
Preservation Costs Survey

Final Report

prepared for
Civil Justice Reform Group

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

February 18, 2014

Executive Summary	3
1 Introduction	4
2 Methodology.....	9
2.1 Overcoming Challenges to Measuring Preservation Costs	9
2.2 Survey Design.....	13
3 Sample Characteristics	17
4 Results	20
4.1 Companies Report Significant Preservation Burdens	20
4.2 Preservation Challenges for Smaller Companies.....	33
4.3 A Small Share of Matters Accounts for Most Litigation Hold Activity.....	39
4.4 Vastly More Data Is Preserved than Used in Discovery	43
5 Conclusion: Preservation Costs and the Federal Rules	47
5.1 Preservation Costs Are Substantial.....	47
5.2 Smaller Companies Face Serious Burdens and Benefit Less from Technology.....	48
5.3 A Small Percentage of Matters Generates Most Litigation Hold Activity.....	49
5.4 Reducing Overpreservation Will Have Little Negative Impact on Discovery	49
6 Appendix	50
6.1 Additional Figures for Phase II Companies	50
6.2 Biographical Information on William H.J. Hubbard.....	53
6.3 Phase III Survey Questionnaire	54

Executive Summary

The Preservation Costs Survey collected data from 128 companies, including companies of all sizes and from a broad range of industries. Data includes survey responses as well as interviews and detailed, case-level information on litigation hold activity from a subset of companies. No prior work has collected quantitative costs data from a cross-section of companies.

The Survey generated conservative estimates of costs that are solely attributable to preservation obligations. Among the largest companies in the sample, the estimated costs exceed \$40 million per company per year.

Both larger and smaller companies report similar preservation burdens. Over 79 percent of respondents reported a “great extent” or “moderate extent” of preservation burdens. Further, smaller companies are far less likely than large companies to have specialized resources to address the risks and costs of preservation. Thus, smaller companies are more vulnerable to legal uncertainty in this area, including the possibility of sanctions with severe effects on their ability to do business.

A small percentage of litigation matters generate a disproportionate share of preservation costs. Five percent of litigation matters account for more than half of all litigation hold notices issued.

Companies report “overpreserving” to protect against serious uncertainty in the case law. Rules amendments that better define the standards for sanctions for failure to preserve could address this phenomenon.

Only a fraction of preserved data is ever collected. On average across all survey respondents, slightly less than half of all preserved data is ever collected, processed, and reviewed. Even less is produced or eventually used in litigation.

Rule changes with even modest effects would generate meaningful cost savings. For the largest companies in the sample, a 3 percent reduction only in employee time spent on litigation holds would equate to savings of over \$1 million per company per year.

Because so little preserved data is ever used, reducing overpreservation in a reasoned fashion is unlikely to have much, if any, negative impact on the production and use of data in litigation. Rules amendments that rein in overpreservation will likely have essentially no adverse impact on discovery and the ultimate resolution of litigation.

1 Introduction

The Advisory Committee on Civil Rules has published for public comment proposed amendments to the Federal Rules of Civil Procedure (“Rules”) that would address, among other things, the preservation of documents and electronically stored information (“ESI”) in federal litigation.¹ This activity comes amid widespread calls for rules reform arising out of frustration with the patchwork of case law that currently governs preservation and sanctions for spoliation in federal court litigation. While there has been considerable debate about the merits of various proposals to amend the Rules, there has been consensus on the need for further empirical research on the magnitude and nature of the costs associated with civil litigation, including the costs of discovery and, in particular, preservation.

The current state of knowledge on discovery costs—let alone preservation costs—is limited. While many practicing attorneys have rich and detailed knowledge of their own experience with preservation, commentators have struggled to collect and organize this anecdotal information 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 low² or surprisingly high.³ As another example, there is anecdotal evidence that many companies fear spoliation sanctions arising out of unclear preservation obligations, yet there is also evidence that the imposition of sanctions is rare.⁴ There is a clear need for more information on the magnitude and nature of the problems with preservation and spoliation.

Lack of data has been a long-standing impediment to constructive dialogue and reforms addressing the costs of discovery. As a recent report has

¹ Note that throughout this paper I will use “documents” and “data” interchangeably to refer both to paper records and ESI.

² An FJC study reports that the median defendant’s discovery costs for civil cases in federal court are \$20,000. Emery G. Lee III and Thomas E. Willging, *National, Case-Based Civil Rules Survey* 35 (FJC 2009) (herein, “*Civil Rules Survey*”).

³ A study by the Institute for the Advancement of the American Legal System (IAALS) reports discovery costs of \$3.5 million for a “midsize” case. IAALS, *Electronic Discovery: A View from the Front Lines* 5 (U. Denver 2008).

⁴ 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* (FJC 2011).

noted, the “actual costs of discovery have rarely been quantified in empirical studies.”⁵ Over the last few years, 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. One limitation of studies such as the ABA Study and the *Civil Rules Survey* is that they are essentially surveys of outside counsel. Consequently, they cannot quantify costs that are internal to the client or costs associated with legal disputes that never reach the point that outside counsel becomes involved. Yet preservation activities tend to involve personnel of the client, rather than outside counsel, and often begin before a lawsuit is filed. 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.

One recent study does attempt to shed light on the costs and of preservation. A RAND Institute for Civil Justice study gathered information on electronic discovery costs directly from companies who were parties (plaintiffs or defendants) to litigation. Even this study, however, despite collecting detailed information on collection, processing, and review of data in discovery, was unable to present *any* quantitative data on preservation costs. Rather, the researchers were limited to in-depth, qualitative interviews with

⁵ RAND Institute for Civil Justice, *Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery* 4 (RAND 2012) (herein, “Where the Money Goes”). See also *id.* at 3 (“A repeated lament in the academic and legal literature is that there has been little or no research into the costs imposed on the larger judicial system by the discovery process.”) (internal quotation marks omitted).

⁶ ABA Section of Litigation, *Member Survey on Civil Practice: Detailed Report* (ABA 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 2010) (herein, “Litigation Cost Survey”).

eight companies.⁸ Indeed, no empirical research prior to the Preservation Costs Survey has gathered quantitative data on preservation costs.⁹

To address the need for empirical data on the costs of preservation, the Civil Justice Reform Group commissioned me in the spring of 2011 to design and implement a study of preservation costs.¹⁰ This survey, the Preservation Costs Survey (“Survey”), is the first, and to date only, systematic effort to measure the extent and costs of preservation activity across a cross-section of companies.

The Survey has collected more data, and richer data, on preservation costs than any prior work. To do so, the Survey first identified the obstacles that have prevented sustained empirical research on this topic in the past. It then developed a suite of approaches to gather a wide array of quantitative and qualitative information on preservation activities and their associated costs. In Part 2 of this Report, I describe this process in greater detail.

The Survey ultimately collected information from 128 companies from a wide spectrum of industries. These companies vary from small companies without in-house litigation counsel to Fortune 100 companies who have entire staffs of attorneys and other professionals devoted full-time to preservation work. Part 3 describes the sample of companies that responded to the survey.

The heart of this Report is Part 4, which presents and discusses results from the Survey.¹¹ These results are organized into four themes.

⁸ *Where the Money Goes* at 15 (“Our approach here was qualitative in nature because it was clear that gauging the magnitude of preservation expenses in individual cases would present some daunting hurdles.”).

⁹ *Where the Money Goes* at 86 (“Despite the costs of preservation having become one of the most discussed topics in the legal press of late, we are not aware of any empirical research that has collected quantitative information about such costs across significant numbers of actual cases.”).

¹⁰ 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 the Appendix.

¹¹ Results from Phase I of the Survey were presented in the Preliminary Report on the Preservation Costs Survey submitted to the Discovery Subcommittee in connection with the September 9, 2011 Dallas Mini-Conference. All findings in the Preliminary Report are consistent with the complete results now available in the Final Report.

First, companies report significant preservation burdens. Most companies report a “great extent” or “moderate extent” of burdens related to preservation. This qualitative assessment is strongly supported by the quantitative data on two main categories of preservation costs, which I refer to as “fixed costs” and “variable costs.”

The fixed costs of preservation are the costs of investments in people and technology in response to preservation obligations. These include the costs of maintaining a staff of attorneys, IT specialists, and other professionals devoted to preservation activity, as well as the costs of automated systems to manage litigation holds and preserve data. These costs are “fixed” because they do not arise in the context of individual litigation matters, but represent ongoing expenses of a company.¹² In Part 4.1, I estimate the fixed costs of preservation for a large company to exceed \$2 million per year.

The variable costs of preservation are those costs that arise in the context of individual litigation matters (and thus vary with the volume of litigation activity). The primary variable cost, and the one that this Survey studies, is the time that non-legal employees who are subject to litigation hold obligations must divert from business activities to compliance with a litigation hold.¹³ While virtually all prior reported information on the costs of preservation reflect the fixed costs of technology, I find in the Survey data that the cost of lost employee time is *by far* the largest source of preservation costs—perhaps 90 percent of total costs due to preservation obligations. Thus, the millions spent on automated systems to manage litigation holds are no more than the tip of the iceberg of total preservation costs.

I quantify this cost by estimating the average number of litigation matters, and the average number of employees subject to hold in each matter, at companies of different sizes. I then take into account *conservative* estimates of the time spent on litigation holds per employee, and the value of employee

¹² I will refer to “litigation matters” rather than “cases” throughout, as preservation costs are incurred both in the context of filed lawsuits and in reasonable anticipation of litigation.

¹³ The primary measure of preservation activity in this study is the number of litigation hold notices issued. A “litigation hold” is set of actions taken by a company to comply with preservation obligations in a given litigation matter. A litigation hold will define the scope of documents and data that must be preserved. A “litigation hold notice” is an instruction from legal counsel to an employee that the employee must retain all documents and data in her custody that are within the scope of the litigation hold (for example, in a products liability case, the scope might be all documents relating to the safety of a particular product that the company produces). The usual practice is to send a litigation hold notice to the set of “key players” who are likely to have data relevant to the dispute in question.

time, to express this cost in dollar terms. This variable cost of preservation is substantial: at the smallest companies in the sample, I estimate an average cost of over \$12,000 per company per year; for the largest companies in the sample, it is over \$38 million per company per year.

Second, smaller companies report a similar extent of preservation burdens to larger companies. The data reveals that, although smaller companies face fewer lawsuits, they are also less well equipped to respond to the preservation challenges that do arise. In particular, they benefit less from technological solutions to the costs of preservation, which tend to be cost-effective only for large companies.

Third, a small share of litigation matters accounts for the majority of litigation hold activity. Most litigation matters involve a relatively moderate number of employees on hold, but a fraction of matters involve large numbers of employees on hold. This distribution of litigation hold activity is so skewed that a majority of all litigation hold notices are due to only 5 percent of all matters.

Fourth, most of the data that is preserved is never collected, processed, and reviewed. This finding is consistent with the numerous reports of companies in interviews that they “overpreserve” data in order to protect themselves against the current environment of highly uncertain legal standards governing preservation and sanctions for failure to preserve.

Part 5 discusses some implications of this Survey in relation to proposed Rules amendments addressing preservation. The four themes described above suggest four implications for these proposed Rules amendments. First, preservation costs are substantial, and thus Rules amendments have the potential to generate meaningful cost savings. Second, both smaller and larger companies stand to benefit from such savings, and although technology alone might partially address high preservation costs, most small companies are not in a position to benefit from technology. Third, because the burdens of preservation activity are not distributed uniformly across litigation matters, Rules amendments need not affect a large share of cases in order to affect a large share of preservation activity. Fourth, because most preserved data is never reviewed in the discovery process, modest reductions in the extent of “overpreservation” activity will not generate adverse consequences for the production of relevant data in discovery.

Part 6 contains the appendix, which includes additional figures illustrating the distribution of litigation hold burdens across litigation matters; biographical information on the author of the study; and a copy of the final survey questionnaire.

2 Methodology

2.1 Overcoming Challenges to Measuring Preservation Costs

In order to measure the costs associated with preservation obligations, this Survey had to overcome a number of challenges that prevented prior research from determining the nature and scale of preservation costs. As I will describe in Part 2.2, the Survey took a phased approach to data gathering, recognizing that any effort to quantify costs of preservation would have to begin by identifying *which* costs of preservation are even susceptible to practical measurement. Thus, the first phase of the survey design focused on understanding which aspects of the costs of preservation are most amenable to study and which would be difficult, or as a practical matter impossible, to estimate. Not surprisingly, every company interviewed for the Survey expressed that estimating the costs of preservation is difficult.¹⁴ In this Part, I explain the main obstacles to quantification of the costs of preservation and the ways in which the Preservation Costs Survey attempted to overcome them.

The monetary costs of preservation-related activity requires highly detailed investigation

Preservation can involve the use of automated systems for issuing holds and otherwise assisting the preservation of ESI. Anecdotal evidence from individual companies makes clear that companies may spend millions or even tens of millions of dollars on a single automated system to help manage some aspects of data preservation. Nonetheless, extensive data on these costs remains elusive, and no prior research has attempted to collect empirical data on monetary outlays of companies for preservation-related technologies. This is largely because of two challenges:

First, identifying systems and their cost requires time-consuming, individualized investigation of each company. Each company has different computer systems, different internal business flow, and different technology needs. For this reason, the preservation-related systems that are developed or purchased will vary from company to company. Nor is the true price tag of a discrete system easy to identify. While an “off-the-shelf” solution from an

¹⁴ This observation is consistent with the recent RAND study. See *Where the Money Goes* at xix (“Most interviewees did not hesitate to confess that their preservation costs had not been systematically tracked in any way and that they were unclear as to how such tracking might be accomplished.”).

outside vendor comes with an invoiced price, the full cost of that solution includes costs to company time and resources for design and bidding for the project, implementation, and maintenance over time. Systems that are developed partially or entirely in-house are even harder to price.

