associated with civil litigation, including discovery and in particular preservation. In response to this need, the Civil Justice Reform Group commissioned me in the spring of 2011 to design and implement an empirical survey of preservation costs borne by large companies in civil litigation.[2] This survey, which I will refer to as the "Preservation Costs Survey" in this report, is part of a larger research agenda in which I am studying the size and distribution of discovery costs, and preservation costs in particular. While this report will focus primarily on the Preservation Costs Survey, I will discuss preliminary results from other aspects of my research to the extent that they are relevant.

Many of the questions that the Preservation Costs Survey seeks to shed light on are the same questions raised by Judge David Campbell and Prof. Richard Marcus as discussion points for the September 2011 Dallas Mini-Conference.[3] These include the following:

- What is the nature of the problem [of preservation of electronically stored information (ESI)], and how are you addressing it?

---

[1] Scheindlin, Shira A., John M. Facciola, Thomas Y. Allman, John M. Barkett, Joseph D. Garrison, Gregory P. Joseph, Dan H. Willoughby, Jr. 2010. Elements of a Preservation Rule. online at http://civilconference.uscourts.gov/.

[2] The Civil Justice Reform Group describes itself as an organization formed and directed by general counsel of Fortune 100 Companies concerned about America's justice system. For biographical information on the author of this report, please see Appendix A.

[3] David Campbell and Richard Marcus, Memorandum (June 29, 2011). The following bullet points are all quoted from this Memorandum.

- In what percentage of lawsuits or potential lawsuits is the problem arising?

- Are problems confined to very large, information-intensive cases, or do they arise in medium and small cases as well?

- What do the problems cost your organization and similar organizations on an annual basis?

- Where are the costs incurred—in identifying and segregating relevant ESI, in storing ESI, in reviewing ESI before production in litigation, in litigating ESI issues in court, in other ways?

  . . .

- The FJC study [Lee (2011)] suggests that spoliation of ESI is rarely raised in federal motions practice. Is that consistent with your experience?

  . . .

- Is there a significant cost associated with storing information preserved for litigation? . . .

- How will technology help reduce the cost of dealing with ESI in litigation over the next few years?

  . . .

- Are cost savings more likely to be achieved through advances in technology than through a rule of civil procedure?

By presenting these questions, Judge Campbell and Professor Marcus highlight the crucial reality that the first order of business in developing sound rules to govern preservation is fact-finding. The current state of knowledge on discovery costs—let alone preservation costs—is rudimentary. While many practicing attorneys have rich and detailed knowledge of their own experience with preservation, commentators have struggled to collect and organize this anecdotal expertise into a coherent empirical picture.

Indeed, to this day there is not even consensus on what litigation costs are for a typical case, with reputable sources providing numbers that may seem surprisingly low (e.g., median defendant's discovery costs of $20,000 in

the *Civil Rules Survey*[4]) to surprisingly high (e.g., discovery costs of $3.5 million for a "midsize" case in the *View from the Front Lines*[5]). As another example, there is anecdotal evidence that many companies fear spoliation sanctions arising out of unclear preservation obligations, yet—as alluded to in the bullet points above—there is also evidence that the imposition of sanctions is rare. Clearly, we need a better handle on the magnitude and nature of the problems with preservation and spoliation before deciding how to address them.

Ongoing research on the discovery process, of which the Preservation Costs Survey is a part, serves to advance our understanding of preservation costs, with the ultimate objective of a better-informed rulemaking process. Indeed, preliminary results that I present below already begin to reconcile some of the disparate results from earlier studies. Nonetheless, the Preservation Costs Survey is currently in its early stages, and more time is required before a more complete picture of the scale and scope of preservation costs emerges.

This preliminary report has four parts, which correspond to its four objectives:

(1) To assess the need for empirical work in this area,

(2) To preview the contributions that this study of preservation costs can provide,

(3) To provide an outline of the design of the Preservation Costs Survey, which includes an initial phase of gathering data from a small sample of companies, followed by a determination of whether a second phase, involving a survey of a broader spectrum of companies, is feasible, and

(4) To describe the preliminary results from the first phase of the Preservation Costs Survey, which involved detailed interviews with, and data gathering from, counsel at four large companies.

I. The Need for Empirical Study of Preservation Costs

Lack of data has been a long-standing impediment to constructive dialogue and reforms addressing the costs of discovery. Over the last few years,

---

[4] Emery G. Lee III and Thomas E. Willging, *National, Case-Based Civil Rules Survey* 35 (FJC 2009). I will refer to this study throughout as the "*Civil Rules Survey*."

[5] Institute for the Advancement of the American Legal System (IAALS), *Electronic Discovery: A View from the Front Lines* 5 (U. Denver 2008).

however, growing awareness of the importance of quantifiable evidence on the benefits and burdens of procedural rules has led to increasingly ambitious efforts to empirically study the costs of civil litigation. Several such studies were presented at the May 2010 Duke Conference. These included the *Civil Rules Survey* by the Federal Judicial Center (FJC), the *Member Survey on Civil Practice* by the ABA Section of Litigation,[6] and the *Litigation Cost Survey of Major Companies.*[7]

Existing studies, to varying degrees, address aspects of the costs of discovery, such as attorney's fees in litigation, document review and production costs, and costs associated with the processing of ESI. These studies provide very little discussion, however, of the costs of preservation.

