



CHICAGO JOURNALS

The University of Chicago
The University of Chicago Law School

Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts

Author(s): Yannis Bakos, Florencia Marotta-Wurgler, and David R. Trossen

Source: *The Journal of Legal Studies*, Vol. 43, No. 1 (January 2014), pp. 1-35

Published by: [The University of Chicago Press](#) for [The University of Chicago Law School](#)

Stable URL: <http://www.jstor.org/stable/10.1086/674424>

Accessed: 04/02/2015 15:44

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press, The University of Chicago, The University of Chicago Law School are collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Legal Studies*.

<http://www.jstor.org>

Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts

Yannis Bakos, Florencia Marotta-Wurgler, and David R. Trossen

A cornerstone of the law and economics approach to standard-form contracts is the informed-minority hypothesis: in competitive markets, a minority of term-conscious buyers is sufficient to discipline sellers from using unfavorable boilerplate terms. This argument is often invoked to limit intervention or regulate consumer transactions, but there has been little empirical investigation of its validity. We track the Internet browsing behavior of 48,154 monthly visitors to the Web sites of 90 online software companies to study the extent to which potential buyers access the end-user license agreement. We find that only one or two of every 1,000 retail software shoppers access the license agreement and that most of those who do access it read no more than a small portion. Since the cost of comparison shopping online is so low, the limiting factor in becoming informed thus seems not to be the cost of accessing license terms but reading and comprehending them.

1. INTRODUCTION

Standard-form contracts, often called fine print or boilerplate, are the most common type of economic contract. They apply to untold billions of commercial transactions per year. In a typical scenario, a buyer pur-

YANNIS BAKOS is Associate Professor of Management at the Leonard N. Stern School of Business at New York University. FLORENCIA MAROTTA-WURGLER is Professor of Law at New York University School of Law. DAVID R. TROSSEN is an associate at Knobbe Martens Intellectual Property Law. This research was supported by grants from the Networks, Electronic Commerce, and Telecommunications Institute (<http://www.NETInst.org>) and the Filomen D'Agostino and Max E. Greenberg Research Fund. The authors are grateful for comments received from Barry Adler, Jennifer Arlen, Oren Bar-Gill, Shmuel Becher, Omri Ben-Shahar, Kevin Davis, Clayton Gillette, Michael Levine, Ronald Mann, Yoram Margoliath, Jacob Nussim, Jeffrey Wurgler, and participants in the 2008 conference on Legal Institutions and Entrepreneurship, the 2009 Workshop of Information Systems Economics,

[*Journal of Legal Studies*, vol. 43 (January 2014)]

© 2014 by The University of Chicago. All rights reserved. 0047-2530/2014/4301-0001\$10.00

chases a good or service and is presented with a preprinted form contract with terms pertaining to dispute resolution, remedies for product failure, and warranties, among others, with little opportunity to negotiate the terms. Examples appear everywhere and can include safety disclaimers noted on the backs of sporting tickets, warranties packaged with consumer goods, privacy policies and terms of use on Web sites, and photocopying restrictions appearing in the front matter of this journal. Every reader of this paper has entered into thousands of standard-form contracts, sometimes unknowingly.

Academics, courts, and policy makers have long debated the degree to which standard-form contracts should be enforced and whether their content or disclosure should be regulated. All sides in this debate realize that, in many circumstances, a majority of buyers do not read fine print. For many buyers, too much time is required to read and give meaningful assent, and fine print can be too difficult to understand or may seem unimportant. The central economic question is whether the fact that a majority of buyers enter standard-form contracts with this imperfect information results in a market failure: if buyers do not factor contract terms into their purchase decisions, sellers lack incentives to provide anything more than the minimally required legal protections.¹

Defenders of freedom of contract have generally rejected intervention by relying on reputational constraints and the informed-minority argument. In this paper, we focus on this latter argument, which derives from some classic law and economics contributions. The articulation by Schwartz and Wilde (1979) of the informed-minority argument in this context is a specific application of work on imperfect information by

the 2009 Harvard-Texas Conference on Commercial Law Realities, the 2009 New York University colloquium on Law and Economics, the 2009 University of Virginia School of Law Legal Studies Workshop, the 2009 George Mason University–Microsoft conference Innovation: Online Markets versus Traditional Markets, the University of Kansas Law School, the 2009 Hebrew University of Jerusalem conference Contract Law: Interdisciplinary Perspectives, the 2009 Georgetown University Law Center Workshop in Law and Economics, the 2009 colloquium series at the University of Alabama School of Law, the 4th annual Conference on Empirical Legal Studies, the 2010 Fordham University Law and Information Society Faculty Workshop series, the 2010 Swiss Federal Institute of Technology (ETH-Zurich) Workshop and Lecture Series in Law and Economics, and the 20th annual meeting of the American Law and Economics Association. Dennis E. Hermreck, Daniel Priest, Robert Taylor, Michael Tonkinson, and Joel Willcher provided excellent research assistance.

1. For a comprehensive review of the factors that might contribute to consumer information problems and subsequent market failures, see Salop (1976) and Beales, Craswell, and Salop (1981).

Spence (1977) in the context of product liability and by Salop and Stiglitz (1977) in the context of price dispersion and search. Schwartz and Wilde (1979) argue that sellers will not necessarily offer one-sided terms even when the majority of buyers do not read standard-form contracts. In their model, nonreading buyers benefit from an informed minority whose willingness to pay for the product is sufficiently sensitive to the quality of the standard terms. When all buyers have the same taste for quality and sellers are unable to discriminate between reading and nonreading buyers, sellers will offer the terms preferred by all buyers. This competitive-markets logic has often been used to resist regulation and to advocate rules limited to facilitating search by those aspiring to join the informed minority (see, for instance, Baird 2006; Gillette 2005; Beales, Craswell, and Salop 1981).

Although the informed-minority argument has been influential in the law and economics literature, it has not been studied much empirically. While it is widely agreed that standard-form contract readers are in the minority, the literature offers no evidence of whether this minority remains large enough to plausibly enforce efficient terms, as assumed by some theorists and courts. In fact, we are not aware of any systematic studies of the extent to which consumers become informed about contract terms. We present large-sample evidence on the extent to which buyers actually do read standard-form contracts in a specific setting, and we identify factors that affect the probability of readership.

In particular, we examine the extent to which potential buyers of software read end-user license agreements (EULAs). For a sample of software companies who offer products online, we use potential buyers' "clickstream" information (that is, the series of uniform resource locator [URL] information for the page visited while browsing the Web and the precise timing of such visits) to study their readership of the EULA. We tracked 48,154 visitors to the Web pages of 90 software companies over a period of 1 month and recorded their detailed browsing behavior. For each such user, we observe the exact sequence of Web page addresses (URLs) accessed in a particular visit and the time spent on each page. The data also include the demographic characteristics of each user, such as age, gender, income, and geographical location. Our main finding is that regardless of how strictly we define a shopper, only one or two in 1,000 shoppers access a product's EULA for at least 1 second, which yields an informed minority of .2 percent that is orders of magnitude smaller than the required informed-minority size in realistic market settings and in theoretical examples suggested in the literature.

To put our findings in economic perspective, we perform a simple calibration exercise. We estimate the fraction of informed shoppers needed to induce sellers to offer good terms in the software market. We estimate the marginal cost of providing one pro-buyer term, maintenance and support, and find that sellers would find it more cost-effective to lose all informed buyers (that is, under the conservative assumption that each would decline to buy if the given term is not offered) than to offer this one term. This conclusion would likely persist for a fraction of informed buyers that is one or two orders of magnitude higher than .2 percent.

We then focus on the factors affecting the probability that a EULA will be accessed. We find that shoppers are more likely to access EULAs of smaller companies or companies that offer potentially suspicious products, such as freeware. The few shoppers who choose to become informed might be rationally deciding to ignore the EULAs of larger, more established companies, relying instead on company reputation or familiarity. We also find that older and higher-income shoppers are more likely to access EULAs. This may be because these consumers have lower search and reading costs (for example, because they have a lower opportunity cost for their time, or because they are more educated and thus find it easier to read contract terms). Thus, a higher fraction of consumers read EULAs when expected benefits are likely to be higher or reading costs are likely to be lower, which suggests that the fraction of informed consumers is limited by the high search and reading costs of standard-form contracts.

Our main contribution is the first large-sample study of the extent to which consumers actually inform themselves of important rights and obligations before entering a standard-form contract. We find that the fraction of consumers who read such contracts is so small that it is unlikely that an informed minority alone is shaping software license terms. We note, however, that the absence of an informed minority does not immediately imply that EULA terms will be inefficiently biased in favor of sellers. Some sellers, at least prominent ones, could be disciplined by other mechanisms, for example, the aforementioned concern for their reputation if onerous terms might eventually be discovered. Furthermore, shoppers may assume that no matter what the EULA terms state, they will be protected by the courts and thus will rationally choose not to become informed about the EULA terms. In other words, it may be rational not to read; what our evidence suggests is that EULA terms in our setting are not being policed by an informed minority of readers.