Second, individualized investigation is required to ensure that the costs being measured are properly attributable to preservation obligations, rather than other motivations. Not all hardware and software that stores, manages, or archives data exists solely or partly for compliance with preservation obligations. Indeed, even some tools which are acquired to mitigate the costs of preservation are not necessarily solely attributable to preservation obligations. For example, Company D reports spending approximately \$340,000 (plus about \$60,000 per year) for a license on a tool for collection and processing. While the tool is used for collection and processing, the stated need for this tool is to mitigate some of the costs of overbroad preservation. It is not necessarily clear how one should allocate the costs of such a system among the preservation, collection, and processing stages of discovery.

To address these challenges, the Preservation Costs Survey took a conservative approach. I relied on detailed, in-depth interviews with companies to accurately identify specific systems whose *sole* purpose was compliance with preservation obligations. (Thus, the Company D system described above is *not* treated as a cost of preservation.) And although all aspects of the costs of implementing these systems cannot be measured, the in-depth interviews ascertained the company's best estimates of the implementation and maintenance costs for these systems. As a consequence of this approach, my data on these costs generates a *conservative* estimate of the total costs of technologies adopted in response to preservation burdens.

Because the costs of automated systems to manage litigation holds or store preserved data do not depend on the number of individual litigation matters of a company, but instead represent costs incurred to address preservation obligations across all matters, I will refer to these as "fixed costs" or "non-case-specific" costs of preservation. Also in the category of fixed costs are the costs of maintaining legal, IT, and other professional staff who are devoted to handling a company's preservation obligations.

The human cost of preservation-related activity in terms of lost work time has never been measured

One major cost of preservation obligations is the cost in lost employee time spent complying with duties imposed through the issuance of litigation hold notices. This cost to employees not in the legal department of a company is in addition to the cost of maintaining a legal staff devoted to compliance with preservation obligations, including the issuance of litigation hold notices. Unlike the cost of maintaining litigation staff, which is a fixed cost for a company (i.e., not a cost attributable to any single matter), the time cost to

non-legal employees of dealing with litigation hold notices is case-specific, in that each litigation matter that generates a litigation hold creates a new set of litigation hold notices that require a new set of actions by employees in response to the hold. Indeed, a single employee may be subject to many separate litigation hold notices at any given time, each imposing different obligations in terms of the duration, date, and scope of data that she must preserve.

Because the cost of compliance with litigation hold notices is dispersed throughout a company, and because the cost primarily takes the form of lost time rather than monetary payments, measuring the magnitude of this cost is difficult. The time and energy that employees must divert towards preservation is never recorded or compensated, unlike the time spent by dedicated lawyers, such as outside counsel.¹⁵

No prior research study has attempted to measure the costs of non-legal employee time lost to preservation activities. In the Preservation Costs Survey, I measure this cost by collecting detailed information on the number of matters with litigation holds, and the number of employees subject to each litigation hold, at a sample of companies. I combine these counts of employees subject to litigation hold compliance obligations with estimates of time lost per employee, and the hourly cost of employee time, to quantify in dollar terms the value of employee time that is diverted from business purposes to compliance with preservation obligations.

Because the costs of employee time lost to litigation hold compliance depend on the size and number of litigation holds issued by a company, I will refer to the costs of employee time diverted from business activities as the “variable costs” or “case-specific costs” of preservation.

Other costs associated with preservation are diffuse and cannot be directly measured

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

¹⁵ Compare *Where the Money Goes* at 85 (“Part of the reason for a lack of existing information in this area appears to be that much of preservation involves expenditures incurred internally, such as the costs of IT staff time, law department attorney and paralegal time, other employees’ time (such as the effort required of custodians to comply with legal-hold notices), and purchases and licensing of applications and hardware to handle preservation.”).

irregularly and sometimes require ad hoc solutions. For example, departing employees leave behind data which is still subject to preservation obligations. Data may need to be collected from hard drives or loose media or identified and collected from collaboration-type information systems. For large companies, this is a serious burden, as the interviewed companies see thousands of employees leave each year.

Other issues arise less frequently, but are even trickier. Obsolete data formats or storage systems need to be replaced, and migrating data to new systems without the loss of data on hold can be difficult, requiring workarounds tailored to the specific systems. Basic business processes like rolling out new computers to employees is slowed or prevented because of concerns about the preservation of data stored on hard drives.¹⁶

These indirect costs of preservation obligations likely cost the interviewed companies millions in expenses and lower productivity, but measuring these diffuse costs and unobserved lost opportunities presents challenges so great that the Survey did not attempt to fully quantify them. Thus, the preservation costs measured by the Survey do not exhaust the universe of costs imposed by preservation obligations. The cost estimates in this Final Report, therefore, should be treated as a lower bound on the cost of preservation.

Companies are unable or reluctant to share sensitive and confidential information about litigation-related costs

A recurring obstacle to gathering data on the costs of preservation is the reluctance of companies to provide data on their costs.¹⁷ This is due in part to the fact that in many cases, the companies simply do not have the information, or cannot gather it at reasonable cost.¹⁸ But it is also largely due to an understandable fear that disclosing information about the company's

¹⁶ Compare *Where the Money Goes* at 86 (“[T]here may be economic impacts resulting from a decision not to adopt certain IT products (such as instant messaging or social-networking platforms) that might present significant difficulties when preserving information, from not implementing more-efficient data systems due to the need to maintain older legacy platforms and processes, from slower computer-system performance caused by halting the routine deletion of obsolete information in transactional databases, or from a reduced ability to recover lost but nevertheless important data due to a shift from a long-term data backup process to a short-term disaster-recovery system primarily because of preservation concerns.”).

¹⁷ See *Where the Money Goes* at 4.

¹⁸ The Preservation Costs Survey asked respondents whether they track the costs of litigation holds. Only 14 percent responded that they did.

litigation experiences and expenses could be used strategically against the company in litigation.

For this reason, all information collected for the Preservation Costs Survey was gathered subject to assurances of strict confidentiality for each survey participant. Several means of submitting survey responses anonymously were provided to participants, and all results reported herein are anonymous. In some cases, exact numbers are rounded or topcoded (e.g., employee counts larger than 100,000 are reported as "> 100,000") as an additional safeguard of anonymity. To ensure the integrity of the data and to avoid manipulation of the Survey, I checked response data for duplicate submissions and data inconsistencies before decoupling the data from company identifiers.

2.2 Survey Design

The Survey was commissioned by the Civil Justice Reform Group ("CJRG"), a group of in-house counsel at large, U.S. corporations.¹⁹ CJRG asked a number of large companies to participate in the Survey and coordinated with other associations of businesses (including small and medium-sized businesses) to request their members to participate in the Survey. CJRG has had no control over the design of the survey questions. It does not have access to survey responses or data collected in the course of the Survey. Nor was CJRG involved in the analysis of the data.

The Survey has involved the collection of three types of data. First, I conducted detailed, in-depth interviews with in-house attorneys, legal staff, and IT personnel to assess their experiences with preservation obligations, their estimates of the costs of preservation, and the extent of objective data on preservation costs that might be available to collect and analyze.

Second, I sought quantitative, company-specific data to be used for statistical analysis. This data would include data quantifying the volume of preservation activity at the individual matter level, as well as any available aggregate statistics on the volume of data involved in the different stages of litigation (preservation, collection, processing, review, and production).

Third, I developed a survey instrument to gather information about companies' experiences with preservation. The survey would collect basic information about the respondent company (e.g., industry classification, number of employees), its legal department (e.g., number of attorneys and number of litigation matters), and its preservation activity (e.g., number of litigation

¹⁹ I have been compensated for my time and expenses associated with designing the survey, interviewing respondents, and processing response data for the Survey.

holds and systems for managing litigation holds). It would also collect qualitative information on the company's experiences with preservation, including questions about the types of cases in which preservation-related issues are most prevalent.

Given the complexity of the topic, and the largely unprecedented nature of a study focused on preservation costs, I employed a three-phase study design. Each phase of the Preservation Costs Survey involved the gathering of information from companies on a strictly anonymous basis to ensure that responses were as candid and complete as possible.

Phase I involved a set of four, in-depth case studies of large companies. These case studies involved both qualitative interviews and requests for quantitative data to be used for statistical analysis. One important aspect of Phase I was developing the survey instrument. I began with an extensive written survey coupled with follow-up interviews to obtain feedback on the clarity and practicability of each question. This information was used to draft the survey instruments used with larger samples of companies during Phases II and III.

Phase II broadened the sample of companies to thirteen and continued to employ an in-depth, case-study approach. A revised questionnaire was combined with interviews and the collection of matter-level datasets of preservation activity in order to create as complete as possible a picture of the sources and amounts of preservation costs for large companies. (By "matter-level," I refer to datasets in which information on the number of litigation holds is provided for each individual litigation matter.) In addition to survey and interview responses, Phase II yielded six unique databases of matter- and employee-level preservation activity within specific companies. These databases of preservation activity were provided on a strictly confidential, anonymous basis. These datasets together provide information on over half a million litigation hold notices issued to individual employees in individual matters. They are the first large samples of case-specific preservation activity data ever compiled for research purposes.

Even with strict assurances of anonymity and data security, inducing participation in the Preservation Costs Survey required considerable time and effort. Companies expressed concerns about sharing sensitive, confidential data about their internal legal operations. For example, companies did not want to inadvertently compromise any confidential attorney work product or attorney-client communications. Thus, all interview and survey responses and statistical data are being reported only on a strictly anonymous basis. Further, all case-specific information provided by companies was redacted of all identifying information and stripped of any potentially privileged or work-product-protected content before being provided for use in the Survey. Given that the process of identifying, collecting, reviewing, redacting privileged material, and obtaining internal approval to disclose the data could involve

days or weeks of employee time at a company, as a practical matter, Phases I and II of the survey were limited to a smaller number of respondents.²⁰

Phase III involved a shortened survey questionnaire and no interviews or requests for data. This Phase was deliberately designed to be distributed to a larger number of companies, which would be able to respond with a much smaller investment in human resources. The goal of Phase III was to obtain survey responses from a large sample of companies, including small and medium-sized businesses, in order to draw inferences about the preservation activity of a broader cross-section of civil litigants. Phase III was publicized to companies through groups such as the National Association of Manufacturers, Lawyers for Civil Justice, and the Association of Corporate Counsel. The surveys could be completed on a printable form or by an online survey instrument hosted on research.net. The Phase III survey was open from October 2013 to January 2014.²¹ By the conclusion of Phase III, a total of 128 unique companies had completed survey questionnaires. The Phase III questionnaire is reproduced in the Appendix.

It is necessary to note that although this study is by far the largest and most detailed survey of preservation costs yet conducted, this study's methodology, by its very nature, cannot guarantee a representative sample of all companies with preservation obligations.²² Because sensitive and sometimes hard-to-gather data is the subject of this Survey, any study with this design will rely on the self-selection into the sample of companies who are willing and able to respond. Nonetheless, the Survey results provide several indications that the sample is not necessarily unrepresentative of the larger population of companies.

First, the results from each phase of the Survey are remarkably consistent with each other, despite substantial differences in the process by

²⁰ The RAND study for similar reasons was limited to eight companies. *Where the Money Goes* at xiii.

²¹ Two surveys which were distributed before the close of the Survey period were returned in February 2014. They are included in the results reported below. (Excluding them has little effect on the reported results.)

²² Compare *Where the Money Goes* at xiii–xiv (“We asked participants to choose a minimum of five cases in which they produced data and electronic documents to another party as part of an e-discovery request. . . . Because the participating companies and cases do not constitute a representative sample of corporations and litigation, we cannot draw generalizations from our findings that apply to all corporate litigants or all discovery productions.”).

which companies were solicited for participation and the degree of effort required by the companies to complete their participation. This suggests that the amount of effort required to participate is not strongly correlated with the characteristics of the company.

Second, many of the patterns that one would predict to see in the data based on strong *a priori* justifications do, in fact, appear in the data. For example, smaller companies have very few (often zero) litigation attorneys and report dramatically fewer active cases. This pattern might not emerge if only the most sophisticated (or most embroiled in litigation) smaller companies participated in the Survey.

Third, unlike prior studies which also depended on the willingness of companies to provide data on discovery costs (*Litigation Cost Survey*) or to provide interview responses on preservation (*Where the Money Goes*), the Preservation Costs Survey did not allow participating companies to select specific cases for inclusion in the sample. Rather, the questionnaire asked for information only about cases in the aggregate, and the requests for databases of preservation activity included all litigation matters with litigation holds (excluding asbestos cases). Thus, the Preservation Costs Survey provides analysis of the first *truly representative* samples of the within-company distribution of litigation activity.

3 Sample Characteristics

The 128 survey respondents represent a broad cross-section of companies in the United States. The participating companies come from a wide variety of industries, including every industry category listed in the Survey.²³ The most heavily represented categories were health care, insurance, technology, and conglomerate, each with at least 10 respondents.

The number of people employed worldwide by each company ranges from 18 to over 100,000.²⁴ Importantly, although large companies were the focus of Phases I and II, smaller companies are well represented in the sample. About a quarter of all respondents have 1,000 or fewer employees worldwide; the same proportion have 500 or fewer U.S. employees, the threshold usually used to define a small or medium-sized enterprise (“SME”). The largest companies, those with over 100,000 employees worldwide, make up about one-sixth of the sample.²⁵ See Table 1.