Relatedly, there is little evidence on the costs associated with legal disputes that do not result in a filed lawsuit. For most categories of legal disputes, many or most disputes never escalate into full-blown litigation—but the possibility of litigation means that preservation obligations and other litigation-related rules impose costs in matters that never even reach the courthouse. One limitation of studies such as the ABA Study and the FJC Study is that they are essentially surveys of outside counsel, and consequently cannot begin to quantify costs that are internal to the client, or costs associated with legal disputes that never reach the point that outside counsel becomes involved.[8]

Understanding the full scope of preservation costs, therefore, requires a close examination of preservation from the potential litigant's perspective, to investigate the time and money devoted to preservation both before and after lawsuits are actually filed. For many individuals and small businesses, of course, litigation is unusual, but for large companies, litigation is an inevitability, with hundreds or thousands of matters (lawsuits or potential lawsuits) active at any given time. Thus, while large companies' preservation activities may not be representative of all litigants, studying large companies provides the best opportunity for the collection of data on preservation costs across a large number of matters, including both actual and potential litigation. The

---

[6] ABA Section of Litigation, *Member Survey on Civil Practice: Detailed Report* (ABA: Chicago, IL 2009) (herein, "ABA Study").

[7] Civil Justice Reform Group, Lawyers for Civil Justice, and U.S. Chamber Institute for Legal Reform, *Litigation Cost Survey of Major Companies* (Searle Center on Law, Regulation, and Economic Growth: Chicago, IL 2010) (herein, "*Litigation Cost Survey*").

[8] A preliminary report from one of the companies participating the Preservation Costs Survey indicates that 44 percent of matters with preservation hold notices do not involve a filed lawsuit.

Preservation Costs Survey intends to collect more data, and richer data, on preservation costs than is currently known.

II. Contributions a Study of Preservation Costs Can Provide

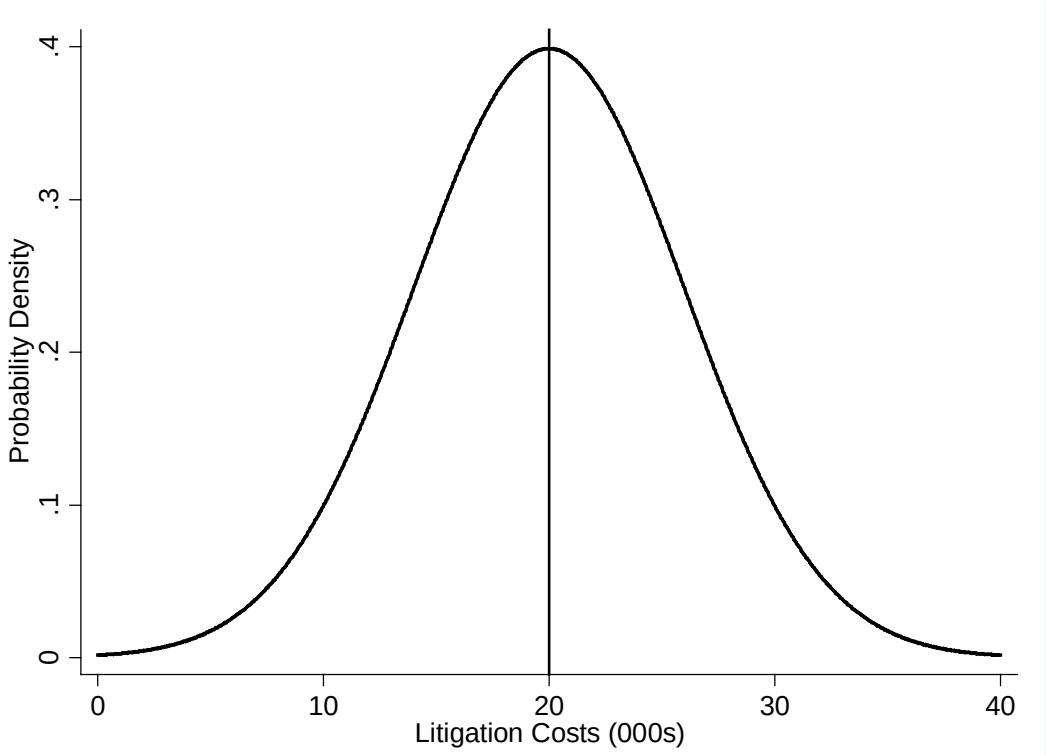
As noted above, this preliminary report is intended to preview the Preservation Costs Survey in light of the need for more empirical evidence on the costs of preservation. Evidence alone, however, is not sufficient to fill the gaps in our knowledge of the challenges presented by preservation obligations. What is also essential is an ongoing dialogue on how to interpret the evidence. My research intends to contribute to this dialogue by applying some fundamental statistical and economic insights to the results of various studies, including the Preservation Costs Survey, and providing context to what otherwise might be conflicting or incomplete statistical accounts.