To summarize, while other factors may discipline sellers from offering adverse EULA terms, we show that the informed-minority hypothesis, the most widely applied argument for the efficiency of standard-form contract terms, does not seem compelling here. Our findings call into question the promise of policies to ameliorate potential market failures by requiring increased or mandatory disclosure, particularly in a setting where the informed minority is most likely to arise because of the low cost of online comparison shopping.² Future empirical work should focus on readership in other contexts to establish whether an informed minority might exist, as well as the potential validity of alternative mechanisms to discipline sellers in our context.

Section 2 offers theoretical and empirical background on the informed-minority hypothesis. Section 3 explains our methodology, Section 4 presents our data, and Section 5 discusses the results. Section 6 concludes.

2. FORM CONTRACTS AND COMPETITION FOR INFORMED BUYERS: BACKGROUND

Despite the transaction-cost-reducing benefits associated with the use of form contracts, such as reduced drafting and negotiation costs, academics and policy makers have debated their fairness and the desirability of their enforcement. Concern for consumer welfare has resulted in numerous articles, laws, and initiatives to regulate these contracts. For example, in addition to existing contract law doctrines to protect buyers from abusive terms, such as unconscionability and unfair surprise, several state consumer laws prohibit the use of forum selection clauses and disclaimers of implied warranties in consumer contracts (see, for example, *Gatton v. T-Mobile USA, Inc.*, 152 Cal. App. 4th Supp. 571, 585 [1st Dist. 2007]; *Fidelity & Deposit Co. v. Gainesville Iron Works, Inc.*, 125 Ga. App. 829 [1972]; Idaho Code sec. 29-110; N.C. Gen. Stat. sec. 22B-3; Mont. Code 36 sec. 18-1-403). Federal laws such as the Truth in Lending Act (15 U.S.C. 41) and the Magnuson-Moss Warranty Act (15 U.S.C. 2301) seek to decrease reading and search costs by requiring standardized disclosure of mandated terms. More recently, as can be seen in Lemley (2006), there has been heated debate about whether online contracts, such as terms of use, privacy policies, and

2. For an analysis of whether increased contract disclosure is associated with increased readership, see Marotta-Wurgler (2012). See also Ben-Shahar and Schneider (2011).

software license agreements, should be enforceable or subject to mandatory disclosure rules or provisions.

2.1. The Informed-Minority Hypothesis

The concern that standard-form contracts are likely biased toward drafters is based on the view that many buyers do not read or understand the terms. Salop and Stiglitz (1977) explore the conditions under which a market with consumers who are heterogeneous in their willingness and ability to become informed about product prices might reach a perfectly competitive price equilibrium and find that even in the presence of many uninformed consumers, a market can yield a competitive equilibrium if enough informed consumers shop for the lowest price.

In a widely cited contribution to the theory of standard-form contracts, Schwartz and Wilde (1979) extend this argument to settings in which consumers vary in their ability to become informed about contract terms. They show that if a sufficient number of buyers are informed about the price and contract terms of a given product, sellers who cannot discriminate between buyer types will offer the product with efficient terms at a competitive price to all buyers, as the cost to the seller of losing a critical mass of informed consumers outweighs the benefits of offering inferior terms to the uninformed inframarginal consumers. Imperfect information alone is thus not sufficient to warrant market intervention. This conclusion has become the cornerstone of the law and economics view of standard-form contracts. For example, Priest (1981), Baird (2006), and Hillman (2006a), among others, have repeatedly relied on the informed-minority argument to support freedom of contract. Instead of intervention, consumers should be given a meaningful opportunity to become informed about the terms prior to purchase.

Others have expressed doubts about this mechanism. Katz (1990) posits that not reading may be rational, given the low probability of adverse events triggering unfavorable clauses. Bar-Gill (2004) and Gabaix and Laibson (2006) question the underlying assumptions of the informed readers and offer behavioral accounts of a failure to read and understand terms. Eisenberg (1995) and Goldberg (1997) argue that sellers might find it more profitable to take advantage of nonreaders than to cater to readers. Ben-Shahar (2009) argues that because nobody reads fine print, regardless of its accessibility, rules requiring increased disclosure are useless, if not dangerous.

The informed-minority argument is reflected in contract doctrines and current proposals to increase consumer protection in mass-market trans-

actions online that stress that consumers must be given a meaningful opportunity to read.³ For instance, in *Principles of the Law of Software Contracts*, the American Law Institute (2010, p. 117) seeks to “promote reading and the opportunity to read terms” to alleviate market failures. Its goal is to reduce the cost of accessing the contract to promote a sufficiently numerous informed minority of buyers.

While this paper limits its scope to assessing whether readership levels are consistent with the informed-minority hypothesis, we note that other mechanisms may incentivize sellers to offer terms preferred by buyers even if none read. Sellers might be constrained by reputation or the threat of litigation, or they might offer one-sided terms to all consumers but might relax them to accommodate reasonable complaints (see Gillette 2004; Bebchuk and Posner 2006). In the case of experience goods or repeat purchases, buyers who do not read terms might ultimately become familiar with the contents of boilerplate. Our data cannot speak to the relevance of these mechanisms.

2.2. Prior Evidence

Despite the theoretical prominence of the informed minority, there has been little empirical investigation of its validity, presumably because observing readership is difficult. However, there is some related survey evidence. Hillman (2006b) surveys 92 contracts students and finds that only 4 percent of those who purchased products online claim to read standard-form contracts as a general matter. Becher and Unger-Aviram (2009) survey 147 students and find that 60 percent of respondents claim that they skim or read parts of a standard-form contract before entering a transaction. Plaut and Bartlett (2012) survey 182 undergraduates and find that about 80 percent claim not to read contracts and that much of the remainder claim to skim them. These surveys are suggestive but somewhat limited; they are based on self-reported behavior or hypothetical commercial scenarios, and the survey subjects are not representative; for example, in some cases they include law students who will write boilerplate for a living.

Other studies show that standard-form contract terms are less one-sided in favor of sellers than might be possible if buyers were completely

3. A notable case reflecting this view is *Specht v. Netscape Communications Corp.* (306 F.3d 17 [2d Cir. 2002]), in which the courts refused to enforce an arbitration agreement made available via a browse wrap (using hyperlink somewhere on the seller’s Web site) a few screens later because it did not provide sufficient notice. Ben-Shahar (2009) offers a comprehensive evaluation of this doctrine.

uninformed and thus are consistent with the existence of the informed minority as well as other disciplining mechanisms. Priest (1981) studies 62 product warranties and finds that they are not biased toward sellers but rather reflect the relative ability of buyers and sellers to prevent and insure against loss. Marotta-Wurgler (2007, 2008) analyzes the terms of 647 online EULAs and shows that while almost all of them are more restrictive than the relevant default rules, they do not all converge to the legal minimum. In a study of contracting practices by online retailers, Mann and Siebeneicher (2008) find that few sellers offer excessively one-sided contracts.

3. RESEARCH FRAMEWORK

We explore the presence of an informed minority of buyers by studying the browsing and shopping behavior of online consumers. In particular, we tracked the behavior of visitors to the Web sites of 90 software companies, and we examine the rate at which shoppers choose to become informed about the EULAs that govern the featured software.

Online software purchases provide an apt setting in which to look for the informed minority. First, while nonprice features, such as associated contractual rights and restrictions, are important for all types of products, they are a particularly significant consideration for information goods such as software, because terms form an integral part of the way the product is or may be used. Second, some of the terms in EULAs have been the subject of litigation in the past decade. (See, for instance, *Mortenson Co. v. Timberline Software Corp.*, 998 P.2d 305 [Wash. 2000]; *Davidson & Assoc. v. Internet Gateway*, 344 F. Supp. 2d 1164 [D. Mo. 2004]). For instance, as end users increasingly rely on software to perform a variety of routine tasks and critical functions, damages from software failure can be significant. Third, shopping for competing goods and the terms that govern them is cheap and easy online relative to most commercial settings. Search costs are also low (Bakos 2001). To the extent that the informed minority exists, this is among the settings where we are relatively likely to find it, especially given our access to clickstream data. Finally, several recent debates on legal reform in standard-form contracts focus on electronic contracts in general and software contracts in particular. Our study of the informed minority in online software markets places us at the center of these debates.