TABLE 1: RESPONDENT SIZE DISTRIBUTION²⁶

	Share	Number
Total Employees: 1–1,000	24%	31
Total Employees: 1,001–10,000	29%	37
Total Employees: 10,001–100,000	30%	39
Total Employees: 100,001+	16%	18

²³ The categories are: Automobiles & Parts; Banks; Chemicals; Conglomerate; Financial Services; Food & Beverage; Health Care; Industrial Goods & Services; Insurance; Media; Oil & Gas; Other; Personal & Household Goods; Retail; Technology; Telecommunications; Travel & Leisure; Utilities.

²⁴ In order to protect the anonymity of some respondents, exact employee counts above 100,000 are not reported.

²⁵ Herein, I will occasionally refer to companies with close to or more than 100,000 employees worldwide as “large companies.”

²⁶ Percentages do not sum to 100 due to rounding. Not all companies provided an exact count of total employees, but all companies provided sufficient information for categorization in Table 1.

The volume of litigation varies widely across these companies; the number of suits currently active varies from 0 to over 10,000.²⁷ Asbestos litigation was specifically excluded from the Survey.²⁸ There is great variation in the number of litigation holds that companies report to have active. As the duty to preserve may arise before a lawsuit is filed, the number of matters subject to litigation holds may be greater than the number of lawsuits. Conversely, a single litigation hold may suffice for a number of related lawsuits, and thus a company may have fewer litigation holds than lawsuits. The number of in-house litigation attorneys ranges from 0 to over 50.²⁹ Most in-house litigation teams are small—the median is 4, and 17 out of the 128 companies have no in-house litigation counsel. See Table 2.³⁰

One category of litigation that was anticipated to affect companies of all sizes was employment litigation, defined in the Survey to include employment discrimination and retaliation claims. Feedback from judges on results from Phase I (presented in the Survey's Preliminary Report) included a request that the Survey specifically ask about employment litigation, in order to generate results that would be more appropriate to extrapolate to smaller companies. The working assumption here is that individual employment discrimination suits should look fairly similar across companies of different sizes and industries. As Part 4.2 will show, this assumption is supported by the data. For employment litigation, we again see great variation across companies. Notably, however, the scope of preservation in employment cases is much smaller than it is across all litigation matters. This is unsurprising, as one might expect that from a discovery perspective, employment claims typically are relatively straightforward, requiring litigation holds for the aggrieved employee's (or ex-employee's) supervisors and immediate coworkers only, although some require a broader set of holds.

²⁷ In order to protect the anonymity of some respondents, exact counts of lawsuits and litigation holds above 10,000 are not reported. Five companies did not report numbers of suits and seven companies did not report numbers of matters with holds.

²⁸ While asbestos litigation remains an important part of the federal civil docket, it is *sui generis* with respect to preservation: at this point in the history of asbestos litigation, virtually every document in the possession of a company defendant that could possibly be relevant to asbestos claims has long ago been preserved and produced.

²⁹ In order to protect the anonymity of some respondents, exact counts of litigation attorneys above 50 are not reported. Three companies did not report number of litigation attorneys.

³⁰ For Total Employees and U.S. Employees, $N = 126$. Median numbers of employees are rounded by up to 1 percent to protect respondent anonymity.

TABLE 2: SURVEY DESCRIPTIVE STATISTICS

Panel A: Employees, Lawsuits, and Litigation Hold Matters				
	Mean	Median	Min	Max
Total employees	43,454	8,000	18	> 100,000
U.S. employees	21,678	6,100	0	> 100,000
In-house litigation attorneys	12	4	0	> 50
Active suits	1,399	33	0	> 10,000
Active employment suits	32	5	0	> 500
Open matters with holds	686	33	0	> 10,000
Open employment matters with holds	80	5	0	> 500
Panel B: Share with Preservation Resources or Practices				
Issues litigation holds notices			100%	
Has formal preservation policies			84%	
Tracks litigation holds and notices			63%	
Has e-discovery team			40%	
Has legal IT group			31%	

4 Results

This Part presents the findings of the Survey, organized around four main themes. First, respondents reported significant preservation burdens. Indeed, for the largest companies, I estimate total preservation costs of over \$40 million per company per year.

Second, smaller companies report preservation burdens comparable to larger companies. This is especially true for types of claims, such as employment discrimination, where cases tend to look similar for companies of different sizes. Although smaller companies face fewer lawsuits, they are also less able to make large-scale investments in technology to control the costs of preservation.

Third, a small percentage of litigation matters generate a disproportionate share of case-specific litigation costs. Using matter-level data on litigation holds issued within six companies, I find that over half of all litigation hold activity takes place in 5 percent of matters.

Fourth, most data that is preserved is never collected, processed, and reviewed. In interviews, companies explain that the uncertainty under the current case law on preservation leads companies to “overpreserve” data.

4.1 Companies Report Significant Preservation Burdens

Consistent with the great weight of anecdotal evidence and prior qualitative studies, surveyed companies generally report significant preservation burdens, although a small fraction of companies report little or no preservation burdens. Beyond this general finding of substantial problems arising out of preservation obligations, the Survey reveals fairly systematic patterns in the types of litigation matters and data that generate the greatest problems.

Preservation problems arise frequently

Survey respondents generally reported that preservation generated difficulties or burden in litigation.³¹ Over 79 percent (102 of 128) of respondents reported a “great extent” or “moderate extent” of burdens from preservation activity. See Table 3.³²

³¹ Compare *Where the Money Goes* at xix (“All interviewees reported that preservation had evolved into a significant portion of their companies’ total e-discovery expenditures.”).

³² $N = 128$.

TABLE 3: OVERALL EXTENT OF BURDENS FROM PRESERVATION³³

Difficulties or Burdens from Preservation	Share
Great extent	41%
Moderate extent	38%
Small extent	14%
Not any extent	6%

Federal cases are a source of concern for companies of all sizes

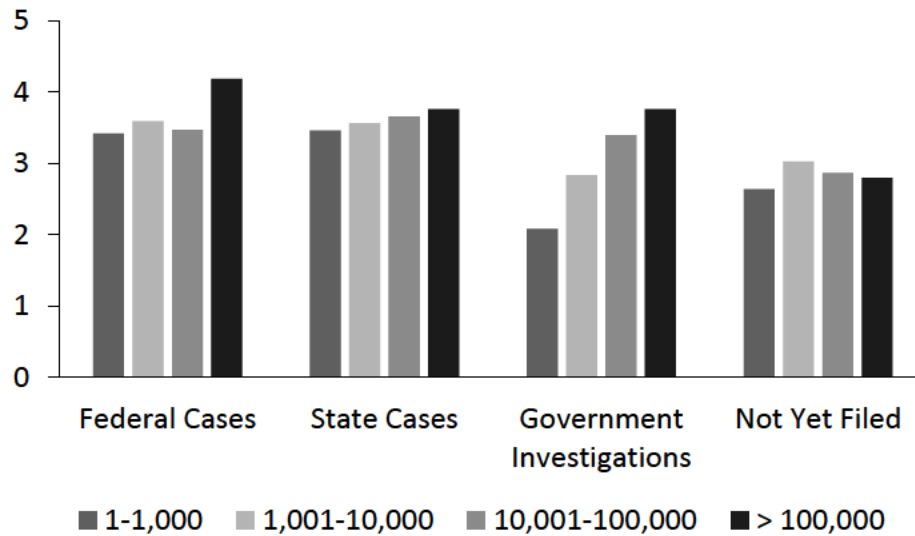
The Survey reveals that not all contexts for preservation activity are created equal. For companies of all sizes, filed actions and particularly filed actions in federal court generate preservation burdens with the greatest frequency. Figure 1 illustrates the average reported frequency of preservation problems and burden in different settings, broken down by company size.

As Figure 1 makes clear, larger companies differ from smaller companies in two notable respects. First, even more so than for other companies, the largest companies perceive *federal* courts as the locus of more frequent preservation-related burdens. Second, while smaller companies are not as greatly concerned by government investigations, for larger companies, preservation obligations in connection with government investigations are a major source of concern.³⁴

³³ Percentages do not sum to 100 due to rounding.

³⁴ One interviewee noted: “A lot of our angst comes from subpoenas . . . and administrative requests” that have no end date and no clarity about scope. Company H Interview 7/12/2013. Another interviewee said that with government investigations, “proportionality is out the window,” due to compressed time frames, unclear parameters for the scope of the investigation, and no indication from the government whether and when the investigation has concluded. Company M Interview 9/24/2013.

FIGURE 1: REPORTED FREQUENCY OF PROBLEMS OR BURDENS FROM PRESERVATION, BY FORUM TYPE AND COMPANY SIZE
(5 = "VERY OFTEN" AND 1 = "VERY RARELY")



Complex cases and cases involving individuals generate the greatest incidence of preservation problems

A natural hypothesis in studying the incidence of preservation costs is that they are disproportionately great in the largest, most complex cases. This hypothesis is supported by the data. Respondents report that large, complex cases generate preservation-related problem more frequently than smaller, routine matters. See Table 4. These results are statistically significant.³⁵ Large, complex matters have the potential to generate a far greater scope of preservation and greater difficulty for companies to identify relevant data sources for preservation. Results below (see Part 4.3) further substantiate this hypothesis, as I find that the majority of the human costs associated with preservation activities are imposed by only a small percentage of litigation matters.

Less obvious are the relative burdens of business-versus-business litigation and individual-versus-business litigation. One hypothesis would be that cases pitting companies against each other will involve the most extensive

³⁵ I test the hypothesis that the means for large, complex matters with individuals and for small, routine matters with individuals are the same with a paired, two-tailed t-test. The null hypothesis is rejected at the 1 percent level. I repeat the test to compare means for matters involving businesses, and again the null hypothesis is rejected at the 1 percent level.

“paper trails” (or, more accurately, “data trails”) and the most sophisticated clients, both of which will lead to greater controversies over preservation and a broader scope of preservation. An alternate hypothesis would be that, as individuals rarely have significant amounts of data to preserve, the highly asymmetrical nature of preservation obligations in litigation between companies and individuals will give individual litigants the incentive to press preservation issues and push for sanctions without concern that such obligations will be enforced against the individuals. The data support the latter hypothesis.

TABLE 4: PRESERVATION-RELATED PROBLEMS BY OPPOSING PARTY TYPE
(5 = “VERY OFTEN” AND 1 = “VERY RARELY”)³⁶

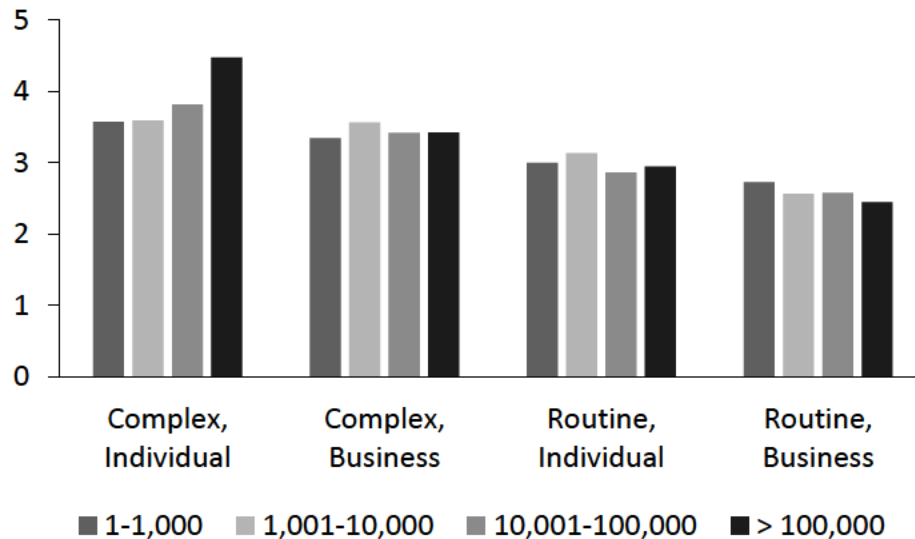
Configuration of Parties	Average Rating
Large, complex matters, individuals on other side	3.81
Large, complex matters, businesses on other side	3.45
Small, routine matters, individuals on other side	2.99
Small, routine matters, business on other side	2.59

As Table 4 shows, litigation in which individuals are on the other side (usually, but not always, as plaintiffs) generate higher levels of reported preservation-related problems than matters with businesses on the other side. These differences are statistically significant.³⁷ Importantly, while most of the anecdotal accounts of asymmetrical litigation generating preservation burdens comes from large companies, the pattern reported in Table 4 holds for companies of all sizes. See Figure 2.

³⁶ $N = 122$.

³⁷ I test the hypothesis that the means for large, complex matters with individuals and for large, complex matters with businesses are the same with a paired, two-tailed t-test. The null hypothesis is rejected at the 1 percent level. I repeat the test to compare means for small, routine matters, and again the null hypothesis is rejected at the 1 percent level.

FIGURE 2: PRESERVATION-RELATED PROBLEMS BY OPPOSING PARTY TYPE
(5 = "VERY OFTEN" AND 1 = "VERY RARELY")



Email and hard drives are the most common sources of preservation difficulties

Regarding data types, respondents consistently reported the greatest incidence of preservation-related problems for email and hard drives.³⁸ This is probably a function of their ubiquity across firms of all sizes and their frequent salience in litigation. See Table 5.³⁹

Notably, this pattern persists across companies of all sizes. Figure 3 illustrates the averages for each data type, broken down by companies with different numbers of employees.