Litigation Costs and the Long Tail

For example, consider the fundamental question: what does the distribution of litigation costs look like? This is a question that recent studies have not specifically taken up—but, as I will explain, is essential to understanding the nature of the costs that discovery and preservation obligations impose. An important source of information to date on the costs of litigation (but not preservation) is the *Civil Rules Survey*. One of the most striking results of the survey is that in the median case—specifically, the median case *with* discovery—the costs of litigation are (arguably) modest, $15,000 for plaintiffs and $20,000 for defendants. And of these costs, only a fraction (20 to 30 percent) are due to discovery.

Given these numbers, it would be fair to ask whether discovery is in fact such a significant source of costs. If the median cost of discovery for defendants is $20,000, we are likely to visualize a distribution of costs that looks something like a bell curve, or normal distribution, as illustrated in Figure 1. Our intuition is that, given a median cost of $20,000 for defendants (the vertical line in Figure 1), most defendants experience costs close to that median amount, in the same way that most test scores are close to the median score and students' grades tend to fall into a bell curve.

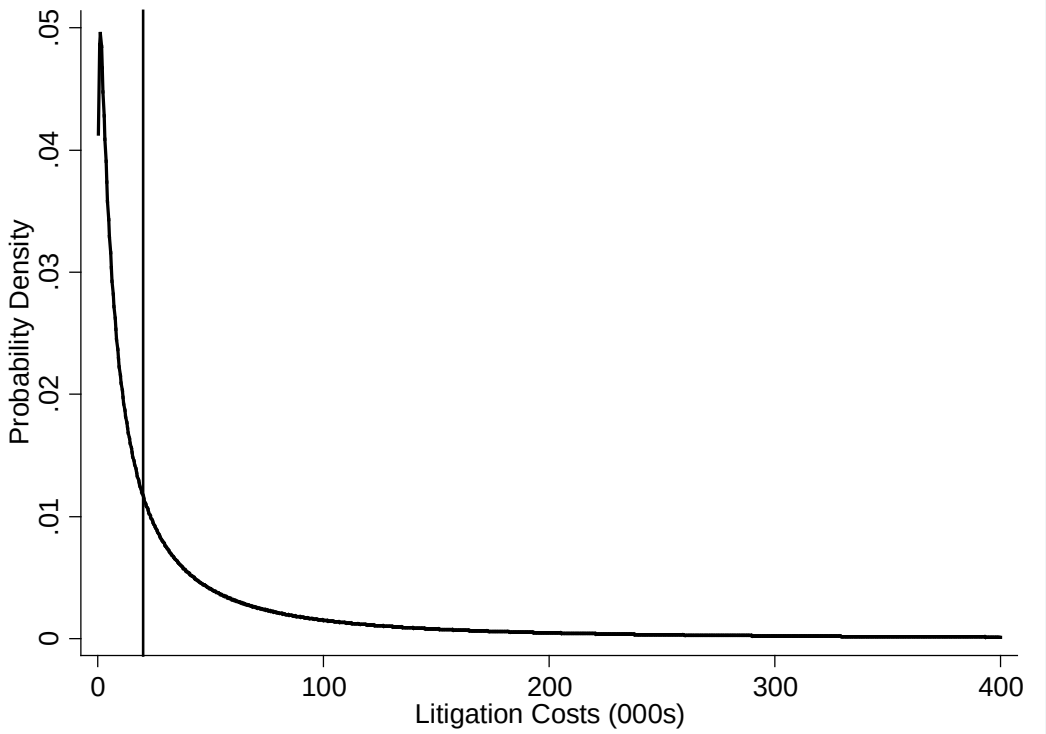
FIGURE 1: LITIGATION COSTS WITH MEDIAN OF $20,000,
ASSUMING A NORMAL DISTRIBUTION OF COSTS



This intuition, however, would lead us astray. Litigation costs are not normally distributed. The clue to seeing this is to look at the *Civil Rules Survey* reports of the 10th and 95th percentiles of the distribution of costs. The 10th percentile, $5,000 is one-fourth the median, but the 95th percentile, $300,000, is fifteen times the median! In other words, this is a clue that litigation costs are not like test scores, with a normal distribution of costs clustered close to the median, but instead more like the distribution of income, or the distribution of stock returns—in other words, a "long tail" phenomenon, where there is a large mass close to zero, but also a long tail of extreme, and extremely important, outliers.

How does this change our intuition about litigation costs? Let's fit the data from the *Civil Rules Survey* to the log-normal distribution, which is a distribution used to describe the distribution of income and which fits the data published in the *Civil Rules Survey* quite well. This is what the distribution of costs looks like:

FIGURE 2: LITIGATION COSTS WITH MEDIAN OF $20,000,
ASSUMING A LOG NORMAL DISTRIBUTION OF COSTS



Once again, the median is marked with a vertical line. But we now see that while the bulk of cases are still close to the median, there is also a "long tail" of extremely costly cases that are nowhere close to the median. How important is this "long tail"? Consider the following: in the distribution illustrated above, *the top 5 percent of cases accounts for 60 percent of all litigation costs.*

In this light, it is helpful to consider the *Civil Rules Survey* together with the *Litigation Cost Survey*. The *Litigation Cost Survey* can be (rightly) criticized as *not* a representative sample of all lawsuits, or even of all lawsuits at large companies. It focuses on the cases with the highest litigation costs. But the *Civil Rules Survey*, which does canvas a representative sample of lawsuits, reveals that the distribution of litigation costs is such that the largest, most expensive cases carry great weight in the calculus of litigation costs.