To empirically investigate the presence and size of the informed minority, we classify visitors to the Web sites of the companies in our

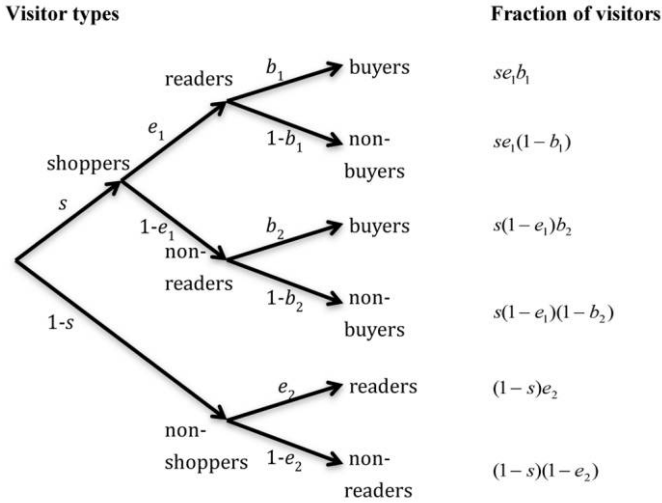


Figure 1. Empirical framework

sample, described below, as potential buyers and those visiting for other reasons, such as looking for user forums or troubleshooting information. We denote by s the fraction of potential buyers (shoppers); nonshoppers make up the remaining fraction $1 - s$. We denote by e_1 the fraction of shoppers and by e_2 the fraction of nonshoppers who read the online EULAs. Finally, we denote by b_1 the fraction who purchase the product (buyers) among shoppers who read the EULA and by b_2 the fraction of buyers among shoppers who do not read the EULA. This framework is depicted in Figure 1. In this setting, the informed minority corresponds to the fraction e_1 of shoppers who read the online EULA.⁴

4. It is possible that, for some shoppers, accessing the end-user license agreements (EULA) will not affect their probability of buying the product. For instance, some shoppers either may not know what a EULA is or may discover after accessing the EULA that they are not capable of comprehending its language, or shoppers may access the EULA accidentally or out of curiosity. To the extent that such accesses of the EULA do not make a shopper part of the informed minority, e_1 will overestimate the fraction of shoppers who are part of the informed minority. We can explore the significance of this to some degree by studying the time spent on the EULA page by those who access it. On the other hand, because we do not consider other ways in which shoppers might become informed about the terms (for example, by word of mouth or repeat purchases), there is a possibility that e_1 will underestimate the size of the informed minority. We comment on the likely significance of this effect in Section 5.

Next we estimate the number of visitors in our sample for each of the six categories in Figure 1. We estimate the number of readers and nonreaders among visitors classified as buyers, shoppers, and nonshoppers. We use access to a EULA page for more than 1 second to identify readers. This creates an upward bias of our estimate of the truly informed readers in that some accesses are accidental, inconsequential to the buying decision, accessed so briefly that little content could have been grasped or read but not understood. We use initiation of a secure checkout process to identify buyers and other contextual information to distinguish shoppers from nonshoppers.

We can break down readers into se_1b_1 readers who buy and $se_1(1 - b_1)$ readers who do not buy. In addition, $s(1 - e_1)b_2$ buyers are not readers, and $s(1 - e_1)(1 - b_2)$ shoppers neither read nor buy. A priori, we expect that few nonshoppers read EULAs, and thus we expect $(1 - s)e_2$ to be small. Finally, the fraction of nonshoppers who do not read EULAs is $(1 - s)(1 - e_2)$, which, as expected and as we confirm, is large. With these inputs, we can estimate the fraction e_1 of shoppers who constitute the informed minority, by writing this fraction as

$$e_1 = \frac{se_1b_1 + se_1(1 - b_1)}{se_1b_1 + se_1(1 - b_1) + s(1 - e_1)b_2 + s(1 - e_1)(1 - b_2)}.$$

We then analyze the seller's choice of which terms to offer, to assess whether our estimates are plausibly consistent with an informed-minority equilibrium.

4. DATA

Our large clickstream data set represents the browsing behavior of 92,411 U.S. households in January 2007. This data set was made available to us by a major online research company, which has recruited a representative panel of U.S. households that have agreed to install on their computers a data collection plug-in that records the URL address of each Web page visited. The data collected include the exact sequence of Web pages visited and the amount of time spent on each page. In raw form, this is a data set of significant size.⁵

The panel of households was selected to be demographically and geographically balanced and representative of the population of U.S.

5. Information was captured for 6,355,922 user sessions and 461,027,284 corresponding page views.

households with Internet access.⁶ The information captured in the raw data for each Web page visited by a panelist is coded with both a user identifier that anonymously but uniquely identifies each panelist and a session identifier that delimits each panelist's Web browsing into separate sessions. Additional information captured includes the URL of each Web page visited, the time that each Web page was accessed, the length of time spent on that page, whether that page was within a secure (that is, encrypted) connection, the Web server delivering the Web page, and a unique identifier for the company or division owning that Web server. The recorded page views compose the bulk of the data, but we were provided with useful additional files that include nonpersonally identifiable demographic information about the panelists and a corporate hierarchy identifying the parents, if any, of the divisions or companies that own the Web servers appearing in the data (for example, Office and Outlook are properly identified as companies or divisions having the same corporate parent: Microsoft).

4.1. Sample Construction

We consider the central standard-form contract in one important market. We study user visits to the Web pages of software companies that sell or distribute their products through their corporate Web sites and that make their EULAs available on their site for users to peruse at their option (prior to making any purchase decision). We use the data provider's classification of markets to identify visits to software companies only. We subsequently identify in our data two types of software companies that make their products available for online purchase or downloading: retailers and freeware providers. Retailers license their software for a price through their corporate Web site. Freeware providers offer their software for free to anyone wishing to download it; examples include browser toolbars and plug-ins.⁷ We are interested in observing

6. This data provider's panel is one of the largest representative media research samples in existence. During the period the data were collected, the sample of participants was defined using random-digit-dialing principles: the company selected a random set of phone numbers from all available residential numbers in the United States and attempted to recruit each for at most 15 times at different times of the day and on different days. The panel also included university students and individuals in the workplace. The company updated its demographic information regularly, had implemented various procedures to keep the panel updated, and ensured that tracking was unobtrusive to prevent any distortions in behavior.

7. We classify a company as a retail company if it offers its core or much of its software for sale, even if it also offers free software.

users' propensities to become informed about the terms of these two types of software.

For the purpose of having a sufficiently homogenous sample of sellers, we exclude subcategories such as vendors not making their products available for online purchase or downloading, peer-to-peer software providers, and Web hosting companies. We exclude companies with fewer than 50 unique visitors who viewed at least two pages during their visit; our interest is in users with intent or potential intent to purchase (shoppers), and users who view only a single page are less likely to have such intent. We identified 197 companies that satisfied the above conditions.

For each of these companies, we obtained the URLs of all EULAs available on the company's Web site. To find the EULAs, we visited each company's Web site and used manual browsing, Google searches within the Web site, and, if available, searches of the Web site provided by the company. In addition, we searched all page views in the clickstream data corresponding to these companies to identify possible EULA pages (for example, pages whose Web address contained the words "EULA," "legal," or "terms"), which we then investigated manually.

Some EULAs were presented as browse wraps—that is, they were posted as a hyperlink somewhere on the seller's Web site. We included all these companies because we can easily measure whether users voluntarily clicked on the EULA hyperlink.⁸ A minority of companies presented their EULAs as click wraps. This mode of presentation requires consumers to click on an "I agree" button acknowledging the EULA terms before they can purchase a product. There are two types of click-wrap sites. One type presents EULA terms via a hyperlink adjacent to the "I agree" button and thus requires an additional click to access the EULA. In this case, while all buyers are forced to acknowledge the EULAs, we can measure what fraction takes the extra step and actually clicks on the contract link, which is a necessary step in becoming informed about the terms. The other type presents the terms in a scrollable

8. A possible concern is that these contracts are not prominent enough to be binding. As noted in Section 2.1, courts have been reluctant to enforce browse wraps. However, this mode of contract presentation is not too problematic in the online software market because most sellers also present the EULA prominently after purchase, at the time of installation of the software. This mode of contract presentation, or "pay now, terms later" contracting, has been held to be valid by most courts. Even though the contract is available after purchase, consumers who aspire to become members of the informed minority would have an incentive to check contract terms before buying because it is less costly to comparison shop this way than to purchase and return software just to see the terms of the license.

text box above the “I agree” button; we removed from our data set companies that use this type of click wrap because we do not have a way to measure whether buyers read the terms. While scrolling through and reading EULA terms would likely increase the total time spent on the corresponding checkout pages, which we can observe, there are several other actions that users typically are required to take on the same pages. As a result, measures based on the total time spent on the checkout page where the terms are presented were too noisy to be useful for this type of company. Finally, we removed from our data set companies that did not make their EULAs available online.