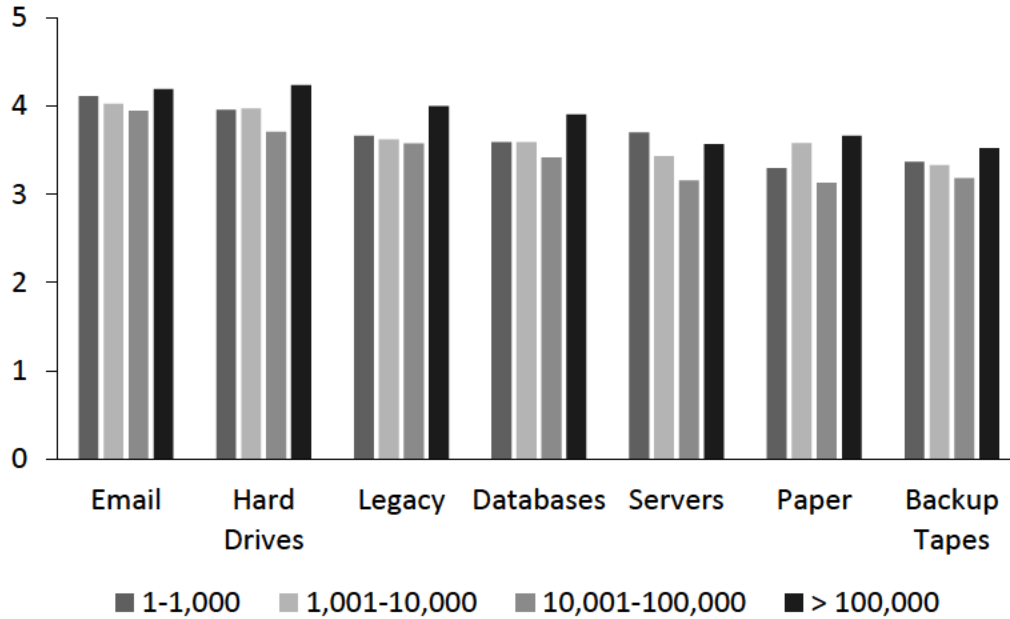
³⁸ The means for email and hard drives are statistically significantly different from the means for the remaining data types at the 5 percent level. I test for equality of means using paired, two-tailed t-tests. The null hypothesis is rejected at the 5 percent level or better when comparing email and hard drives to each of the other data types.

³⁹ $N = 122$ for hard drives, paper documents, and backup tapes. $N = 123$ for all other categories.

TABLE 5: PRESERVATION-RELATED PROBLEMS BY DATA TYPE
(5 = “VERY OFTEN” AND 1 = “VERY RARELY”)

Data Type	Average Rating
Email	4.05
Hard drives	3.93
Legacy data	3.68
Databases	3.59
Central servers	3.43
Paper documents	3.39
Backup tapes	3.33
Collaboration tools	3.28

FIGURE 3: PRESERVATION-RELATED PROBLEMS BY DATA TYPE
(5 = “VERY OFTEN” AND 1 = “VERY RARELY”)⁴⁰



⁴⁰ The “collaboration tools” category is omitted in the interests of space. The pattern by size is consistent with the remaining categories.

For large companies, the fixed (non-case-specific) costs of preservation are high

As discussed in Part 2.1, the full costs of preservation include the “fixed costs” of hiring personnel and implementing the hardware and software systems to manage preservation activities across all matters. Companies invest in these fixed costs to lower the per-matter cost of preservation activities through the application of legal and technical expertise and automation.

While collecting specific estimates of these fixed costs was infeasible for Phase III of the Survey, I collected information on these costs in interviews with a number of large companies in Phases I and II. (As explained below in Part 4.2, smaller companies generally make fewer fixed-cost investments to control the costs of preservation.)

The first type of fixed cost is personnel. Legal IT departments in particular, and to a lesser extent e-discovery groups, are devoted primarily—and in many cases nearly exclusively—to preservation activity. These groups are composed of lawyers, paralegals, IT specialists, and other professionals whose efforts have been diverted from other aspects of their company’s business and legal affairs. Having technology and legal professions devoted full-time to preservation is costly to companies, but the benefits for the companies that create these groups outweigh the costs. By having full-time, in-house personnel devoted to preservation, these companies cultivate expertise with their company’s preservation needs, and pursue policies, training, and automated system development that reduce the total costs of complying with preservation obligations. And because preservation efforts are centralized and systematized, preservation efforts in any given case are more defensible in the face of a motion for sanctions. (In contrast, for most smaller companies, it will not be cost effective to hire personnel with this expertise and dedicate them to preservation support. See Part 4.2.)

The efficiencies from personnel specialized in preservation are substantial, as a relatively small number of individuals are able to coordinate preservation activities for a large number of litigation matters. Companies that participated in Phases I and II of the Survey and had legal IT or e-discovery groups reported the numbers of full-time attorneys and paralegals in these groups. On average, the companies had 2 attorneys and 4 paralegals or other legal professionals working full-time in a dedicated legal IT or e-discovery group.

The second type of fixed cost is technological. Estimates of the costs of individual computer systems (both hardware and software) implemented by those companies to handle aspects of preservation range from hundreds of thousands to tens of millions of dollars *per system*. One important type of system in this area is the automated litigation hold management system. These systems automate the process of distributing, tracking, and monitoring

litigation hold notices that are created by in-house counsel. While not all companies have precise cost estimates for these systems, Company A implemented a system to partially automate the issuing and tracking of litigation holds at a cost of approximately \$900,000. Company B estimates the cost of its new system to be \$800,000. In addition to implementation costs are upkeep, staffing, and maintenance costs, in the ballpark of \$150,000 per year.

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, are used and stored in different ways on a company's computer systems and are not usually integrated into a unified system. To preserve all types of ESI, therefore, requires multiple preservation solutions.

Precisely because there are no uniform solutions to preservation problems, identifying those technological solutions that *specifically* address preservation obligations requires detailed interviews. In Phases I and II of the Survey, I undertook this in-depth investigation and identified some (but not all) of the costs of automated preservation systems. For example, the tools used by Company A to collect data to be preserved at the outset of litigation 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. Testimony at the February 7, 2014 public hearings provides an additional data point. Michael J. Harrington, General Counsel of Eli Lilly and Co., stated that his company spent \$40,000,000 on email archiving systems for preservation purposes.

These reports suggest that for large companies, the per-year fixed costs associated with preservation activity run into the millions of dollars. Table 6 presents a rough calculation using numbers on the *low* end of the ranges reported by companies. The total (measurable) fixed costs of preservation for a single, large company exceed \$2.5 million per year.

TABLE 6: APPROXIMATE FIXED COSTS OF PRESERVATION
FOR A HYPOTHETICAL LARGE COMPANY

Preservation Solution	Per Year Fixed Cost
A legal IT and/or e-discovery team	\$1,000,000
Litigation hold management system (implementation cost amortized over a 5 year expected life)	\$160,000
Maintenance of litigation hold management system	\$150,000
Automated data preservation system (implementation cost amortized over a 5 year expected life)	\$1,000,000
Maintenance of automated data preservation system	\$200,000
Total	\$2,510,000

For all companies, variable (per-case) costs associated with preservation are significant

As high as the fixed costs of preservation may be, the largest share of preservation costs are the costs in human time and effort to address preservation obligations on a case-by-case basis. While sophisticated investments in expert personnel and automated systems can streamline the process of preserving data and reduce per-case costs, it remains the case that individual employees who are placed on hold or otherwise asked to engage in preservation activities must divert time and attention away from normal business activities. In this respect, a day spent responding to litigation hold notices is just as significant a drain on worker productivity as a day spent sick with the flu, for example.

Prior to this Survey, however, the magnitude of this aspect of preservation costs was completely unknown. This Survey therefore provides the first estimates of the cost, measured by the value of lost employee time, of preservation activity. Phase III of the Survey asked each company to identify the number of separate matters with currently active litigation holds. It did not ask for the number of employees on hold, due to the determination in Phases I and II of the Survey that many companies (particularly those without automated hold tracking systems) will not be able to provide that information without time-consuming internal investigation. The data from the companies that participated in Phase I or Phase II, however, allowed me to determine

the number and distribution of employees on hold among those large companies.

Averaging across the total number of matters and employees on hold in this sample, I find that there are an average of 185 employees receiving notices per litigation matter. (Of course, some holds may affect all the employees in a given company, while other holds may affect only a handful.) These sampled companies have on average approximately 150,000 employees worldwide. This implies that for the largest companies, about 0.12 percent of employees on average are subject to a hold for each litigation matter. To generate a conservative estimate of the numbers of employees on hold at smaller companies, I use this 0.12 percent of employees figure, but impose a minimum average of 5 employees on hold per matter at companies in the smallest size category (1 to 1,000 employees).⁴¹ This is because, even for the smallest companies, it would make little sense if the number of employees on hold averaged less than five.⁴²

I then multiply the reported numbers of matters with litigation holds by the number of employees per matter to obtain an estimate of the total number of employee-holds per company. (Note that in practice, some employees may be subject to no holds, while other employees may have to comply with a large number of distinct holds.)

To estimate the total amount of employee time lost to complying with litigation hold notices, I must estimate the time spent per employee per year complying with a litigation hold. This time includes time spent reading, confirming receipt, and asking questions about a litigation hold notice; time spent changing personal device settings and other work practices to comply with the litigation hold notice; and time spent reviewing electronic and paper files to mark, copy, or set aside files for preservation.

I use an estimate of 3 hours per employee per year spent on each litigation hold. This number is intended to be a conservative estimate, and it is based on estimates reported by interviewed companies. (No surveyed company had a precise estimate, however. As noted in Part 2.1, this type of cost is

⁴¹ I call this method conservative because one would expect the percentage of employees subject to any given hold to rise as company size falls. To be more precise, my formula multiplies the midpoint of each size range times 0.0012 and adds 4.5. This is calibrated to set a floor at 5 for the smallest companies and to generate 185 for a company with 150,000 employees. For example, for the size range 1,001–10,000, I multiply $5500.5 \times 0.0012 + 4.5 \approx 11$. I use 185 for companies in the 100,001+ range.

⁴² As noted in Part 4.2 below, using employment litigation as a benchmark for comparing smaller and larger companies suggests that this minimum of five is quite conservative.

not generally tracked by companies.) It is also smaller than the estimate for time spent on litigation holds per employee provided in a public comment by Exxon Mobil, which undertook its own estimate of employee time lost due to hold activity.⁴³

I then multiply this number by my estimate of the average cost of employee time: \$52.20/hour, which is the average hourly wage of workers in management occupations (across all sectors and all business sizes) provided by the latest data from the Bureau of Labor Statistics.⁴⁴ This is a *conservative* estimate of the cost of employee time, because this figure includes only wages and not benefits (nor does it include the lost value of equipment and capital from lost employee time). Indeed, this estimate is about 30 percent lower than an estimate based on detailed salary and benefits data provided by a Phase II company. As with all of the parameters used in this calculation, I use the lower, publicly available figure for the sake of a *lower-bound* estimate.

The results appear in Table 7. For companies of all sizes, the costs in lost employee time are significant. For the smallest companies in the Survey, the costs average over \$12,000 per company per year. The estimate of costs for the largest companies exceeds \$38.6 million per company per year.⁴⁵ Indeed, despite the fact that larger companies spend millions per year on legal and IT staff and automated preservation systems, these fixed costs are dwarfed by the cost of lost employee time due to litigation hold notice obligations. Including both fixed and variable costs, the average costs of preservation for the largest companies in the sample to exceed \$41.1 million per company per year.

⁴³ See Testimony of Robert L. Levy of Exxon Mobil Corporation, *Transcript of Public Hearing on Proposed Amendments to the Federal Rules of Civil Procedure* at 159–161 (Nov. 7, 2013), available online at <http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/public-hearings/civil-hearing-transcript-2013-11-07.pdf>. Mr. Levy estimated 10 minutes per day per employee spent on litigation hold notice compliance. Note, however, that as a given employee may be subject to multiple holds, this time estimate reflects the total time spent on *all* litigation holds that apply to a given employee. My estimate of 3 hours per year (which is less than 1 minute per day) is the time spent on *each* litigation hold that applies to a given employee.

⁴⁴ See Bureau of Labor Statistics, *May 2012 National Occupational Employment and Wage Estimates*, available online at http://www.bls.gov/oes/current/oes_nat.htm.

⁴⁵ In a letter to the Hon. David G. Campbell, I provided initial estimates of the costs of employee time for preservation activities based on publicly disclosed data from the Microsoft Corporation and the estimate of \$70/hour for employee time based on Phase II data. William H. J. Hubbard, *Letter to the Hon. David G. Campbell* (Nov. 3, 2011). Data from Phase III reveals that, among large companies, Microsoft is in fact on the low end of the range of preservation costs.