In short, one response to the *Civil Rules Survey* is to ask, "If most cases have low discovery costs, why should we devote resources to rules reform that may affect only the 5 percent of cases with high discovery costs?" But perhaps a better question would be, "Should we explore rules reform, if a reform that affected only 5 percent of cases could help control 60 percent of litigation costs?"

Do Preservation Costs Have a Long Tail?

The next question that arises is whether we find a similar, "long tail" pattern for preservation costs. To answer this question, the Preservation Costs Survey will be essential. Without data, we won't know whether preservation costs have a skewed distribution in the same way that litigation costs do.

After all, we might expect the skewness of the distribution of litigation costs to arise out of the litigation process itself. Many cases settle early with little discovery, while a few cases go all the way to trial. The low median of litigation costs could merely reflect the fact that most cases settle early.

This factor, however, should not affect the distribution of preservation costs, because the preservation obligation attaches at or before the onset of litigation. Most preservation costs will be imposed on the parties regardless of whether the case settles early or goes all the way to trial.

A second factor is that case complexity may have a highly skewed distribution, so that the long tail of litigation costs partly reflects a long tail of very complex disputes. To the extent that the skewness of litigation costs is driven by case complexity, we might expect preservation costs to have a distribution with a long tail as well.

Some preliminary results from the Preservation Costs Survey offer suggestive evidence in this regard. Two of the companies participating in Phase I of the Survey (described in more detail below) have provided data on a sample of litigation matters opened during two recent sample periods. In the Company A data, for each matter there is information on the number of hold notices issued and interviews conducted during a two-year window. In this sample, there are 112 distinct matters representing actual or anticipated civil litigation.[9] During the sample period, a total of 5021 distinct actions were taken—these include issuances of a litigation hold notice to an individual, interviews, and revisions to and terminations of litigation holds. Of the 112 sample matters, the top five (which is 4.5% of the total) account for 1410 of the 5021 actions—which is more than 28 percent of all actions. Indeed, more than half of all preservation activity was generated by only 16 (or 14.3%) of the matters. As Figure 3 illustrates, preservation activity across cases as Company A has a long tail, although not as extreme as the long tail for litigation costs in the *Civil Rules Survey*.

[9] Note that this sample excludes certain categories of cases, such as asbestos cases, but is otherwise representative of civil matters requiring litigation holds.

FIGURE 3: DISTRIBUTION OF PRESERVATION ACTIONS TAKEN, COMPANY A LITIGATION HOLD SAMPLE

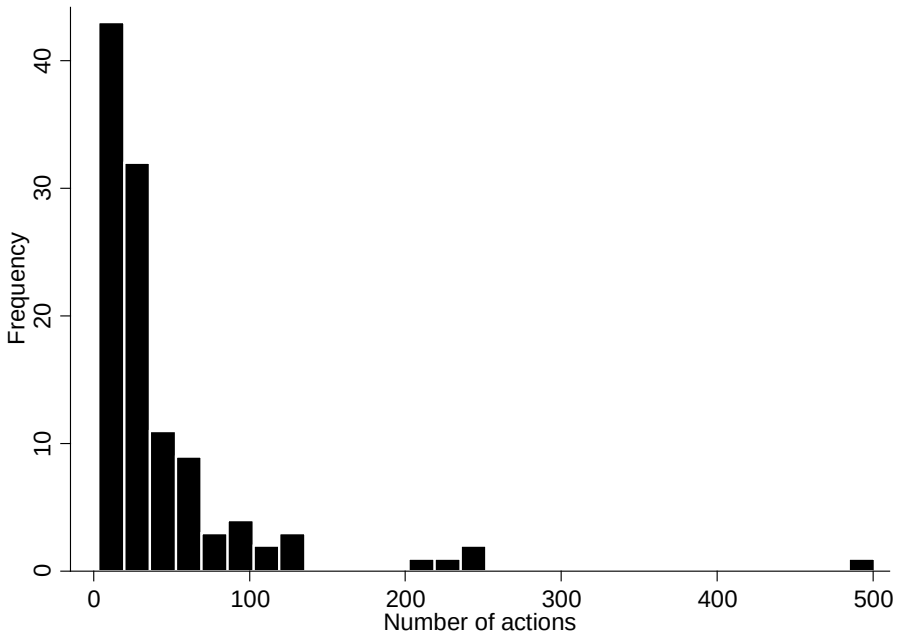
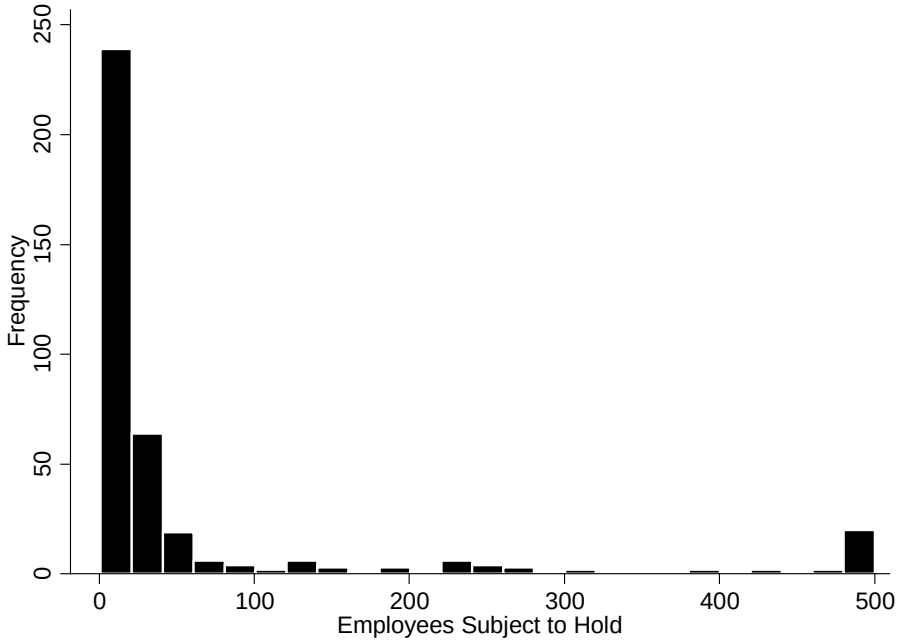