After excluding companies for which we do not have enough data or that are otherwise inappropriate for our tests, we arrived at a final sample of 78 retail and 12 freeware companies. We have no reason to believe that our basic results or conclusions would change significantly were we to increase the number of companies in the sample or the time window during which panelists were followed. In addition, the size of our sample is probably more usefully characterized in terms of the tens of thousands of company visits that we track, described below, because each of these represents an opportunity to access a EULA and is thus the essential unit of observation.

4.2. Company and Product Characteristics

All else being equal, consumers may feel less need to scrutinize the terms in EULAs from companies that are large or old because they assume that such companies are more trustworthy and fair. To test this hypothesis, we obtain information about each company’s annual revenue, year of incorporation, and public or private status. These data were obtained from Yahoo! Finance, Hoover’s,⁹ or direct communication.

Table 1 reports summary statistics for the company characteristics of the two types of companies analyzed. For each company, we note the number of products that it offers (counting each distinctly named product as a separate product), which allows us to calculate the average revenue per product.

We also collect several product characteristics, and we record one flagship product per company. Many small- and medium-size companies market one main product, in which case we select that product as the flagship. For larger companies, we select the product accounting for the

9. Proprietary business information was purchased from Hoover’s, Inc. (<http://www.hoovers.com>).

Table 1. Company and Product Characteristics

	Mean	SD	Min	Median	Max
Company:					
Retail:					
Revenue (\$ millions)	1,560	6,980	.1	5.2	51,100
Average revenue per product (\$ millions)	99.3	406	.004	.99	2,550
Age (years)	15.84	10.10	3	14.5	56
Public company	.27	.45	0	0	1
Freeware:					
Revenue (\$ millions)	1,160	4,010	.1	.1	13,900
Number of products offered	9.25	28.58	1	1	100
Average revenue per product (\$ millions)	12.5	39.8	.1	.1	139
Age (years)	9.08	6.72	3	6.5	27
Public company	.08	.29	0	0	1
Product:					
Retail:					
Consumer product	.68	.47	0	1	1
Number of products offered	14.78	25.02	1	6	150
Price (\$)	407.95	1,042.92	9.97	65.12	5,295
Median price (\$)	364.84	1,032.23	1	49	5,000
Trial, featured product	.83	.38	0	1	1
Trial, most products	.78	.42	0	1	1
Freeware:					
Consumer product	.83	.39	0	1	1

Note. The sample consists of 78 retail companies and 12 freeware companies.

largest fraction of sales or, when this information is not available, the product most prominently featured on the Web site, as these might be the products about which most consumers care.¹⁰ Consumers might be less inclined to read the EULAs of the most popular (or flagship) products, as they are likely to have established stronger reputations. Still, the choice of flagship product should not affect our overall findings, as the readership and visit metrics that we report are for all products, not just flagship products.

It is possible that users are more inclined to become informed about the EULA terms of more costly products, so we record the price of the flagship product as well as the median price of all products available on the Web site for that firm. We record whether the product is a single or multiuse license, because multiseat licenses are likely to have higher prices, and whether the product is offered to developers. We note whether the company offers a trial version of the flagship product, and also of the majority of its products, because that may also affect users' propensity to read terms.¹¹ We also note whether the product is oriented toward business users or the general public. Finally, we classify each product into one of 150 software product categories (for example, antivirus or word processing), on the basis of the characterizations of software at Amazon.

4.3. Contract Characteristics

We want to measure the fraction of shoppers who become informed about EULA terms. We thus collected all the EULA URLs available on a company's Web site. As noted above, many firms sell only one product, and thus they make available online only the EULA that governs the use of that product. Other firms sell many products that are all governed by a single EULA posted on their Web site, and others post different EULAs for different products. Finally, some firms post the EULAs for all their current and past versions of all their products. We found 240 unique URLs corresponding to EULAs for our sample companies.

10. We used a flagship product to collect product-related statistics, as our data set did not allow us to identify the actual product considered by most shoppers, and we were not able to obtain detailed per-product sales data that we could use to weigh the products offered by different sellers.

11. Trial versions are generally offered with limited functionalities over a limited period. Marotta-Wurgler (2007) finds that the majority of trial licenses are noticeably different (for example, the trial license reads "trial license" and is generally shorter than the product license), such that a user would not consider them substitutes.

4.4. Defining Shoppers and Shopping Visits

Among panelists in our data who visit a given company's Web site, we need to define shoppers (namely, visitors with some potential to purchase), since a potentially large fraction of visitors may be browsing without any intent to purchase. We define a user visit as all page views (URL accesses) from a company's Web site during a single user session.

The first definition of user visit identifies shoppers by examining the intensity of a visit to a company Web site. A user with intent to purchase is likely to view several pages on the retail side of the company's Web site. We follow Moe and Fader (2004) and Catledge and Pitkow (1995) and define our broadest definition of a shopping visit as one with at least two page views on a company's Web site. A second, more restrictive definition includes all visits by users who accessed at least five pages on a given company's Web site. Bucklin and Sismeiro (2003) find that this is progressively more likely to exclude casual browsers.

At the other extreme, a visitor who has selected a product and initiated a checkout or payment process has demonstrated intent to purchase. Thus, we use the initiation of the checkout process as the strictest criterion to identify visits with intent to purchase. We identify such events by identifying for the 90 companies in our sample the Web page addresses that would be accessed only during the checkout and payment process and by subsequently recognizing visits that access such pages. While knowing that a user started a checkout or payment process provides no guarantee that the transaction was completed, it indicates a high likelihood that a transaction was at least contemplated. This definition of shopping visit is likely to be overly restrictive, as it excludes visits that do not result in the initiation of a checkout process.¹²

To summarize, the three measures described above establish the shopping intent of a session with increasing strictness. As our definitions of a shopping visit become stricter, we expect that estimates of the informed minority become more conservative, and the actual number is likely to lie somewhere between the three estimates that our methodology provides.

4.5. Defining Shopping Visits: Single Sessions versus Monthly Aggregates

To define a shopping visit, we adopt the two approaches that are standard in the literature using clickstream data. The first approach, used

12. Given the low conversion (of visitors to buyers) rates in e-commerce, such visits likely represent the majority of shopping visits.

by our data provider and the industry in general, defines user sessions as periods of Web browsing activity separated by at least 30 minutes of inactivity. Under this definition, as summarized by Moe and Fader (2004), a user can have multiple visits to a given company in a day, a week, or a month.

The second approach recognizes that a user's shopping activity on a given company's Web site can span several days or even weeks given the low cost of access. Johnson et al. (2004) find that repeated visits to a company's site within a month typically correspond to the same shopping cycle. We thus aggregate visits to a unique company's site in a given month and present these aggregated sessions as an alternative measure of company visits with intent to purchase.

4.6. Demographic and Geographic Data

We use personal information about our panelists to identify characteristics of shoppers and shopping households that are associated with becoming informed about standard terms. Our data set includes the age and sex of the head of the household, household income, household size, and whether there are children present in the household.

In Table 2, we report summary statistics. For the sample of 48,154 visitors who accessed a minimum of two pages on at least one company's Web site in the sample during a single uninterrupted session, the average age is 46.22 years, and the reported age range is 18–99 years. Average income for heads of households is \$60,502 (standard deviation, \$39,704). Income (and perhaps age) is top coded; median income (\$37,500) better describes the sample. About half of the heads of households are male. The average number of household members is 2.78. There are children in 41 percent of these households. Table 2 also shows summary statistics for the sample of visitors who accessed a minimum of five page visits on at least one company's Web site and for unique visitors who selected a product for purchase and began the checkout process.

5. RESULTS

Our analysis here is based on shopping visits to company Web sites in which the user accessed a EULA. We identify these visits by matching the URLs corresponding to all the EULAs we collected to the clickstream of URLs accessed by users during their company visits. We compute descriptive statistics about company visits and EULA accesses under

Table 2. Characteristics of Users Visiting at Least One Sample Company's Web Site for Three Definitions of a Visit

	Mean	SD	Min	Median	Max
User accessed at least two pages (N = 48,154):					
Age (years)	46.22	13.78	18	46	99
Gender (1 = male)	.50	.50	0	0	1
Income (\$)	60,502	39,704	12,500	37,500	150,000+
Household size	2.78	1.27	1	3	5+
Children (1 = yes)	.41	.49	0	0	1
User accessed at least five pages (N = 33,655):					
Age (years)	46.37	13.70	18	46	99
Gender (1 = male)	.50	.50	0	1	1
Income (\$)	60,612	39,782	12,500	37,500	150,000+
Household size	2.79	1.27	1	3	5+
Children (1 = yes)	.41	.49	0	0	1
User initiated checkout (N = 2,831):					
Age (years)	47.39	14.03	18	47	99
Gender (1 = male)	.52	.50	0	1	1
Income (\$)	63,008	41,373	12,500	75,000	150,000+
Household size	2.77	1.23	1	3	5+
Children (1 = yes)	.40	.49	0	0	1

alternative definitions of a visit with intent to purchase. Finally, we present regressions to study the determinants of the (as it turns out, low) probability that a EULA will be accessed.