TABLE 7: ESTIMATED PER-COMPANY COSTS OF EMPLOYEE TIME
LOST TO LITIGATION HOLDS, BY COMPANY SIZE⁴⁶

(1)	(2)	(3)	(4)	(5)
Total Employees	Matters with Holds	Employees per Matter	Employee Hours Spent per Year	Cost of Lost Time per Year
1–1,000	16	5	240	\$12,528
1,001– 10,000	249	11	8,217	\$428,927
10,001– 100,000	1,245	71	265,185	\$13,842,657
> 100,000	1,333	185	739,815	\$38,618,343

Preservation costs are perceived as leading to unjust litigation outcomes

Although relatively few Survey participants provided comments in addition to questionnaire responses, among those that did, a repeated refrain was that costs associated with preservation and discovery are not merely placing burdens on companies, but they are affecting the outcomes of cases. Here are examples of what some respondents said:

The cost of discovery in IP cases results in early suboptimal settlement and for patent troll cases results in payoff settlements mainly to avoid protracted expensive discovery.⁴⁷

The costs and burdens (e.g., fear of sanctions despite best efforts and extreme use of internal resources and cost of external resources) of discovery and preservation have become an important factor in whether to litigate or settle.⁴⁸

⁴⁶ Column (4) is the product of column (2) and column (3) times 3 hours per employee-hold. Column (5) is column (4) times \$52.20 per hour.

⁴⁷ Respondent 2866719471, Phase III Survey.

⁴⁸ Respondent 2866957321, Phase III Survey.

Desperately need reasonable requirements that are fairly and reliably enforced by all courts. Also need mechanism to curtail the use of e-discovery as a litigation weapon rather than a discovery device. Serious issue becoming increasingly problematic for businesses and their in-house legal departments.⁴⁹

These subjective impressions gain some support in quantitative data gathered by Emery Lee of the Federal Judicial Center. In a recent and oft-cited study, Lee found that *motions* related to spoliation are filed in only 0.15 percent of federal cases.⁵⁰ Lee points out, however, that the *Civil Rules Survey* found that a party raised a claim of spoliation—even if no motion was filed—in 3 percent of plaintiffs’ cases and 2 percent of defendants’ cases. These numbers jump to 7.7 percent and 5 percent, respectively, in cases involving ESI.⁵¹ Thus, preservation and discovery of ESI are associated with higher rates of claims of spoliation.

⁴⁹ Respondent 3I, Phase III Survey.

⁵⁰ 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 4* (FJC 2011).

⁵¹ *Id.*

4.2 Preservation Challenges for Smaller Companies

Most of the anecdotal evidence regarding preservation costs that has been offered to the Advisory Committee at public hearings or in written submissions has come from larger companies. Thus, one can question the extent to which smaller companies' experiences with preservation obligations resemble larger companies' experiences. As Part 4.1 indicates, smaller and larger companies both gave similar responses to questions asking the extent and frequency of problems and burdens arising from preservation obligations.

Nonetheless, it remains the case that smaller companies do differ from larger companies for the simple reason that they face fewer lawsuits. See Table 8. One consequence of this is that smaller companies invest much less than larger companies in technology and human expertise to address preservation costs. As discussed below, this has the consequence of making smaller companies more vulnerable to preservation burdens when they do arise.

TABLE 8: MEAN NUMBER OF ACTIVE LAWSUITS, BY COMPANY SIZE

Total Employees	Active Lawsuits
1–1,000	27
1,001–10,000	228
10,001–100,000	2,563
> 100,000	3,404

The fact that smaller companies have smaller numbers of litigation matters also calls into question the extent to which the detailed, litigation-matter-level data collected from large companies in Phase II is informative of the types of litigation matters faced by smaller companies. Below, I discuss how an examination of employment litigation matters, which are likely to be similar across companies of all sizes, suggests that results from larger companies can be extrapolated to smaller companies as well.

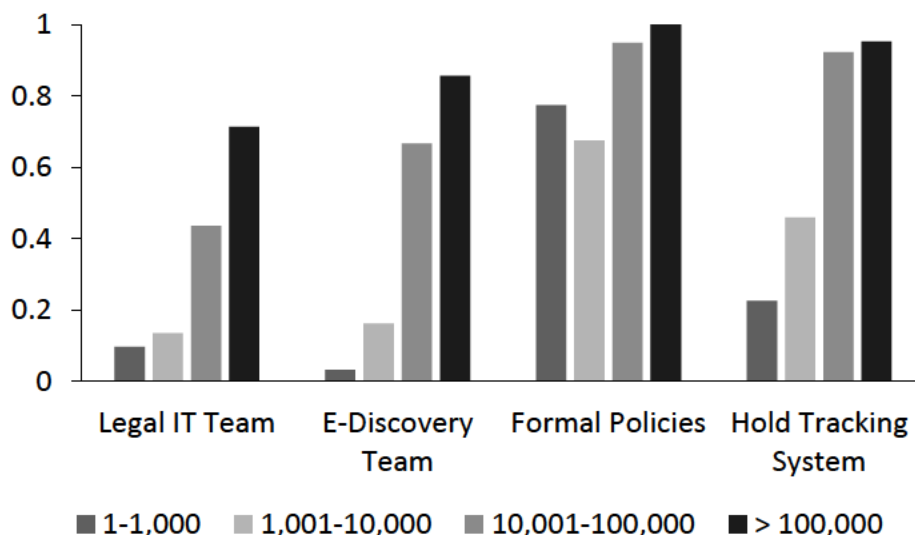
Smaller companies are more vulnerable to preservation costs

As noted above, in many respects, smaller companies report similar burdens when compared to larger companies. Many large companies respond to the burdens of preservation obligations by making investments that benefit from large economies of scale: automated litigation hold tracking systems, company-wide preservation policies and training, and full-time legal and IT

personnel dedicated to preservation.⁵² As discussed above in Part 4.1, these investments are very costly for large companies. For small companies, such investments may simply be impossible.

The Survey results bear out this concern. See Figure 4. While the very largest companies almost uniformly gain the benefits of in-house preservation experts for managing litigation holds, virtually none of the smallest companies have separate legal IT or e-discovery staff. Similarly, automated litigation hold tracking software is virtually standard practice among large companies, but this technology is uncommon among smaller companies. Formal preservation policies, which aid companies in defending against claims of spoliation, are also somewhat less common among smaller companies.

FIGURE 4: SHARE REPORTING PRESERVATION PRACTICES, BY COMPANY SIZE



These results suggest that many smaller companies may be *more* vulnerable to preservation burdens and spoliation claims than larger companies. Precisely because they are smaller and face fewer lawsuits, it is not generally cost effective for smaller companies to make expensive, but beneficial, investments in sophisticated automated systems or in-house expertise.

One smaller company respondent emphatically expressed this sense of vulnerability:

⁵² For example, Company D spent around \$1 million to implement and maintain software to assist in indexing and searching preserved data, but the interviewed employee of Company D saw this cost as a fraction of the savings it has generated.

Our relatively small company has been embroiled in seven years of patent litigation with a patent troll competitor. Discovery is ongoing during this entire time. In fear of all the outrageous sanctions against companies for e-discovery mistakes, we dare not delete any records from our systems until this matter is completely over. . . . Piecemeal deleting of non-relevant records like emails is not practical either and is fraught with risk of costly, embarrassing sanctions if someone makes a mistake along the way.⁵³

This conclusion places a damper on the hope that continued technological innovation will be sufficient to alleviate the costs of compliance with preservation obligations. Because technological solutions generally involve expensive systems that are most cost-effective when employed at large scale, smaller companies are poorly situated to reap the benefits that new technologies may provide in this area. This is true even for companies in the technology field itself; one respondent, a tech company with about 100 employees, explained:

We are a small company, but we are in a space where we need to protect our IP and also to prevent customers from eluding payment. We manage most of the process in house, but it is a huge burden on our IT. We are looking at vaulting solutions for e-mail, which should be a big help. But the costs are enormous, and vendors are unwilling to give us a good demo or trial vault.⁵⁴

Instead, smaller companies may have to rely on ad hoc, outside assistance, which may be less efficient and more expensive on a per-case basis. One respondent, an industrial company with about 200 employees, explained:

Our company along with every other company in our industry is involved in several suits concerning one toxic tort-related issue. We are a very small player in this field. Yet, we have to produce the same documents as the big guys. In our case, our IT employee, our President, our Accountant, our Attorney, etc. has to devote all of their time to answer discovery. We also employ an outside law firm at an hourly rate to help us. It is very costly.⁵⁵

⁵³ Respondent 2865842363, Phase III Survey.

⁵⁴ Respondent 2865509178, Phase III Survey.

⁵⁵ Respondent 2867300205, Phase III Survey.

Employment cases provide a benchmark for extrapolating from larger to smaller companies

Employment cases, defined in the Survey as employment discrimination and retaliation cases, tend to involve a relatively small number of “key players.” For this reason, even employment cases in very large companies tend to have a modest scope in terms of litigation holds, a fact confirmed by detailed litigation hold data from Phase II survey participants. See Table 9.⁵⁶

TABLE 9: EMPLOYEES ON HOLD PER MATTER, SELECTED MATTER TYPES, SAMPLES FROM TWO PHASE II COMPANIES⁵⁷

Type	Employees on Hold	
	Mean	Median
Company A		
Contracts	44	58
Environmental	71	38
Personal injury	24	16
Employment	24	10
Company K		
Patent/product liability	9,095	> 10,000
Gov’t Investigation/subpoena	2,962	100
Other	2,771	136
Employment	40	8

For this reason, this type of litigation is more likely to have similar characteristics across company size than other types of litigation that may be concentrated among large companies (for example, antitrust litigation). By looking specifically at employment cases, then, we can better perform an apples-to-apples comparison of the responses of larger and smaller companies. To the extent that the responses of larger and smaller companies are

⁵⁶ To protect respondent anonymity, not all matter types reported are displayed. (Unreported categories are consistent with reported results.)

⁵⁷ The Company A data includes 85 matters in the 4 listed categories. The Company K data includes 102 matters in the 4 listed categories.

consistent, this will in turn allow us to more confidently extrapolate our results based on detailed litigation hold data from large companies to companies of all sizes. As shown herein, the responses for employment cases are, in fact, consistent across companies of different sizes.

The Survey results reveal that respondents consistently report that employment cases are a relatively more frequent source of preservation-related burdens. See Table 10. While it might seem surprising that employment litigation would generate greater preservation-related problems than antitrust litigation, one must remember that the numbers in Table 10 reflect an average across all company sizes. Not surprisingly, smaller companies are largely unconcerned with certain types of higher-stakes litigation, such as antitrust, product liability, and consumer protection. See Figure 5. Thus, the high rating for employment cases is informative not because it shows that employment cases are more burdensome than other types of cases, but rather that employment cases are *consistently burdensome across companies of all sizes*. (In this regard, contract cases appear to generate similar levels of preservation burdens across companies of all sizes as well.)

TABLE 10: PRESERVATION-RELATED PROBLEMS BY CLAIM TYPE
(5 = “VERY OFTEN” AND 1 = “VERY RARELY”)⁵⁸

Preservation Type	Average Rating
Employment	3.34
Contract	3.26
Product liability	2.68
Consumer protection	2.58
Antitrust	2.55

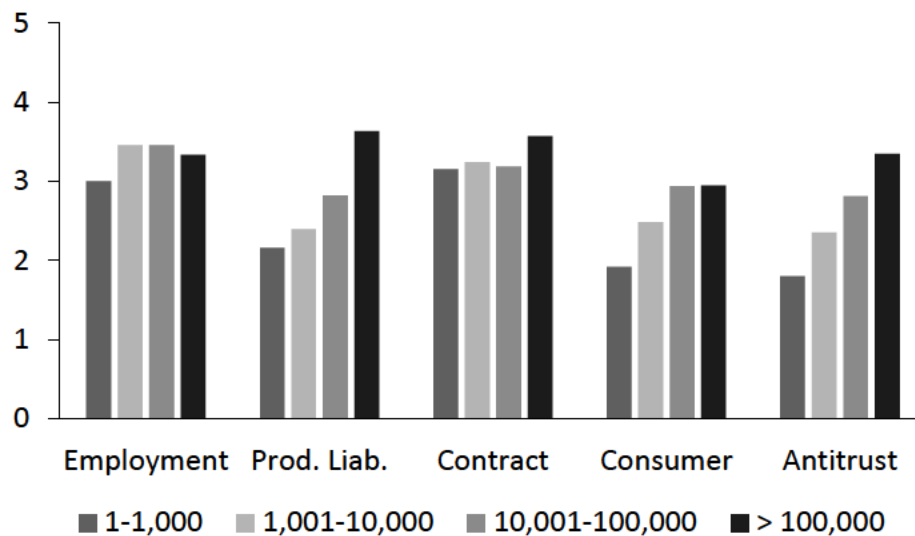
A number of results in this Final Report are based on detailed case-level data collected from Phase II companies, all of which are larger companies. The patterns in this data appear both across all cases and across only employment cases. This similarity between the patterns for all cases and for

⁵⁸ Not all companies provided responses to each category. For employment and contract, $N = 119$. For product liability, consumer protection, and antitrust, $N = 109$, 107, and 106, respectively.

employment cases provides assurance when extrapolating results for smaller companies using this data from large companies.

Indeed, the employment litigation data suggest that the estimates in Part 4.1 of the variable costs of preservation for smaller companies are very conservative. The estimates in Part 4.1 assume the average number of employees on hold per matter is 5 at the smallest companies (those with 1 to 1000 employees). The data in Table 9 on the number of employees on hold in employment matters, however, suggest this number should be higher. For the companies in Table 9, the *median* numbers of employees is 8 or 10. (I use median numbers to reflect typical, routine employment disputes and to exclude the influence of large, complex employment matters.) Using these median numbers indicates that the estimates of variable cost for the smallest companies in the sample could easily exceed \$20,000 per company per year.