FIGURE 4: DISTRIBUTION OF NUMBER OF EMPLOYEES ON HOLD PER MATTER, COMPANY D LITIGATION HOLD SAMPLE



The data from Company D covers 390 distinct matters representing actual or anticipated civil litigation. For each matter the dataset provides the number of individuals subject to a litigation hold in that matter. During the five-year sample period, a total of 43,011 holds were issued. In this sample, five percent of the matters account for more than 62 percent of the holds issued (26,864 holds out of 43,011). See Figure 4. [10]

This preliminary data suggests that preservation costs, like litigation costs, are highly skewed, with a long tail in which a small number of highly complex and burdensome cases account for a large share of the total costs borne by individuals subjects to holds. It may therefore be productive to think in terms of steps that can address the burdens of large, information-intensive cases in particular.

The "Fixed Costs" of Preservation

Existing surveys of litigation costs, such as the *Civil Rules Survey* and the *Litigation Cost Survey* focus on the costs of litigation on a per-case basis. As the figures above illustrate, the Preservation Costs Survey seeks to measure the per-matter costs of preservation as well. But a study of preservation costs has to account for a second type of cost as well. While many costs of preservation, such as the costs of responding to litigation holds, accrue on a per-case basis, other preservation costs are not tied to a particular matter, but instead reflect the costs for a company to create internal systems to handle preservation across all cases. These "fixed costs" include expensive investments in technology that companies make in order to control what would otherwise be even higher per-case preservation costs.

Importantly, while fixed costs are not captured at all by the figures above, the Preservation Costs Survey is measuring fixed costs separately. I have initial data from two companies on the costs of computer systems (both hardware and software) implemented by those companies to handle aspects of preservation. One fixed cost is the cost of systems to handle litigation hold notices. Company A implemented a system to partially automate the issuing and tracking litigation holds at a cost of approximately $900,000. Company B is in the processing of implementing a new system with similar goals, and at a similar cost (estimated to be $800,000). In addition to implementation costs are upkeep and maintenance costs, which Company A estimates to be $150,000 per year.

[10] In Figure 4, note that for graphical clarity, matters with more than 500 employees subject to hold have been included in the category for 500 employees subject to holds.

By far the largest fixed costs, however, are associated with the preservation of data itself. Every large company that I have encountered, both in my practice experience and in connection with the Preservation Costs Survey, has had a diverse set of systems used to address preservation obligations. This is because of the large variety of types of ESI, many of which have distinct business purposes and are used and stored in different ways on a company's computer systems. To preserve all types of ESI, therefore, requires multiple preservation solutions.

Gathering data on the costs of all of the systems used for preservation in any given company is a daunting task; it may not be feasible for the Preservation Costs Survey to collect such a comprehensive set of costs data. But Phase I of the Survey has been able to identify costs for specific, recently-implemented systems for which individual companies have information on costs. For example, the tools used by Company A to collect data to be preserved at the outset of litigation—which is only a fraction of the data preserved—cost $4,800,000 to implement. The data vault system that Company B uses to preserve certain types of ESI, including email, cost $12,000,000 to implement and maintain in 2010.