5.1. Company Visits and End-User License Agreement Accesses

We want to measure the fraction of buyers who seek to become informed about EULA terms in deciding whether to purchase, and as noted we define the sample to include only those company Web sites where EULA access is possible but optional. Table 3 and Table 4 summarize the characteristics of visits to such companies, measured either as uninterrupted sessions (Table 3) or as visits by unique users, aggregating all the monthly sessions by individual users (Table 4). In each case, the data are presented for each definition of a shopping visit to a company's site. We separate visits according to the type of company visited, noting that only retailer visits include secure checkout page views; there is no need for a secure checkout process for a free product. In addition to company visits, the tables show information on page views and visits with EULA access. The number of pages viewed before the first EULA access and the length of time spent viewing EULAs give us some indication of shoppers' level of care or intent in accessing EULA pages. For simplicity, we report the definition of a company visit of intermediate strictness (five page views) for the uninterrupted sessions in Table 3, but the reader can explore alternative definitions in Tables 3 and 4 and see that our results are robust to alternative definitions of shopper and shopping visit.

When a visit is defined as requiring five or more pages accessed at the company visited, there were 72,282 visits during uninterrupted sessions to software retailers and 13,715 to freeware companies. The median number of pages viewed in a given visit to a retailer was 10 pages, and the median visit length was 183 seconds (3.05 minutes). Distributions of page views and visit durations are skewed. End-use license agreements were accessed 57 times among software retailers (.08 percent), and 30 visits were made among freeware companies (.22 percent). The median number of pages seen before accessing a EULA was eight for retailers and four for freeware providers. These numbers are already telling, but another consideration is whether shoppers who access the EULA actually read it. For users in this group, the average length of time spent on the EULA page was 62.7 seconds, and the median time was 32 seconds. (Note that we are defining "access" as a EULA visit of at least 1 second, for the purposes of obtaining a conservatively high number of EULA accesses.) Half of the accesses to EULAs lasted less

Table 3. Company and End-User License Agreement (EULA) Visits: Uninterrupted Sessions

Visit Definition	Company Visits		Page Clicks per Company Visit		Company Visit (Seconds)		EULA Visits		Company Visits (%)		Pages Viewed before EULA Access		EULA Access (Seconds)		EULA Accesses (%)		
	Visits	Mean	Median	Mean	Median	Mean	Median	Visits	Visits (%)	Mean	Median	Mean	Median	Mean	Median	Seconds	Minutes
User accessed at least two pages:																	
Retail	131,729	12.1 (26.7)	5	303.4 (698.4)	101	63	.05	19.1 (62.3)	7	59.4 (91.7)	46	92					
Freeware	28,663	13.4 (36.5)	4	164.3 (616.3)	43	44	.15	6.9 (14.8)	2	99.1 (227.3)	50	86					
User accessed at least five pages:																	
Retail	72,282 (34)	19.8 (34)	10	436.8 (891.6)	183	57	.08	20.9 (65.3)	8	59.6 (96.3)	49	91					
Freeware	13,715	25 (50.3)	12	241 (855.2)	68	30	.22	9.5 (17.4)	4	60.6 (104.5)	57	93					
User accessed at least one secure checkout page:																	
Retail	4,866	13.1 (30.7)	5	568.9 (1,861.8)	207.5	7	.14	15.7 (15.9)	7	166.3 (234.6)	0	71					

Note. Values in parentheses are standard deviations.

Table 4. Company and End-User License Agreement (EULA) Visits: Monthly Aggregates of Uninterrupted Sessions

Visit Definition	Company Visits			Page Clicks per Company Visit			Company Visit (Seconds)			EULA Visits			Company Visits (%)			Pages Viewed before EULA Access			EULA Access (Seconds)			EULA Accesses (%)			
	Company Visits	Mean		Median	Mean		Median	Mean		Median	Mean		Median	Mean		Median	Mean		Median	Mean		Median	Mean		Median
		Visits	Standard Deviation		Standard Deviation	Standard Deviation		Standard Deviation	Standard Deviation		Standard Deviation	Standard Deviation		Standard Deviation	Standard Deviation		Standard Deviation	Standard Deviation		Standard Deviation	Standard Deviation		Standard Deviation	Standard Deviation	
User accessed at least two pages:																									
Retail	69,372	23	(77.6)	7	576.1	152	61	.09	30.1	10	63.4	42	44	90											
Freeware	11,323	35.1	(288.8)	4	416	76	46	.41	10.2	3	101.7	33.5	49	85											
User accessed at least five pages:																									
Retail	43,708	34.9	(95.8)	14	829.2	286	56	.13	32.7	11	62.7	32	46	91											
Freeware	5,509	69.4	(411.4)	11	741.3	150	36	.65	12.6	4	107.8	33.5	49	84											
User accessed at least one secure checkout page:																									
Retail	2,991	37.1	(88.6)	11	1,456.8	457	9	.30	20.7	12	160.4	106	0	67											

Note. Values in parentheses are standard deviations.

than 30 seconds, and 90 percent spent less than 2 minutes on the contract.

To give these numbers some context, the average number of words in EULAs for retail products in the sample (unreported) is 2,277, with a median of 2,187 words and a standard deviation of 1,148. The time spent on the EULAs relative to their length indicates that readers rarely read terms in their entirety, especially as they are generally written in complex legalese. Since consumers are unlikely to be aware of the default rules, even if EULAs do spell out some terms in clear language, they may still be misunderstood. Bailey and Bailey (1999) find that the average (nonlegalese) reading rate of American adults is 250–300 words per minute, so a complete read of the typical EULA would require 8–10 minutes, not 1 minute. In other words, even the small number of EULAs accessed in our sample is still likely to be an overestimation—probably a substantial one—of the number of effectively informed readers. On the other hand, the small number of people who read EULAs may not be representative of the average reader and may have developed the ability to quickly skim the essential information.

Since our results could be biased if companies with relatively few visits in our sample are systematically different in terms of the probability of having visitors access their EULAs (for example, because visitors are less likely to be *ex ante* familiar with the terms offered by such companies), we recalculated the frequency of EULA accesses using frequency weights to adjust our data by the inverse of the total number of visits at each seller. This resulted in lower rates of EULA access across all definitions and thus addressed concerns that the lower frequency of visits to certain types of firms in our sample could be biasing our observed access rates downward.

Aggregating all monthly sessions of an individual user into a monthly visit (shown in Table 4) leads to similar results. The overall results, however, indicate that the impressions from Table 3 are robust to the precise definition of company visits. Ultimately, the highest fraction of readers among retail shoppers, across all shopper and session definitions, is .65 percent, or about six readers per every 1,000 shoppers.

5.2. Interpreting the Results: Can This Be an Informed-Minority Equilibrium?

Coming back to the empirical framework of Figure 1, we see that visitors to the Web sites of the companies in our sample can be classified as potential buyers or users visiting for other reasons. We measure the total

number of page views during each visit, whether a EULA was accessed, and whether a secure checkout session was initiated. These data, reported for individual sessions in Table 3 and for monthly visitors in Table 4, allow us to estimate the number of readers, buyers, and shoppers by using access to a EULA page as a proxy for reading, initiation of the checkout process as a proxy for buying, visits with five or more page views as a proxy for identifying shoppers, and visits with between two and five page views as a proxy for identifying nonshoppers. On the basis of the data in Table 4, we estimate the number of monthly visitors in our sample for each of the six categories shown in Figure 1.

We find that se_1b_1 , or nine, readers who buy and $se_1(1 - b_1)$, or 47, readers who do not buy. There are $s(1 - e_1)b_2$, or 2,982, buyers who are not readers and $s(1 - e_1)(1 - b_2)$, or 40,670, shoppers who neither read nor buy. Few nonshoppers would be expected to read EULAs, so it is not surprising that the value of $(1 - s)e_2$ is small; in our sample, it equals five (of 25,664 visits). Finally, the large majority of nonshoppers do not read EULAs: $(1 - s)(1 - e_2)$, or 25,661, on the basis of the above proxies. We thus arrive at an estimate for e_1 , the fraction of shoppers who constitute the informed minority of $[se_1b_1 + se_1(1 - b_1)]/[se_1b_1 + se_1(1 - b_1) + s(1 - e_1)b_2 + s(1 - e_1)(1 - b_2)] = 56/43,708 = .13$ percent.