FIGURE 5: PRESERVATION-RELATED PROBLEMS BY CLAIM TYPE
(5 = "VERY OFTEN" AND 1 = "VERY RARELY")



4.3 A Small Share of Matters Accounts for Most Litigation Hold Activity

One question that existing evidence, which is entirely anecdotal, cannot answer is whether the cases that have enormous preservation costs are typical or atypical. On the one hand, examples of cases with extremely large preservation costs are abundant. On the other hand, some practitioners have expressed skepticism, noting that the *Civil Rules Survey* found that the median costs of discovery were quite modest in closed civil cases in federal court. While the *Civil Rules Survey* did not attempt to measure preservation costs, but only measured the costs of litigation and discovery that were observed by the counsel litigating the case (and not the costs incurred by the client directly), this nonetheless is some evidence that high discovery costs are not the norm. It is natural to wonder, therefore, whether the same is true for preservation costs.

As I explain below, it is in fact the case that (1) litigation matters with very high preservation costs are *not* typical, but (2) despite being infrequent, these matters are a major part of the total costs of preservation activity.

In most litigation matters, preservation scope is not broad, but in a fraction, it is extremely broad

Six of the companies participating in Phase II of the Survey provided data from their litigation hold tracking systems on matters for which litigation holds were issued. The data from Company D is representative and is illustrated by Figure 6.⁵⁹ The Company D data 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. As Figure 6 shows, most litigation matters involve litigation holds affecting relatively few employees—well over half of the matters had twenty holds or fewer. (The left-most vertical bar in Figure 6 represents the number of matters with 20 employees or fewer on hold.)

Yet, the distribution of litigation holds across matters is highly skewed, and there is a “long tail” of matters in which huge numbers of employees are placed on hold in each case. (For graphical clarity, the distribution of the number of employees on hold per matter in Figure 6 is topcoded at 500. Matters with more than 500 employees subject to hold are included in the right-most vertical bar.) This means that a small percentage of litigation matters can account for the bulk of all litigation hold activity. Indeed, in the

⁵⁹ Figures illustrating the distribution of litigation holds per case for the remaining five companies appear in the Appendix.

Company D sample, 5 percent of the matters account for more than 62 percent of the hold notices issued. As Table 11 shows, Company D is not unusual. Across the six companies for which data is available, 5 percent of litigation matters account for over half of all litigation hold notices.

FIGURE 6: DISTRIBUTION OF EMPLOYEES ON HOLD PER MATTER, COMPANY D SAMPLE, TOPCODED AT 500

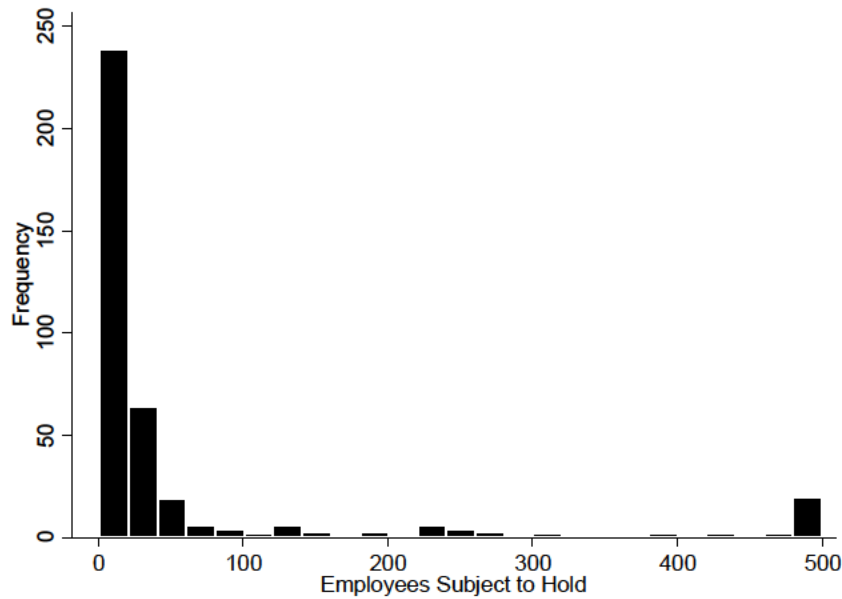


TABLE 11: EMPLOYEES ON HOLD PER MATTER, ALL MATTER TYPES

Company	Median	Max	Top 5 Percent Share of Total
A	25	> 500	28.1
C	11	> 500	54.5
D	17	> 5,000	62.5
E	75	> 10,000	52.3
G	8	> 10,000	86.5
K	15	> 10,000	28.1
Average			52.0

The distribution for employment matters is very similar

Although not all companies that provided data had indexed matters by claim type, three companies did. Looking only at employment matters, the top 5 percent of matters account for almost exactly half of all litigation holds, a share very close to the share for all cases presented above. See Table 12.⁶⁰

TABLE 12: EMPLOYEES ON HOLD PER MATTER, EMPLOYMENT MATTERS

Company	Median	Max	Top 5 Percent Share of Total
A	10	> 50	22.8
D	27	> 1,000	48.3
K	8	> 1,000	77.5
Average			49.5

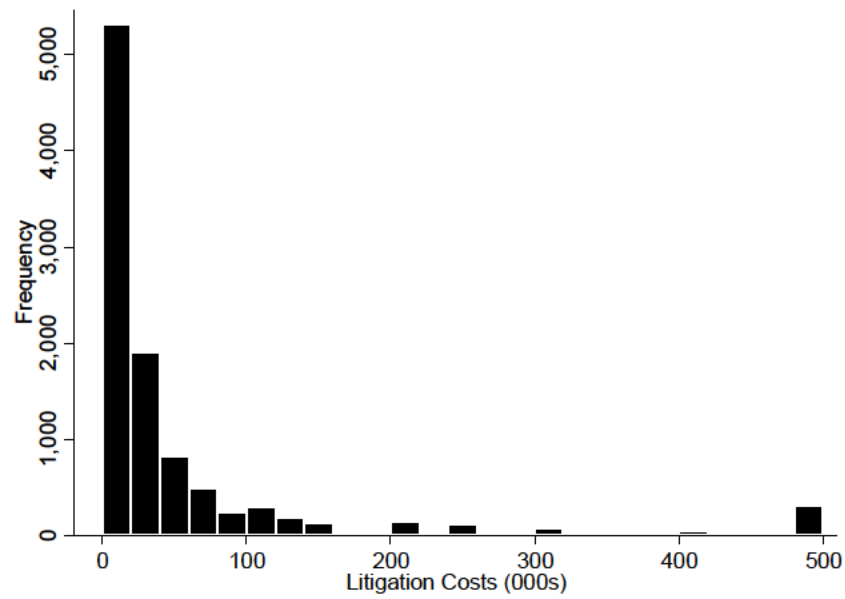
The distribution of litigation costs across cases is very similar to the distribution of preservation costs

In this regard, it is worth noting that the patterns that appear in the Preservation Costs Survey data resemble the patterns for litigation costs found in the *Civil Rules Survey*. Data collected for the *Civil Rules Survey* revealed that while the *median* case had relatively low litigation costs (\$15,000 for plaintiffs and \$20,000 for defendants), the 95th percentile case involved costs of approximately \$300,000 for each party.⁶¹ Figure 7 presents the distribution of litigation costs from the *Civil Rules Survey*. The pattern is virtually identical to the pattern for preservation costs in Figure 6. And indeed, in the *Civil Rules Survey* data, the top 5 percent of cases accounted for about 60 percent of all litigation costs.

⁶⁰ The Company A sample contained only 14 employment matters in its sample. The “top 5 percent” figure for Company A is therefore simply the share of the single matter with the most holds.

⁶¹ *Civil Rules Survey* at 35–37.

FIGURE 7: DISTRIBUTION OF LITIGATION COSTS PER PARTY,
FJC CIVIL RULES SURVEY DATA



4.4 Vastly More Data Is Preserved than Used in Discovery

Unlike collection, processing, review, and production, preservation activity must occur before the appropriate scope of discovery in a case is well-defined—indeed, it often must occur before a lawsuit has even been filed. For a potential defendant, therefore, preservation activity may occur without the benefit of a defined plaintiff or claim, let alone a defined set of discovery requests. This requires the preserving party to be overinclusive in defining the scope of preservation, lest its best guess as to the proper scope of preservation be proven wrong in hindsight. For this reason, it is inevitable that companies will always preserve more than is requested and produced in litigation.

Nonetheless, the rules governing preservation activity can have an effect on the extent to which the scope of preservation activity is greater than it otherwise would need to be in order to effectuate its purpose of ensuring that the discovery process can function as designed. To determine the extent to which preservation activity may be broader than otherwise might be desirable, the Survey gathered data on the extent to which the volume of preserved data exceeds the volume of data used in litigation.

Much preservation activity involves matters with no filed suit

Given that preservation obligations may attach before a lawsuit is filed, it is no surprise that companies report that many litigation holds are opened before a suit is filed. Most companies do not track these numbers, but a few companies did provide such data. One respondent reported that 44 percent of holds are not for active litigation, while another reported that 77 of their holds were issued without a suit being filed or subpoena issued.⁶² These numbers bracket the figure of about 67 percent that was provided by the Microsoft Corporation outside the context of this Survey.⁶³ While some of these litigation hold matters eventually involve a filed lawsuit, others do not. The cost of these holds, of course, is incurred regardless.

Only a fraction of preserved data is ever collected

Most respondents in the survey did not know what fraction of the data that is put on litigation hold ever is collected, processed, and reviewed in the course of discovery. Those that did reported that perhaps half of all data that is preserved is never even looked at in the course of litigation. See Table 13.

⁶² Company A Interview 9/1/2011; Respondent 2973478321, Phase III Survey.

⁶³ David M. Howard, Jonathan Palmer, and Joe Banks, *Letter to the Hon. David G. Campbell* 3 (Microsoft Corp. Aug. 31, 2011).

This result is consistent with a recent survey by an e-discovery vendor, which found that for most companies, legal holds proceed to collection less than half the time.⁶⁴

TABLE 13: AVERAGE SHARE OF PRESERVED DATA
COLLECTED, PROCESSED, AND REVIEWED⁶⁵

	Share
Share of preserved data collected	59%
Share of preserved data processed and reviewed	49%

For larger companies, the drop-off from preservation to collection, processing, and review is even steeper. Figure 8 presents data from a large company on the number of custodians involved in three stages of discovery: preservation, collection, and processing. Out of over 5,000 custodians placed on litigation hold, and thus subject to preservation obligations, fewer than 10 percent ultimately see their data collected, let alone processed.

Figure 9 presents a similar picture with non-anonymous data provided in public testimony on behalf of Microsoft.⁶⁶ In Figure 9, the unit of measurement is the quantity of data preserved, collected, and processed rather than the number of custodians subject to those activities. The Microsoft data also illustrates how little data, relative to the quantity preserved, is ever used in litigation. (Data from Company K on data volumes reviewed, produced, and used is strikingly similar, as well.)

⁶⁴ *Legal Hold and Data Preservation Benchmark Survey 2013* 16 (Legal Hold Pro) (finding that for 64 percent of respondents, legal holds progress to collection less than half the time).

⁶⁵ Forty-nine companies provided figures for share collected; forty-five for processed and reviewed.

⁶⁶ Testimony of David M. Howard on behalf of Microsoft Corp., *Transcript of Public Hearings on Proposed Amendments to the Federal Rules of Civil Procedure* 79–80 (Jan. 9, 2014), available online at <http://www.uscourts.gov/uscourts/RulesAndPolicies/rules/public-hearings/civil-hearing-transcript-2014-01-09.pdf>

FIGURE 8: NUMBER OF CUSTODIANS SUBJECT TO PRESERVATION, COLLECTION, AND PROCESSING, COMPANY A

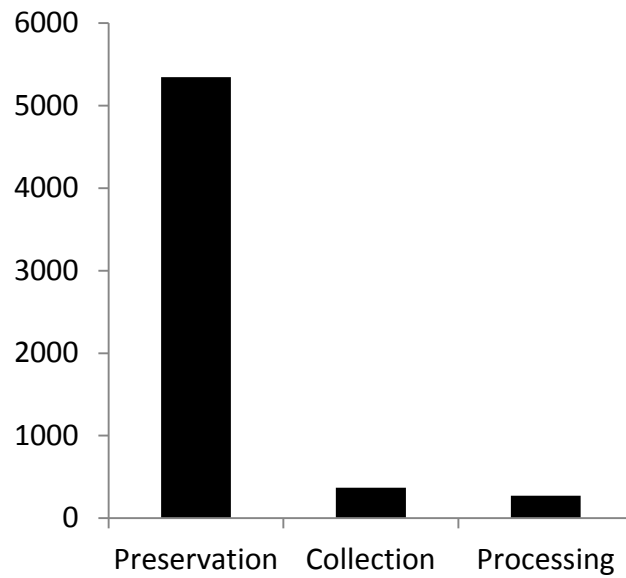
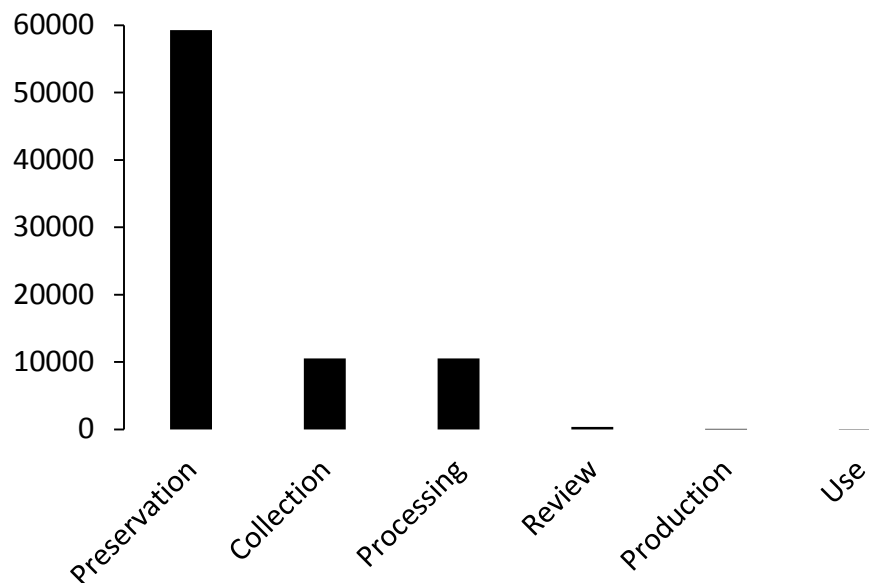


FIGURE 9: NUMBER OF PAGES (IN 1000S) PRESERVED, COLLECTED, AND PROCESSED, MICROSOFT CORPORATION



Note: “Pages” refers to paper pages or data equivalent (approximately 16 KB per page).