These are the costs for individual systems designed to address specific elements of the preservation obligation. A more comprehensive measure of costs is much harder to quantify, both because of the number of systems involved and because so many personnel within a company share responsibilities for preservation, including individuals who otherwise have no connection with the law or litigation. Unlike litigation costs for outside counsel, there are no itemized records of the costs of time spent by company employees on preservation. One of the goals of the Preservation Costs Survey is to measure the cost of time spent on preservation by these individuals.

Ideally, too, we would like data from a larger set of companies to measure both fixed costs and preservation costs associated with individual cases. With this in mind, I will now turn to a description of the Preservation Costs Survey.

III.The Preservation Costs Survey

As noted above, I am currently in the process of undertaking a survey of preservation costs at large companies. There are a number of aspects of preservation costs that are unlike other litigation costs and which are particularly difficult to quantify. These include:

- Costs of discovery borne by in-house counsel and non-legal employees, rather than by outside counsel;

- Costs to IT infrastructure;

- Costs from diversion of resources from non-legal functions; and
- Costs from risk and uncertainty of legal rules governing preservation.

The goal of the Preservation Costs Survey is to obtain quantitative data on these previously unmeasured costs and apply statistical and economic analysis to this data. The desire is to inform the discussion on preservation costs and rules reform.

Given the complexity of the topic, and the largely unprecedented nature of a study focused on preservation costs, I have established a two-phase study design. Both phases of the Preservation Costs Survey involve the gathering of information from large companies on a strictly confidential basis to ensure that responses are as candid and complete as possible.

Phase I has already begun. Phase I has involved a set of four, in-depth "case studies" of large companies. These case studies have involved both qualitative interviews and requests for quantitative data to be used for statistical analysis. The case studies have also included extensive written survey testing in order to explore the feasibility of data gathering on each of the questions above. This information will be used to determine whether a broader survey is feasible, and if so, to draft an effective survey instrument for use with a larger sample of companies during Phase II.

Phase II, if feasible, will begin some time after the Dallas mini-conference. It will involve the creation of a final survey instrument to be used in a survey of a larger number of companies. Together with the administration of this survey, I will continue qualitative interviews and the collection of datasets of preservation activity from selected companies in order to create as complete a picture of the sources and amounts of preservation costs for large companies. The goal of Phase II is to have the survey responses collected by early 2012. Based on analysis of the surveys, interviews, and datasets, I will prepare a report on the Preservations Costs Survey in early 2012.

While Phase I has primarily served to lay the groundwork for Phase II, the case studies I have conducted have already yielded some valuable, even if preliminary, results. I have discussed some of these insights above. Below, I describe other results from Phase I of the Survey.

IV. Additional Results of Phase I of the Preservation Costs Survey

In my initial investigations, I am encountering a few recurring themes in the interviews and responses from companies. I will describe these themes here, with the caveat that these are only preliminary impressions, and that a final report at the conclusion of Phase II will present a more systematic review of the responses of a larger sample of companies.

Quantifying the Costs of Preservation Is Difficult, as the Costs Are Diverse and Borne by Many Groups within a Company

Phase I of the survey design focused not only on quantifying some elements of the costs of preservation, but on understanding which aspects of the costs of preservation are most susceptible to study and which will be the hardest to estimate. Not surprisingly, the interviewed companies expressed that estimating the costs of preservation is difficult. The reasons for this are several:

First, unlike litigation costs such as outside counsel fees, the costs of preservation are borne in-house. Further, although some individuals, particularly in the Legal and Legal IT functions, may spend most or all of their time dealing with preservation issues, the vast majority of individuals affected by the preservation obligation are not connected to the legal function at all. Instead, they are employees devoted to the business function, who happen to be custodians of data that may be relevant to a legal matter or they are employees devoted to the IT function, who happen to be responsible for systems that may contain data relevant to a legal matter. As noted above in Part II, the time and energy they must divert towards preservation is never recorded or compensated, unlike the time spent by dedicated lawyers, such as outside counsel.

Second, in today's environment, preservation essentially requires the use of automated systems for some or all aspects of preservation, including identifying custodians, issuing holds, and facilitating the preservation of ESI. Quantifying the cost of designing, implementing, and maintaining such systems can be difficult. Even systems purchased from outside vendors, for which there is an identifiable price tag, have costs that are hard to quantify, such as the time of in-house lawyers and IT specialists, the time of users, and the costs of upkeep and maintenance.

Third, not only are the individuals affected by preservation diffused throughout a company, but the types of actions that must be taken to preserve data are widely varied as well. Some actions are routine and easily described (even if estimating cost is difficult), such as designing and issuing litigation hold notices, or creating an archive of preserved emails. But other actions arise irregularly and sometimes require ad hoc solutions. These situations may arise in the context of departing employees, from whom data may need to be collected from hard drives or loose media. This may sound like a trivial undertaking if a single employee is involved, but the interviewed companies see thousands of employees leave each year.

Other issues arise less frequently, but are even more tricky. Obsolete data formats or storage systems need to be updated, and migrating data to new systems without the loss of information on hold can be difficult, requiring workarounds tailored to the specific systems. These steps can cost millions.

This is not to say that the costs of preservation should not be, or cannot be, estimated. Rather, the costs of preservation are extensive and varied, requiring further study before we can measure them with any confidence.