It is possible that considering all visitors with five or more page views as shoppers overestimates the number of shoppers. An alternative estimate could be obtained by assuming that, among actual shoppers, the conversion ratio to initiate a checkout session among nonreaders is the same as that for readers, $9/56 = 16.1$ percent (which is higher than the purchase conversion ratios of 2–5 percent cited in the marketing literature but reasonable if not all checkout sessions that we capture result in actual purchases). In that case, the informed-minority fraction of all shoppers would be the same as the fraction of buyers, namely, $9/2,991 = .30$ percent.^{13,14}

13. Most models of the informed minority predict that the conversion ratio for nonreaders would be the same as or higher than the conversion ratio for readers, as the latter is less likely to purchase the product if they are not satisfied with the terms of the EULA. An upper bound on the size of the informed minority can be obtained if we assume that $b_2 = 100$ percent, that is, that 100 percent of nonreaders proceed to purchase the product. In that case, the informed minority would be $56/(2,991 + 47) = 1.84$ percent of the total number of shoppers.

14. If we assume that real shoppers will purchase from some merchant (while they may visit many) and that among these shoppers the ones who constitute the informed minority are equally likely to access a EULA at the sites of any of the merchants they visit, then

The bottom line is that the fraction of visitors who access EULAs is very small, on the order of .1 percent. While a number of alternative estimates can be calculated, these estimates point to that fraction being well under 1 percent. Assuming that no other disciplining mechanisms are at play, is it conceivable that such a small informed minority could protect all buyers and discipline sellers into providing efficient contract terms, thus preventing a market failure? The literature offers few meaningful suggestions as to how large the informed minority needs to be, and these are typically provided in the context of illustrative examples. Schwartz and Wilde (1979) offer an example in which the informed minority needs to be 20–30 percent to be effective. Our estimates here are imperfect, but they are two orders of magnitude smaller.¹⁵

Theoretically, the size of the informed minority required to induce sellers to provide good terms depends on the trade-off between the gross profit from selling to informed buyers (determined from the marginal cost of the product) and the cost of providing better contract terms. Consider a seller who may offer standard contract terms that are more or less favorable to the buyers, which we will call good and bad terms, respectively. The fraction r of buyers who become informed about the terms reflects the cost of finding and reading the standard-form contract and the expected benefit from doing so and is determined by the characteristics of the setting (for example, buyer search strategies as in Schwartz and Wilde [1979]). In our sample, this corresponds to the fraction e_1 of shoppers who are in the informed minority. Buyers value the rights and restrictions incorporated in the standard-form contract (for example, warranty terms, the ability to transfer the product, and so on), and thus good terms are valued more than bad terms. But good terms are naturally more expensive for the seller to provide than bad

the behavior of visitors in our sample who initiate a checkout session is representative of shoppers as a whole, and their likelihood of accessing EULAs (30 percent) provides an estimate of the size of the informed minority among these most determined shoppers.

15. The estimates presented above are based on monthly visits as reported in Table 4. They are conservative in the sense that using visits defined as individual sessions would result in lower estimates for the size of the small minority. Using session data from Table 3 would result in seven readers who buy, 50 readers who do not buy, 4,859 buyers who are not readers, 67,366 shoppers who neither read nor buy, and 59,447 nonshoppers who include only six readers. The fraction e_1 of shoppers in the informed minority would be $57/72,282 = .079$ percent. The fraction of readers who initiate checkout sessions would be $7/57 = 12.3$ percent, and assuming the same conversion ratio for nonreaders would give $7/4,866 = .14$ percent as the informed minority. A conversion ratio of 100 percent for nonreaders would give an upper bound for the informed minority of $57/(4,866 + 50) = 1.16$ percent.

terms, which results in corresponding product costs of c_g and c_b , with $c_g > c_b \geq 0$. Amending our earlier notation, we see that informed buyers purchase with probability b_1 if the terms are good and b_3 if the terms are bad ($b_1 > b_3$), and uninformed buyers still purchase with probability b_2 . The exact values of b_1 , b_2 , and b_3 are determined by the characteristics of the setting, but it is natural to consider $b_1 > b_2 > b_3$. The seller offers good terms if the expected payoff from doing so is higher than under bad terms:

$$[rb_1 + (1-r)b_2](p - c_g) \geq [rb_3 + (1-r)b_2](p - c_b).$$

Equivalently, the fraction of readers required to induce offering good terms¹⁶ is

$$r \geq (c_g - c_b) \left/ \left[\left(1 - \frac{b_1}{b_2} \right) (c_g - c_b) + \frac{b_1 - b_3}{b_2} (p - c_b) \right] \right.$$

This fraction becomes smaller as the incremental cost of providing good terms decreases and as the probability that shoppers who become informed will drop out if they see bad terms increases.

Given certain values for these unknown parameters, any fraction of informed shoppers could support an informed-minority equilibrium. However, the market for software maintenance and support (M&S) can be used to derive very rough estimates of the likely range of one of these parameters, the marginal cost of good terms, and to put our observed fraction of readers into perspective. Maintenance and support is a key term in software EULAs,¹⁷ and thus the cost of supplying M&S should be an order-of-magnitude approximation of the cost of offering good EULA terms.

To estimate the cost of M&S terms, we obtained product price and annual M&S price for 520 software products from the 42 software companies in the sample of Marotta-Wurgler (2007) that provided M&S separately on a periodic basis (that is, did not charge per incident). On average, M&S is priced at 26 percent of the product price (exclusive of the M&S). The median is 20 percent, and the standard deviation is 22 percent. Since there was high intercompany correlation, we focused on

16. An outcome in which the seller offers bad terms is inefficient if, assuming buyer valuations V_g for good terms and V_b for bad terms, $V_g - V_b > c_g - c_b$, as it corresponds to an inefficient provision of terms because the buyers value good terms above the seller's cost of providing them.

17. Marotta-Wurgler (2007) identifies and measures 23 important and common terms that allocate rights and risks between buyers and sellers of software, and maintenance and support (M&S) is one of these terms.

company means. The distribution for the 40 companies remaining after dropping two obvious outliers has a mean of .29, a median of .24, and a standard deviation of .16.

Thus, a year of M&S for software is on average priced at 25–30 percent of the product price.¹⁸ Since M&S costs are primarily variable (labor) costs, if the market for M&S was perfectly competitive, this would provide some indication of the marginal cost of M&S and thus a floor on the marginal cost of pro-consumer EULA terms. There are several reasons why 25–30 percent of the product price may be too high an estimate: (1) consumers may be more likely to purchase M&S from the seller of the software, and thus software companies may price as a two-part tariff, with a lower price for the up-front purchase (the software product) and a higher price for the subsequent purchase (M&S), (2) consumers who purchase M&S are likely to have higher M&S costs due to adverse selection and/or moral hazard, and (3) software companies may have substantial market power in providing M&S because of barriers to entry for competitors that are not as familiar with their product or consumers' propensity to purchase M&S from the seller of the original software. All of these factors would result in an M&S-to-product price ratio that is higher than the cost of providing M&S. On the other hand, M&S is only one of 23 key EULA terms, which include several other types of warranties and permissions to copy or distribute the software that can impose opportunity costs. Furthermore, M&S pricing is similar in enterprise software markets, where significant competition exists from third-party M&S providers and purchase of M&S contracts is almost universal.

On balance, it is reasonable to assume that the cost of this level of M&S is around 20 percent of the product price p , or $.2p$. Dividing the numerator and denominator of the fraction of informed buyers necessary to induce the seller to offer good terms by p , we get

$$r \geq \left(\frac{c_g - c_b}{p} \right) \left/ \left[\left(1 - \frac{b_1}{b_2} \right) \frac{c_g - c_b}{p} + \frac{b_1 - b_3}{b_2} \left(1 - \frac{c_b}{p} \right) \right] \right.$$

If $(b_1 - b_3)/b_2 \leq 1$ —which would be the case if readers purchase with at least the same probability as nonreaders when they discover good terms (bad terms reduce or eliminate this probability)— $b_1 \geq b_2$ and thus $1 - b_1/b_2 \leq 0$, and $c_b \geq 0$ and thus $1 - c_b/p \leq 1$, we get $r \geq (c_g - c_b)/$

18. One year was the most common duration—as well as the median duration—of free M&S for the companies that provided such a period of free M&S.

p , or $r \geq .2$. In other words, if the cost of providing good terms is $.2p$ and the three reasonable assumptions in the previous sentence hold, as is the case in our data, then one would require $r \geq .2$ to support an informed-minority equilibrium. This is 200 times larger than the value of $r \approx .001$ (or .1 percent) that we observe in the data. Alternatively, our data suggest that, for the informed buyers to induce the seller to offer good terms, their incremental cost would have to be almost negligible at less than .1 percent of the selling price.¹⁹ In either case, our data are unlikely to be consistent with an informed-minority equilibrium.