Uncertainty about preservation obligations leads to broader preservation

These results make clear that the vast majority of the data that is preserved is ultimately judged unnecessary to the litigation. But data that is never used still imposes preservation costs. This fact alone justifies careful consideration of whether the degree of preservation is properly calibrated to the needs of the litigation process.

The Survey data, however, cannot directly answer this question. The fact that much more data is preserved than ever collected, let alone used in litigation, does not prove that the preserved amounts are greater than appropriate. I note, however, that companies repeatedly expressed in interviews that they are deliberately “overinclusive” or “overpreserve” to protect themselves against the great uncertainty associated with the current law of preservation, which varies from district court to district court, and even judge to judge, within the federal system.⁶⁷ These reports are consistent with responses from other, qualitative studies on preservation activities.⁶⁸

⁶⁷ The Phase III did not ask about overpreservation. Nonetheless, a number of Phase III companies expressed similar sentiments in the open-ended comments section of the questionnaire.

⁶⁸ See *Where the Money Goes*, p. xx (“A key concern voiced by the interviewees was their uncertainty about what strategies are defensible ones for preservation duties.”); *id.* at 92 (“If there was one consistent theme in what we heard, it revolved around complaints of a lack of understandable legal authority and guidance that could be comfortably relied on when making preservation decisions.”).

5 Conclusion: Preservation Costs and the Federal Rules

Under the proposed amendments, for the first time Rule 37 would expressly address failure “to preserve” as a basis for sanctions. These proposals represent a deliberate, but modest, step toward clarifying and defining the duty to preserve in federal civil litigation. By addressing the standards for sanctions for failure to preserve, the proposed amendments to Rule 37 focus on an issue of expressed need; the costs imposed by the current environment of uncertain and conflicting legal standards was a repeated refrain from companies in interviews.

One looming question is whether it is possible to estimate the potential cost savings from the proposed Rules amendments addressing preservation. To be clear, the objective of the Survey was to gather, for the first time, both qualitative and quantitative information on the costs of preservation from a large sample of companies. It was not designed to directly measure any likely impacts of the proposed Rules amendments, nor could it—the Survey was developed before the current proposed amendments were finalized. With this important caveat in mind, I will conclude by offering some potential implications that can be drawn from the results presented in Part 4.

5.1 Preservation Costs Are Substantial

As detailed in this Final Report, the costs associated with preservation are substantial. For a large company, the costs of complying with preservation obligations will be measured in the tens of millions of dollars per year. These costs include frequent, multi-million dollar investments in technology to track and manage the preservation of an ever-expanding universe of ESI; the creation and maintenance of teams of lawyers, IT specialists, and other professionals devoted to compliance with preservation obligations; and the tens or hundreds of thousands of hours of employee time diverted from regular business operations to respond to litigation hold requirements. Thus, there is potential for significant costs savings from Rules amendments that clarify and define preservation obligations.

Of course, it is important not to overstate the potential for cost savings. While the proposed amendments will reduce these costs, most of these costs will remain. Large companies will still need to invest in expensive, sophisticated technology to track and manage ESI; they will still retain full-time legal IT and e-discovery personnel; and they will still issue thousands of litigation holds per year.

On the margin, however, greater certainty and less risk from the proposed Rules amendments will generate cost savings. The results of the Survey make clear that there is plenty of room to trim the extent of overpreservation without affecting data that would be within the scope of collection, processing, and review. In the Survey, companies often reported that they “overpreserve” due to lack of certainty about their preservation obligations and the perceived risk of being sanctioned for mistakes. Greater stability and less uncertainty in the law of preservation should have its most direct effect on this phenomenon of overpreservation. Less directly, improved certainty and clarity in the law may extend the life of investments in technology, as legal obligations become less of a moving target.

Reduction in overpreservation should manifest itself in reductions in the number of employees subject to litigation holds, which in turn relieves the burdens on their time. As shown in Table 7, and discussed in Part 4.1, **among the largest companies in the sample, the cost of lost employee time due to litigation holds averages over \$38 million per year. A mere 3 percent reduction in the number of employees subject to litigation holds implies over \$1 million in savings for a single company.** Thus, even a small percentage reduction in the costs of preservation activity will translate into economically significant savings to litigants.

5.2 Smaller Companies Face Serious Burdens and Benefit Less from Technology

Notably, smaller companies reported burdens and problems from preservation to the same degree as larger companies, which suggests that smaller companies, like larger companies, stand to gain from Rules amendments that reduce the costs of preservation. The savings in employee time spent on litigation holds will, of course, be proportional to a company’s scale of preservation activity, so the absolute magnitude of smaller companies’ cost savings will be less.

In one important respect, however, smaller companies are *more* vulnerable than larger companies to the costs of preservation. As the Survey results show, larger companies invest in in-house legal and IT expertise and employ automated systems to manage the costs and legal risks associated with preservation. But few of the smaller companies in the Survey do. The fixed costs of expertise and technology are investments that many smaller companies simply cannot afford. The lesson here is that **while technological solutions might partially address the high costs of preservation, technology alone cannot remedy the burdens faced by all companies. Most small companies are not in a position to benefit from current technology.**

5.3 A Small Percentage of Matters Generates Most Litigation Hold Activity

As reported in Part 4.3, a small share of litigation matters generate the majority of litigation hold notices. This has two implications for current debates on the proposed Rules amendments regarding preservation.

First, there have been many anecdotes suggesting that preservation burdens are large, and many anecdotes suggesting that they are not large. What is important to recognize is that these conflicting anecdotes do not pose a credibility contest in which one must choose which account to believe. Instead, these **divergent anecdotes on cost reflect different aspects of the same phenomenon—an enormous amount of preservation activity that is very unevenly distributed across litigation matters.** This is the context in which individual experiences with the costs of preservation must be understood.

Second, because the burdens of preservation activity are not distributed uniformly across litigation matters, **Rules amendments need not affect all cases, or affect all cases equally, in order to affect most preservation activity.** And as noted above, even a modest reduction in the scale of preservation activity will generate, in dollar terms, large savings.

5.4 Reducing Overpreservation Will Have Little Negative Impact on Discovery

Finally, while the proposed Rules amendments should generate cost savings, any benefits of a proposal must be weighed against its potential costs. A natural concern that arises in the context of preservation costs is that reducing preservation costs will reduce the amount of valuable discovery, with consequent impacts on the resolution of cases.

The results of this Survey indicate that there is little risk of any detrimental impacts on the discovery process from a reasoned scaling back of overpreservation in response to Rules amendments that clarify and better define the law governing preservation and sanctions. As reported in Part 4.4, most preserved data is never even collected, processed, and reviewed. Only a small fraction is ultimately produced as responsive and non-privileged, and an even smaller fraction is ever used in litigation. Thus, **there is plenty of room for sensible reductions in the volume of preservation activity without affecting the amount of relevant and useful data that is preserved for litigation.**