Some of the Largest Costs of Preservation Are Related to Relatively Small Categories of Preserved Material

Interviewees in Phase I have explained that many of the largest costs of preservation are related to the less salient aspects of preservation: legacy data, data migration, data that was on hold but which has been released, and data left over when a litigation hold ends. For example, Company A notes that some of the biggest headaches for preservation involve departing employees' hard drives, the migration of legacy data to current systems, the preservation of data on computers and systems that require maintenance, repair, or updates. Attempts to reduce these costs have led to delays in the roll-out of new applications and the delay of roll-out of new computers to employees on hold. This has not only impacted productivity, but invited an understandable backlash from employees on hold. In this way, some of the seemingly obscure aspects of preservation have had outsized effects on business efficiency and employee morale.

Uncertainty about Preservation Obligations Leads to Overbroad Preservation

Another common theme is that uncertainty about the scope of the preservation obligation and the consequent fear of sanctions leads companies to preserve more than would otherwise be justified. Sanctions, of course, can be very costly in monetary terms and can lead to adverse outcomes on the merits in litigation as well. They also have a severe reputational cost, and large companies, no less than individuals, tend to work hard to avoid even the appearance of being a scofflaw. For example, Company A expressed that its policy is to make legal compliance a top priority, and thus the company seeks to avoid sanctions or the perception of spoliation even if it is very costly to do so—and it appears that it often is.

This reluctance to risk sanctions is consistent with a recent study of motions for sanctions, which found a motion related to spoliation of evidence in only 0.15 percent of cases.[11] This figure is supported by initial Phase I survey results, where Company A estimates that motions for sanctions are filed in less than 0.5 percent of its cases.

---

[11] Emery G. Lee III, *Motions for Sanctions Based Upon Spoliation of Evidence in Civil Cases: Report to the Judicial Conference Advisory Committee on Civil Rules* (Federal Judicial Center 2011).

FIGURE 5: NUMBER OF MATTERS WITH PRESERVATION, COLLECTION, AND PROCESSING (COMPANY A)

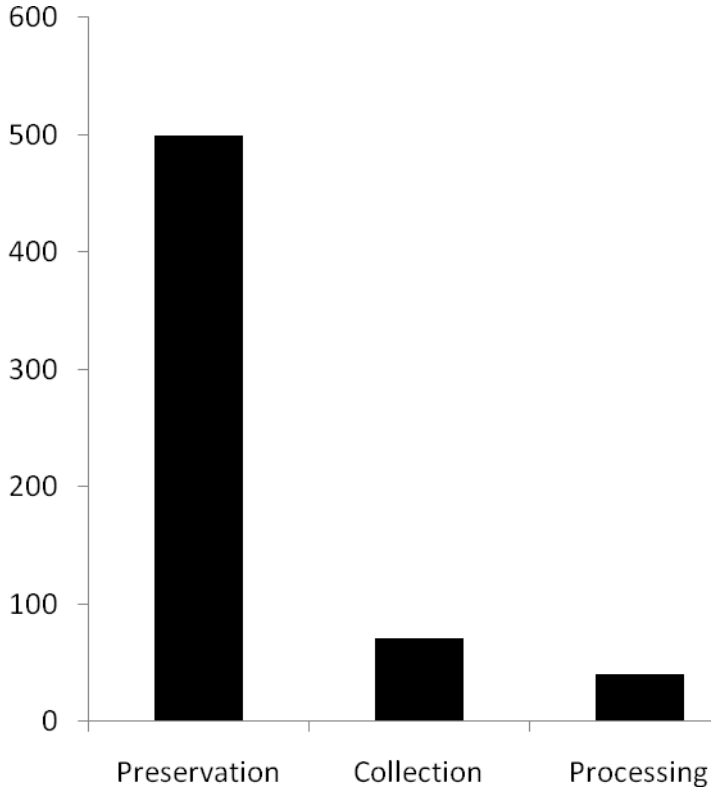
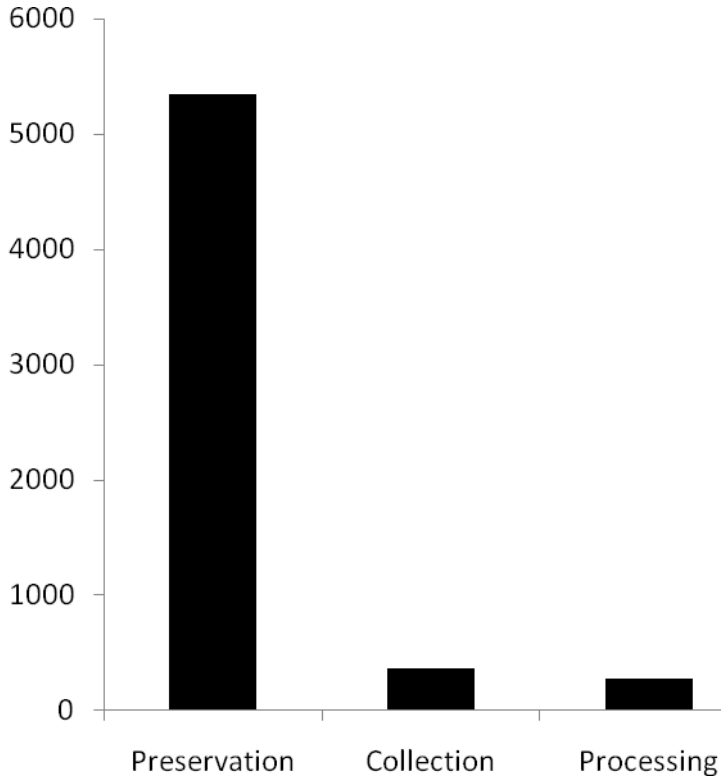