A similar and perhaps simpler approach to whether we might be observing an informed-minority equilibrium is to look at the seller calculus for visitors who initiate a checkout session. In an informed-minority equilibrium, sellers would be offering good terms, and according to Table 4, EULAs are accessed in six of 3,534 visits with checkout sessions. Since initiating a checkout session is a requirement to complete a purchase, if sellers were to offer bad terms, they might lose up to 100 percent of these readers. Thus, if sellers are trading off the net revenue from the sales to the six readers against the cost of providing good terms to the 3,528 nonreaders, as would be the case if we were observing an informed-minority equilibrium, sellers would keep offering good terms if the cost of doing so was less than .17 percent of the selling price.

This argument relies on fewer assumptions than that developed above, and it once again leads to the same conclusion, namely, that our data seem inconsistent with an informed-minority equilibrium. Hence, if the informed minority were the only disciplining mechanism, the natural implication would be that sellers are offering the lowest quality M&S terms. As they do not appear to take such advantage, it must be that other mechanisms are at play and that (perhaps) in the absence of such mechanisms, more buyers would read.

5.3. Robustness Checks: Becoming Informed without Reading?

Consumers can become informed in a number of ways. The Internet contains several consumer product review sites, blog posts with comments about product quality, and specialized news outlets that might

19. Of course, one could argue that offering a pro-buyer dispute resolution clause might cost the seller .1 percent of the price. Although it is theoretically possible, we find it unlikely that sellers will change this clause in response to the threat of losing one in 1,000 buyers. More important, the terms that matter the most to consumers (such as M&S and warranties) are among the costliest ones.

convey the content of standard-form contract terms to a few more shoppers.

From the Alexa Web site,²⁰ we obtained a list of 25 of the most trafficked sites likely to have information about EULA terms, product quality information, and sellers' practices. We then measured the rate at which shoppers in our sample accessed these sites. Very few sites focus exclusively on software and EULA terms. Others that occasionally discuss EULA terms are technology-related news sites, such as *Wired* and *Ars Technica*; sites that offer general consumer protection news and rants about abusive practices by sellers, such as *Consumerist*; and sites with general technology news, such as *PC Magazine*. Finally, there are general consumer information sites that do not focus on software but contain software product reviews, such as *Consumer Reports*.

We reviewed the pages accessed by shoppers visiting each URL associated with these 25 sites to make sure that the pages accessed contain information about EULAs. Of the 131,729 sessions with at least two pages accessed, only three shoppers accessed pages with EULA information in consumer review sites.²¹ At the monthly level of aggregation, 11,657 (16.8 percent) visits accessed at least one of the 25 consumer sites, but none of that activity was related to EULAs. Of this group, only 69 shoppers accessed pages with particular software product reviews or pages containing information such as reviews of tax software in that month. A total of 84 shoppers accessed news or general information pages about software, such as how to obtain certain free plugins. The remaining visits were to pages unrelated to software or EULA terms, so it does not appear that consumers are becoming informed by these alternative means.

It is also possible that the quality of terms is reflected in the seller's price, and thus shoppers do not need to become informed about these terms. As is the tenet of the informed-minority hypothesis, however, for terms to be reflected in prices, a sufficient number of consumers need to be informed about terms when they comparison shop (unless all sellers offer the worst enforceable terms, and consumers know this and therefore have no incentive to become informed about terms). Marotta-Wurgler (2007) finds that most sellers offer terms worse than those provided by default rules but not so bad that no information can be

20. The top Web sites were obtained from Alexa, Top Sites (<http://www.alexa.com/topsites>).

21. Two shoppers downloaded the EULAlyzer software from Javacool Software to obtain a review of a particular EULA. The third accessed a reference to a mock EULA.

gained by becoming informed about them. Our results are consistent with this, although we are not able to draw conclusions about the economic efficiency of this outcome.

Marotta-Wurgler (2007) is particularly relevant to this paper, as it looks specifically at the relationship between EULA term bias and price and finds almost no detectable relationship between them. It is possible that this is because terms for prepackaged software products are a relatively small component of price, and it is hard to measure the term-price relationship without controlling for the many other product attributes that affect price. However, this finding is also consistent with our results, as sellers would have little incentive to adjust their prices to reflect the quality of EULAs if consumers fail to become informed and thus compare terms. In such a situation, we would indeed expect that consumers can infer very little about EULA terms just by looking at price.

In addition, although consumers might be expected to have stronger incentives to become informed for higher priced products, the effect of price on the propensity to read terms is not statistically significant. However, we refrain from interpreting this evidence too strongly, because our investigation of why consumers do or do not read is limited compared with what we can explore directly, namely, the prediction that they read at all.

Another possibility is that the degree of disclosure of the EULA terms (for example, the prominence with which they are displayed) signals their quality. For instance, an equilibrium may arise in which good terms are prominently displayed while poor terms are made harder to find or inaccessible. Our results show that such an equilibrium, if it exists in our setting, is not maintained based on an informed minority. Furthermore, Marotta-Wurgler (2009) finds that prominently disclosed click-wrap contracts were roughly as one-sided as non-click-wrap contracts that required more effort to discover. This suggests that firms are not using the form of disclosure to signal the quality of terms. They do not seem to use click wraps to increase awareness of good terms, and they do not seem to use browse wraps to hide bad ones.

5.4. Determinants of End-User License Agreement Visits

Here we briefly examine those characteristics of the company, product, user, and Web site that distinguish the readers (or, more precisely, the clickers) from the nonreaders. Table 5 reports logit regressions in which

Table 5. Logit Regressions of Determinants of Online Visits to End-User License Agreements (EULAs)

	Uninterrupted Sessions			Monthly Aggregates of Uninterrupted Sessions		
	At Least Two Pages	At Least Five Pages	At Least One Secure Checkout Page	At Least Two Pages	At Least Five Pages	At Least One Secure Checkout Page
Freeware	1.06 (.75)	.71 (.81)		2.05** (.75)	1.95* (.79)	
In Median price	-.004 (.13)	-.02 (.14)	.38 (.39)	.07 (.14)	.06 (.14)	.07 (.41)
In Revenue per product	-.52** (.05)	-.55** (.06)	-.32 (.21)	-.52** (.05)	-.54** (.06)	-.36* (.14)
Public company	2.28** (.35)	2.57** (.43)	.29 (1.16)	1.85** (.35)	2.13** (.39)	1.08 (.89)
Pages viewed	.06** (.01)	.03* (.01)	.13** (.04)	.07** (.01)	.04** (.01)	.12** (.03)
Gender	-.39* (.22)	-.30 (.24)	-1.23 (.94)	-.35* (.20)	-.36* (.22)	-.48 (.71)
In Income	.11 (.15)	.21 (.17)	-.23 (.43)	.13 (.14)	.21 (.15)	-.21 (.42)
In Age	.40 (.35)	.14 (.39)	1.45 (1.31)	.30 (.33)	.19 (.36)	-.49 (1.38)
N	160,392	85,997	4,866	80,695	49,217	2,991
Pseudo-R ²	.11	.11	.20	.13	.13	.15

Note. The dependent variable is EULA access. Standard errors, clustered by visitor, are in parentheses.

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

the dependent variable is a dummy for whether a EULA was accessed during a particular company visit.

Factors that have a positive, albeit very small, effect on propensity to read are whether the product offered is freeware (since consumers might think there is a catch) and the number of pages visited in a session (longer visits might indicate a more serious intent to shop). A factor that has a negative effect on readership is the natural log of a company's revenues (as a proxy for size) divided by the number of products. A reason behind this is that consumers may be less likely to access EULAs for products they know about and thus may trust, relying instead on product familiarity. We use average product revenues as a proxy of consumer trust.²² Still, the main result here is that the most important term in the regression is the constant term: EULAs are rarely read by anyone.

5.5. Why Do So Few Consumers Read?

The small fraction of consumers accessing EULAs suggests a high cost of finding the EULA and reading its terms. If the primary cost lies in locating the contract, then mandating disclosure should increase the fraction of informed consumers. However, Marotta-Wurgler (2012) analyzes this same data set and finds that increased contract disclosure is not associated with increased readership. This suggests that the primary cost facing consumers is in reading and comprehending contract terms; it is possible that consumers do not access these terms even if they can do so with only the click of the mouse, as they expect reading EULA terms to be prohibitively costly. Measures that reduce the cost of comprehending the contract terms are likely to be more successful in increasing the fraction of informed consumers. Thus, a regulatory approach focusing on shortening and simplifying online contracts, standardizing their terms, and providing a standardized summary is more likely to increase readership than an approach focusing solely on disclosure. Regulations that mandate the disclosure of basic credit terms in a standardized manner and large fonts, such as the Schumer box in the United States and the

22. In unreported regressions, we examined whether shoppers are less likely to read the EULAs of products that are more likely to involve repeat purchases (such as upgrades) because they may be familiar with the terms from previous use. Other products, such as test preparation software, are less likely to be purchased repeatedly. The relationship between a dummy indicating whether a product is likely to be purchased repeatedly and the users' propensity to access EULAs is generally positive and, under certain definitions of shoppers, is statistically significant.

summary box in the United Kingdom, can reduce the cost of reading and comprehending contract terms. Simple and plain language requirements can be seen in the same light.