6 Appendix

6.1 Additional Figures for Phase II Companies

FIGURE A1: DISTRIBUTION OF EMPLOYEES ON HOLD PER MATTER,
COMPANY A SAMPLE, TOPCODED AT 500

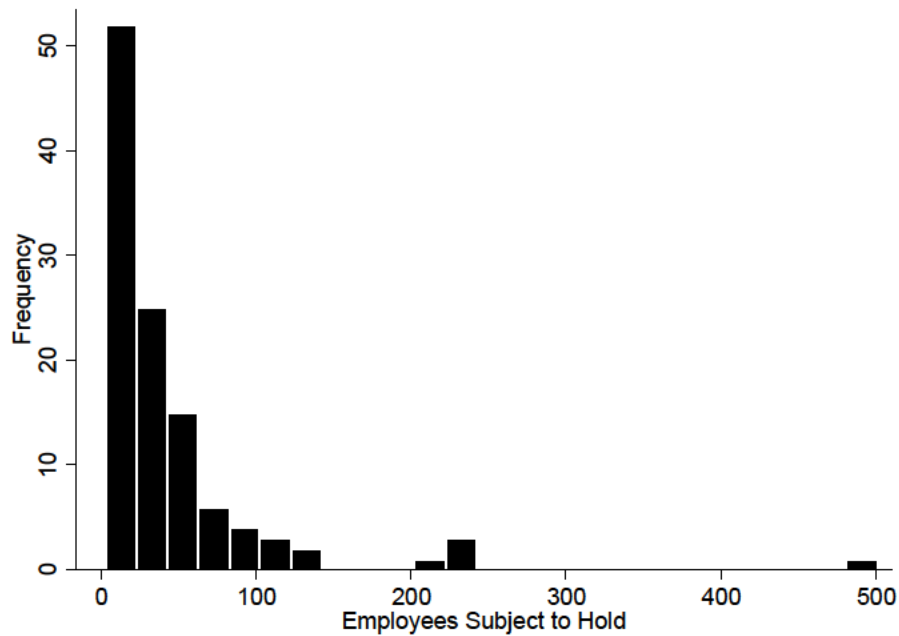


FIGURE A2: DISTRIBUTION OF EMPLOYEES ON HOLD PER MATTER,
COMPANY C SAMPLE, TOPCODED AT 500

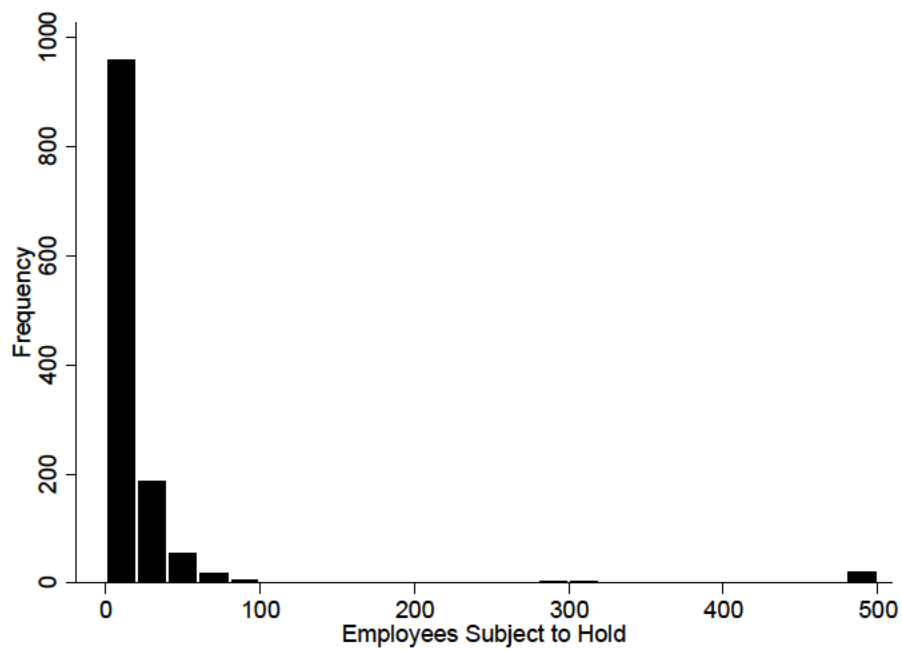


FIGURE A3: DISTRIBUTION OF EMPLOYEES ON HOLD PER MATTER,
COMPANY E SAMPLE, TOPCODED AT 1000

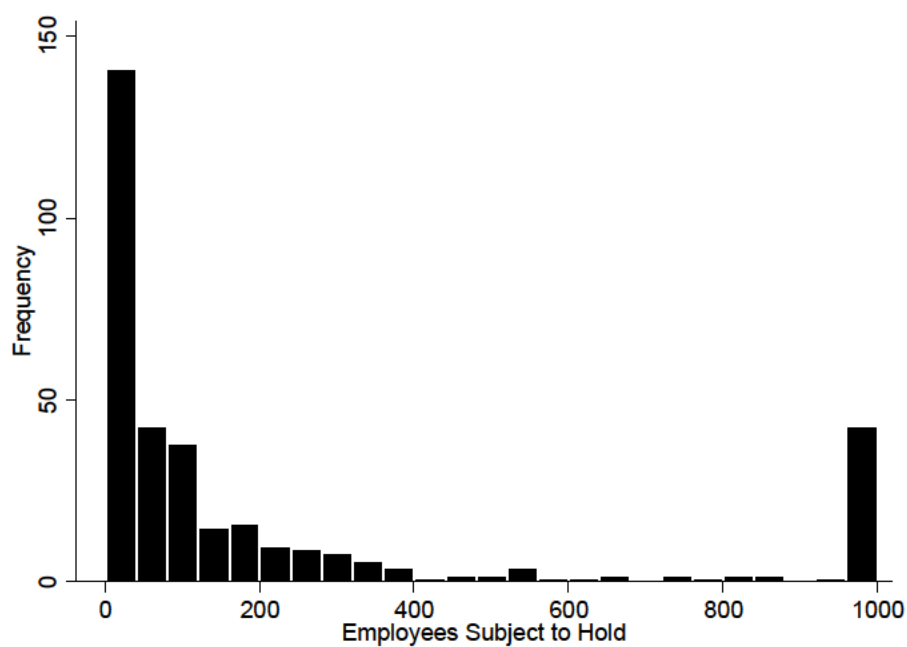


FIGURE A4: DISTRIBUTION OF EMPLOYEES ON HOLD PER MATTER,
COMPANY G SAMPLE, TOPCODED AT 500

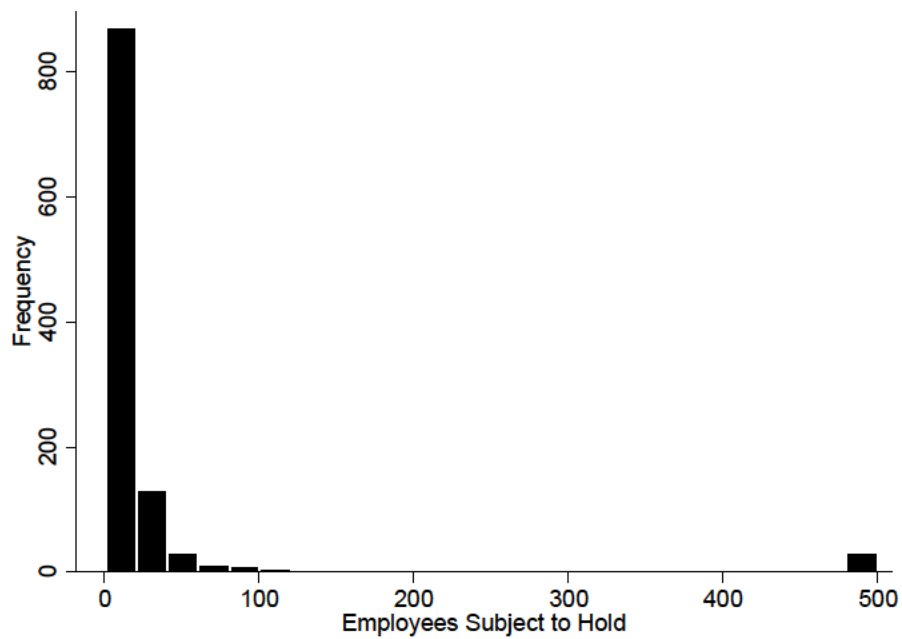
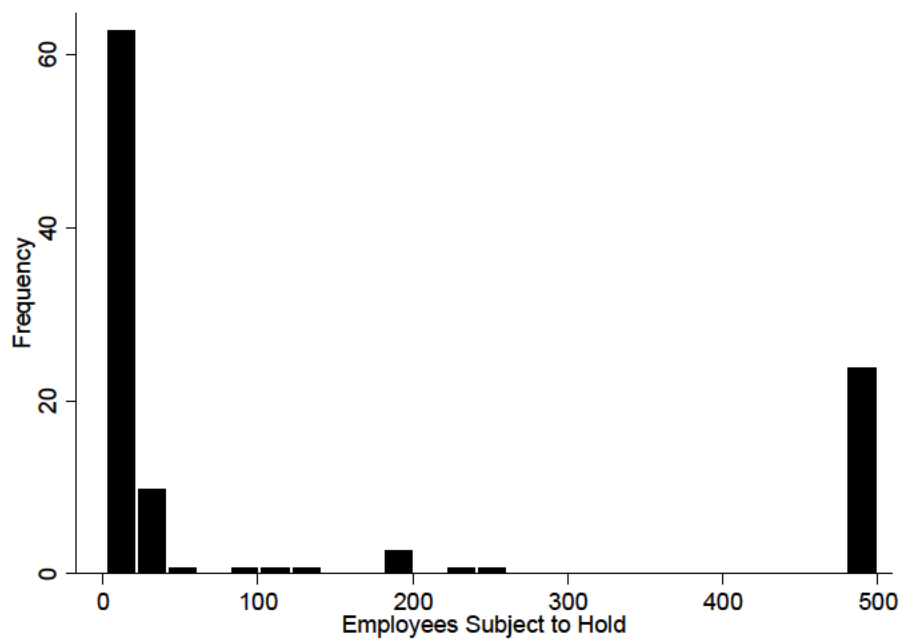


FIGURE A5: DISTRIBUTION OF EMPLOYEES ON HOLD PER MATTER,
COMPANY K SAMPLE, TOPCODED AT 500



6.2 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 for clients. Other aspects of my practice consisted of a broad range of pre-trial litigation and appellate advocacy.

From 2006 to 2011, I completed a PhD in Economics at the University of Chicago. I am an Assistant Professor of Law at the University of Chicago Law School, where I serve as an Editor of the *Journal of Legal Studies* and teach courses and seminars on civil procedure and economic analysis of law. I have presented testimony to the U.S. Congress on costs associated with discovery and preservation.

My research primarily addresses questions related to courts and civil procedure, with a focus on the empirical study of litigation. I have published or forthcoming articles in the *Journal of Legal Studies*; *DePaul Law Review*; *Common Market Law Review*; the *Journal of Law, Economics, and Policy*; the *American Economic Review Papers & Proceedings*; the *Journal of Human Resources*; and the *Journal of Human Capital*. I have presented papers at Harvard Law School, Columbia Law School, Northwestern University Law School, USC Gould School of Law, the University of Illinois College of Law, the Annual Meetings of the American Economic Association, and the Becker Friedman Institute for Research in Economics.

6.3 Phase III Survey Questionnaire

A copy of the paper/PDF version of the Phase III Survey Questionnaire appears beginning on the following page. The research.net version of the questionnaire had identical questions but different formatting. Note that the original deadline of October 28, 2013, which was printed on the paper version of the survey, was extended to January 2014.

Phase III Preservation Costs Survey
of
Anonymous Companies

Short-Form Survey Questionnaire

Prof. William H.J. Hubbard
University of Chicago Law School

Final Version
October 3, 2013

For Researcher Use: _____

Survey Instructions

Thank you for participating in Phase III of the Civil Justice Reform Group study of preservation costs. Your participation is vital to the success of this research effort. Preliminary, anonymous results from Phase I of the study have attracted the attention of the Advisory Committee on the Civil Rules, the U.S. Congress, legal academics, and the media. Phases II and III of the survey are ongoing.

The purpose of this survey questionnaire is to gather basic, anonymous information about your company's experience with the costs associated with legal preservation obligations.

A person who is directly involved in your company's electronic discovery and litigation hold efforts is probably the individual best suited to complete this survey questionnaire. This survey will focus on questions about your company's legal department and its experience with litigation holds and preservation obligations. There will also be a few background questions that require some basic information about the company as a whole.

This survey is designed to require only a limited time commitment. It contains **only 20 questions**. It can be completed electronically (it is a fillable PDF) or on paper. I estimate that you should be able to complete it in a half-hour.

Please do your best to make your responses to this form accurate and complete. Your time and effort in completing this survey is greatly appreciated and will be very valuable!

If have questions about this survey for any reason, please contact me at whubbard@uchicago.edu or 773-834-8999.

Remember, **all of your responses will be kept strictly anonymous** and will only be disclosed in ways that protect complete anonymity of your company. Thank you very much for your time and thoughts.

Please return this survey by October 28, 2013.

Question 1 (Industry)

Which single industry category best describes your company?

- ☐ Automobiles & Parts
- ☐ Banks
- ☐ Chemicals
- ☐ Financial Services
- ☐ Food & Beverage
- ☐ Health Care
- ☐ Industrial Goods & Services
- ☐ Insurance
- ☐ Media
- ☐ Oil & Gas
- ☐ Personal & Household Goods
- ☐ Retail
- ☐ Technology
- ☐ Telecommunications
- ☐ Travel & Leisure
- ☐ Utilities
- ☐ Conglomerate
- ☐ Other (Please specify): _____

Question 2 (Number of Employees)

How many employees does your company employ worldwide? Please include any contractors or others who would be subject to litigation holds issued by your Legal Department.

Question 3 (Employees in U.S.)

How many of your company's employees are in the U.S.? _____

Question 4 (Departing Employees)

In the last year (either the last calendar year or last 12 months), how many of your company's employees left your company, including terminations, retirements, etc.?

Question 5 (Number of Attorneys)

How many litigation (as opposed to transactional) attorneys does your company employ? Please include attorneys in Legal and Legal IT. Include only attorneys involved in matters to which your company is a party (e.g., an insurance company should not include attorneys employed solely to represent insureds against third-parties).

Question 6 (Legal IT)

Does your company have a separately-defined "Legal IT" team, department, or staff?

- ☐ Yes.
- ☐ No.

Question 7 (E-Discovery)

Does your company have a separately-defined "E-Discovery" team, department, or staff?

- ☐ Yes.
- ☐ No.

Question 8 (Current Active Lawsuits)

Right now, how many lawsuits filed by or against your Company are currently pending? *Please exclude asbestos-related lawsuits.*

Right now, how many *employment-related* (e.g., discrimination, retaliation) lawsuits filed by or against your Company are currently pending?

Question 9 (Preservation Issues)

In your opinion, has compliance with preservation obligations been a source of difficult problems or undue burdens to your company?

- ☐ Yes, to a great extent.
- ☐ Yes, to a moderate extent.
- ☐ Yes, to a small extent.
- ☐ No, not to any meaningful extent.

➔ *If you answered “No” to Question 9, please skip Questions 10–13 and continue with Question 14.*

Question 10 (Incidence of Issues—Data Source)

To what extent do you observe problems or burdens from preservation obligations arising for the following kinds of data?

Very Often		Sometimes		Very Rarely	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Paper documents.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Email.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Data on individual employee hard drives.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Data in databases.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Data on SharePoint / collaboration sites.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Documents stored on central servers.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Data on retired systems / legacy data.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Backup tapes.

Question 11 (Incidence of Issues—Forum)

To what extent do you observe problems or burdens from preservation obligations arising in the following settings?

Very Often		Sometimes		Very Rarely	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Cases in federal court.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Cases in state court.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Government investigations that have not yet been filed in court.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Private (non-governmental) matters that have not yet been filed in court.

Question 12 (Incidence of Issues—Case Size)

To what extent do you observe problems or burdens from preservation obligations arising in the following categories of cases?

Very Often		Sometimes		Very Rarely	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Large, complex matters (including class actions) involving individual plaintiffs.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Large, complex matters involving large companies on the other side.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Small, more routine matters involving individuals on the other side.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Small, more routine matters involving large companies on the other side.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Other (please specify): _____

Question 13 (Incidence of Issues—Claim Type)

To what extent do you observe problems or burdens from preservation obligations arising in the following types of disputes?

Very Often		Sometimes		Very Rarely	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Employment litigation.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Products liability.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Contract disputes.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Consumer protection law.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Antitrust.
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Other (please specify): _____

Question 14 (Preservation Policies)

Has your company developed explicit preservation policies with formal documentation and protocols?

- ☐ Yes.
- ☐ No.

Question 15 (Approach to Preservation)

What approach best describes how and where your company preserves documents and data? Please check all that apply.

- ☐ Preservation in place by individuals subject to litigation hold.
- ☐ Collection at the outset and litigation holds for individuals.
- ☐ Central data vault and/or repository capturing copies of data.
- ☐ A combination of the above methods.
- ☐ Other.

Please describe: _____

Question 16 (Issuing Holds)

Does your company issue litigation hold or preservation notices to employees (and non-employees, if applicable) for potential or actual litigation or investigations?

- ☐ Yes.
- ☐ No.

➔ *If you answered "No" to Question 16, please skip Questions 17–19 and continue with Question 20.*

Question 17 (Current Open Matters)

Right now, how many matters with litigation holds are active at your company? In other words, how many matters (including matters for which there is no current filed lawsuit) requiring litigation holds are currently pending at your company? *Please exclude asbestos-related matters.*

Right now, how many *employment-related* matters (including matters for which there is no current filed lawsuit) requiring litigation holds are currently pending at your company?

Question 18 (Cost Tracking)

Does your company separately track its costs associated with preservation and compliance with litigation holds?

- ☐ Yes.
- ☐ No.

Question 19 (Litigation Hold Systems)

Does your company have systems designed to manage and/or track litigation holds and hold notifications?

- ☐ Yes.
- ☐ No.

If yes, how do you manage and/or track litigation holds? Check all that apply.

- ☐ Through an automated system
- ☐ Through spreadsheets maintained by Legal
- ☐ Other (please specify): _____

Question 20 (Share of Matters with Collection, etc.)

Do you know in what percentage of matters in which data or documents are preserved are data or documents *collected, processed, or reviewed*?

☐ Yes.

☐ No.

If you answered "yes," please provide this information. Of all data and documents that are *preserved*:

_____ % are *collected*

_____ % are *processed and reviewed*

Comments on Survey

Feel free to provide any comments, questions, or suggestions you may have.

End of Survey!

Please return this survey by October 28, 2013 to

Prof. William H.J. Hubbard
University of Chicago Law School
1111 E. 60th Street
Chicago, IL 60637

(773) 834-8999
whubbard@uchicago.edu