FIGURE 6: NUMBER OF CUSTODIANS WHOSE DATA IS SUBJECT TO PRESERVATION, COLLECTION, AND PROCESSING (COMPANY A)



What is the cost that Company A must pay in order to avoid the specter of a spoliation claim? More research is required before I can quantify these costs, but some preliminary data provides some insight into the extent of overbroad preservation. At Company A, data is collected in only 14 percent of matters in which data is preserved, and data is processed for review in only about 8 percent of matters. See Figure 5. Looking at individual custodians rather than matters, there is an even more stark difference between the amounts preserved and the amounts ultimately collected and processed. See Figure 6. In short, the vast majority of the data that is preserved is ultimately judged unnecessary to the litigation. But the vast majority of data that is never used still imposes preservation costs.[12]

Technology Both Creates More Efficient Methods of Preservation and Creates New Costs and Complexities

Technology has become a central part of business life, and it has come to dominate the practice of discovery and preservation in particular. My interviews have revealed that rapidly advancing technology for data storage and processing has been both a source of rising costs and of cost savings.

One major cost, alluded to above, is that advancing technology means that companies have to account for an ever-growing number of legacy formats and platforms, which often require expensive and time-consuming data migration and archiving efforts. Even advances in hardware cause problems, because as computers are replaced, special efforts are needed to preserve data on individual hard drives and other storage media.

One cost that is less often discussed is the fact that technology has necessitated the creation of entirely new departments within companies. The companies interviewed all have what could be called (and usually is called) a "Legal IT" function. This is a group or department that spans the space between Legal and IT to ensure that the company's legal obligations with respect to its IT infrastructure are met. As a practical matter, this means that most of what Legal IT does is handle matters relating to the preservation of ESI. For example, Company D has at least seven employees whose time is essentially dedicated to coordinating the IT aspects of preservation and collection in-house.

Of course, it is important to recognize that technology creates opportunities for efficiencies, in addition to creating complexities. Company D describes how it is working with outside vendors to improve the process for defining searches for email, so that a more precise set of emails is preserved in re-

[12] The Preservation Costs Survey is working to determine the extent to which these costs can be quantified.

sponse to a litigation hold. Another example is software designed to assist in indexing, searching, and foldering preserved data for collection and processing. Company D has spent around $1 million to implement and maintain such a system over the last two years, but the interviewees see this cost as a fraction of the savings it has generated.

Conclusion

This preliminary report on the Preservation Costs Survey begins to address the serious need for data and analysis on the nature of preservation costs. While the Preservation Costs Survey is currently in its early stages, some initial results have emerged. For example, the costs of preservation, like the costs of litigation, exhibit a "long tail," meaning that a small fraction of cases account for most of the expenses associated with individual cases. Further, many costs of preservation are "fixed costs," representing multi-million dollar investments in technology to track and manage the preservation of an ever-expanding universe of ESI. Both case-specific costs, and the fixed costs of preservation, could potentially be subjects for rules reform.

Of course, I should reiterate that these results are preliminary, and it would be premature to judge any proposed rules based only on preliminary findings. The Preservation Costs Survey will generate additional results from a larger sample of companies in the coming months. I will prepare a detailed report on the Survey in early 2012 to describe and analyze the full set of results.

Appendix A: Biographical Information on William H.J. Hubbard

After graduating from the University of Chicago Law School with high honors, I clerked for the Honorable Patrick E. Higginbotham of the U.S. Court of Appeals for the Fifth Circuit during the 2000 term. I worked as a litigation associate at Mayer Brown LLP from 2001 through 2006, where I was an original member of the firm's Electronic Discovery and Records Management Group. As a member of this Group, I developed protocols for the preservation of electronically stored information and created materials to be used for defense-of-process in e-discovery disputes. My experience included conducting on-site interviews and investigations related to preservation technology and processes for large companies. Other aspects of my practice consisted of a broad range of pre-trial litigation and appellate litigation.

In 2006, I entered the PhD program in Economics at the University of Chicago. I received my PhD in August of this year. I have published or forthcoming papers in the *American Economic Review Papers & Proceedings*, *Journal of Human Resources*, and *Journal of Human Capital*. I have presented working papers at the Annual Meetings of the American Economic Association, the Milton Friedman Institute, and the University of Chicago Law School.

I am an Assistant Professor of Law at the University of Chicago Law School. I teach courses and seminars on civil procedure and economic analysis of law.