6. DISCUSSION AND IMPLICATIONS

Consumer access to the terms of standard-form contracts has been at the center of a legal and policy debate, and a major question has been whether disclosure of terms in standard-form contracts that govern consumer transactions should be regulated. A related debate has focused on the enforceability of terms and possible need to regulate disclosure for software in general and software purchased online in particular. A central issue in these debates is the validity of the informed-minority hypothesis: the view that comparison shoppers for standard terms help sustain efficient terms in equilibrium. In this paper, we investigate the extent to which consumers actually access the terms of certain standard-form online contracts. Our clickstream data allow us to measure the informed minority with reasonable precision for the first time.

We find that very few consumers choose to become informed about standard-form online contracts. In particular, we estimate that the fraction of retail software shoppers who access EULAs is between .05 percent and .22 percent, and most of the few shoppers who do access EULAs do not spend enough time doing so to have digested more than a fraction of their content. We also document that shoppers rarely access substitute sources of information, such as sites containing consumer product reviews or relevant news, to learn about EULA terms. Even under generous assumptions, it is difficult to envision the probability that EULAs are read (and understood) growing even to 1 percent. Our estimates of the size of the informed minority in this market are one or two orders of magnitude smaller than examples offered in the literature for the size required to sustain an informed-minority equilibrium, and this is further confirmed by simple theoretical calculations.

It is possible that buyers in the market for online software believe that other factors will discipline sellers in terms of the EULA terms they offer and thus rationally decide not to become informed and establish an informed minority in this market. Our data do not speak to whether it is rational or irrational to read or whether or not terms are efficient. Still, the market that we study offers a very favorable setting to look for an informed minority of buyers because of the low costs of accessing

contracts and shopping between firms and the presence of new and unknown firms that have yet to establish reputations.

While our results apply directly only to one context, the fact that online comparison shopping is so cheap and easy raises questions about whether informed-minority mechanisms could be consequential in other contexts. Furthermore, if that is the case, an approach similar to our own should result in readership levels that are substantially higher than what we observed. Assessing the presence of other factors that discipline sellers in the market for online software as well as whether readership levels are adequate to support the informed-minority mechanism in other markets are promising directions for further research.

The low readership of contract terms, even for those EULAs for which the terms are prominently accessible and consumers are required to acknowledge reading and agreeing to the terms before purchasing a product, suggests that it is the cost of reading and assessing these terms, rather than the cost of finding them, that discourages consumers. Hence, regulation that aspires to promote the emergence of an informed fraction of consumers solely through additional disclosure may be too optimistic. If the goal is achievable, it will require making contract terms easier to read, understand, and compare.

REFERENCES

- American Law Institute. 2010. *Principles of the Law of Software Contracts*. Philadelphia: American Law Institute.
- Bailey, R. W., and L. M. Bailey. 1999. Reading Speeds Using RSVP. *User Interface Update*. February.
- Baird, Douglas. 2006. The Boilerplate Puzzle. *Michigan Law Review* 104:933–52.
- Bakos, Yannis. 2001. The Emerging Landscape for Retail E-Commerce. *Journal of Economic Perspectives* 15(1):69–80.
- Bar-Gill, Oren. 2004. Seduction by Plastic. *Northwestern University Law Review* 98:1373–1434.
- Beales, Howard, Richard Craswell, and Steven C. Salop. 1981. The Efficient Regulation of Consumer Information. *Journal of Law and Economics* 21: 491–539.
- Bebchuk, Lucian A., and Richard A. Posner. 2006. One-Sided Contracts in Competitive Consumer Markets. *Michigan Law Review* 104:827–36.
- Becher, Shmuel I., and Esther Unger-Aviram. 2010. The Law of Standard Form

- Contracts: Misguided Intuitions and Suggestions for Reconstruction. *DePaul Business and Commercial Law Journal* 8:199–227.
- Ben-Shahar, Omri. 2009. The Myth of the Opportunity to Read in Contract Law. *European Review of Contract Law* 5:1–28.
- Ben-Shahar, Omri, and Carl Schneider. 2011. The Failure of Mandated Disclosures. *University of Pennsylvania Law Review* 159:647–749.
- Bucklin, Randolph E., and Catarina Sismeiro. 2003. A Model of Web Site Browsing Behavior Estimated on Clickstream Data. *Journal of Marketing Research* 40:249–67.
- Catledge, Lara D., and James E. Pitkow. 1995. Characterizing Browsing Behaviors on the World Wide Web. *Computer Networks and ISDN Systems* 27: 1065–67.
- Eisenberg, Melvin A. 1995. The Limits of Cognition and the Limits of Contract. *Stanford Law Review* 47:240–41.
- Gabaix, Xavier, and David Laibson. 2006. Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets. *Quarterly Journal of Economics* 121:505–40.
- Gillette, Clayton P. 2004. Rolling Contracts as an Agency Problem. *Wisconsin Law Review*, pp. 679–722.
- . 2005. Pre-approved Contracts for Internet Commerce. *Houston Law Review* 42:975–1013.
- Goldberg, Victor P. 1997. The “Battle of the Forms”: Fairness, Efficiency, and the Best-Shot Rule. *Oregon Law Review* 76:155–65.
- Hillman, Robert. 2006a. On-line Boilerplate: Would Mandatory Website Disclosure of E-Standard Terms Backfire? *Michigan Law Review* 104:837–56.
- . 2006b. Online Consumer Standard Form Contracting Practices: A Survey and Discussion of Legal Implications. Pp. 283–312 in *Is Consumer Protection an Anachronism in the Information Economy?* edited by Jane K. Winn. Burlington, Vt.: Ashgate Publishing.
- Hillman, Robert, and Jeff Rachlinski. 2002. Standard Form Contracting in the Electronic Age. *NYU Law Review* 77:429–95.
- Johnson, Eric J., Wendy W. Moe, Peter S. Fader, Steven Bellman, and Gerald L. Lohse. 2004. On the Depth and Dynamics of Online Search Behavior. *Management Science* 50:299–308.
- Katz, Avery. 1990. Your Terms or Mine? The Duty to Read Fine Print in Contracts. *RAND Journal of Economics* 21:518–37.
- King, Gary, and Langche Zeng. 2001. Logistic Regression in Rare Events Data. *Political Analysis* 9:137–63.
- Lal, Rajiv, and Carmen Matutes. 1994. Retail Pricing and Advertising Strategies. *Journal of Business* 67:345–70.
- Lemley, Mark. 2006. Terms of Use. *Minnesota Law Review* 91:459–83.
- Mandell, Lewis. 1971. Consumer Perception of Incurred Interest Rates: An Em-

- pirical Test of the Efficacy of the Truth in Lending Law. *Journal of Finance* 26:1143–53.
- Mann, Ronald, and Travis Siebeneicher. 2008. Just One Click: The Reality of Internet Contracting. *Columbia Law Review* 108:984–1012.
- Marotta-Wurgler, Florencia. 2007. What's in a Standard Form Contract? An Empirical Analysis of Software License Agreements. *Journal of Empirical Legal Studies* 4:677–713.
- . 2008. Competition and the Quality of Standard Form Contracts: The Case of Software License Agreements. *Journal of Empirical Legal Studies* 5: 447–75.
- . 2009. Are “Pay Now, Terms Later” Contracts Worse for Buyers? Evidence from Software License Agreements. *Journal of Legal Studies* 38:309–43.
- . 2012. Does Contract Disclosure Matter? *Journal of Institutional and Theoretical Economics* 168:94–119.
- Moe, Wendy W., and Peter S. Fader. 2004. Dynamic Conversion Behavior at e-Commerce Sites. *Management Science* 50:326–35.
- Plaut, Victoria C., and Robert P. Bartlett. 2012. Blind Consent? A Social Psychological Investigation of the Non-readership of Contracts. *Law and Human Behavior* 36:293–311.
- Priest, George. 1981. A Theory of the Consumer Product Warranty. *Yale Law Journal* 90:1297–1352.
- Salop, Steven. 1976. Information and Monopolistic Competition. *American Economic Review* 66:240–45.
- Salop, Steven, and Joseph Stiglitz. 1977. Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion. *Review of Economic Studies* 44:493–510.
- Schwartz, Alan, and Louis Wilde. 1979. Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis. *University of Pennsylvania Law Review* 127:630–82.
- Spence, Michael. 1977. Consumer Misperceptions, Product Failure and Producer Liability. *Review of Economic Studies* 44:561–72.