NELLCO
NELLCO Legal Scholarship Repository

New York University Law and Economics Working Papers

New York University School of Law

10-1-2009

Does Anyone Read the Fine Print? Testing a Law and Economics Approach to Standard Form Contracts

Yannis Bakos
Stern School of Business, New York University, bakos@stern.nyu.edu

Florencia Marotta-Wurlger
New York University School of Law, fmarotta@law.fordham.edu

David R. Trossen
Boalt Law School, University of California at Berkeley, dtrossen@berkeley.edu

Follow this and additional works at: http://lsr.nellco.org/nyu_lewp

Part of the Commercial Law Commons, Consumer Protection Law Commons, Contracts Commons, and the Law and Economics Commons

Recommended Citation
Bakos, Yannis; Marotta-Wurlger, Florencia; and Trossen, David R., "Does Anyone Read the Fine Print? Testing a Law and Economics Approach to Standard Form Contracts" (2009). *New York University Law and Economics Working Papers.* Paper 195.
http://lsr.nellco.org/nyu_lewp/195

This Article is brought to you for free and open access by the New York University School of Law at NELLCO Legal Scholarship Repository. It has been accepted for inclusion in New York University Law and Economics Working Papers by an authorized administrator of NELLCO Legal Scholarship Repository. For more information, please contact tracy.thompson@nellco.org.

Does Anyone Read the Fine Print?
Testing a Law and Economics Approach to Standard Form Contracts

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

October 6, 2009

Abstract

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 enough to discipline sellers from offering unfavorable boilerplate terms. The informed minority argument is widely invoked to limit intervention in consumer transactions, but there has been little empirical investigation of its validity. We track the Internet browsing behavior of 45,091 households with respect to 66 online software companies to study the extent to which potential buyers access the associated important standard form contract, the end user license agreement. We find that only one or two out of every thousand retail software shoppers chooses to access the license agreement, and those few that do spend too little time, on average, to have read more than a small portion of the license text. The results cast doubt on the relevance of the informed minority mechanism in a specific market where it has been invoked by both theorists and courts and, to the extent that comparison shopping online is relatively cheap and easy, suggest limits to the mechanism more generally.

[*] Stern School of Business, New York University.
[**] New York University School of Law.
[***] Boalt Law School, University of California at Berkeley.
This research was supported with a grant from the NET Institute (www.NETinst.org). The authors are grateful for the comments received by Barry Adler, Jennifer Arlen, Oren Bar-Gill, Shmuel Becher, Kevin Davis, Clayton Gillette, Michael Levine, Ronald Mann, Yoram Margoliath, Jacob Nussim, Jeffrey Wurgler, and the participants in the 2008 Law and Entrepreneurship Conference, the Workshop of Information Systems Economics (WISE), the Harvard-Texas Conference on Commercial Law Realities, the NYU Colloquium of Law and Economics, the University of Virginia School of Law Legal Studies Workshop, the GMU-Microsoft conference on "Innovation: Online Markets vs. Traditional Markets," the University of Kansas Law School, the Hebrew University conference on "Contract Law: Interdisciplinary Perspectives," and the Georgetown University Law Center Workshop on Law and Economics. Dennis E. Hermreck, Daniel Priest, Robert Taylor, Michael Tonkinson and Joel Willcher provided excellent research assistance.

1. Introduction

Standard form contracts, or what laypersons call "fine print" or "boilerplate," apply to most commercial transactions; they are probably the commonest type of economic contract. The degree to which standard form contracts should be enforced, and whether their content or disclosure should be regulated, has long been debated by legal academics, courts, and policymakers. All sides in this debate realize that some majority of buyers, in some majority of circumstances, does not read fine print. It is too long, hard to understand, or seemingly unimportant to take the time to read and give meaningful assent. The central economic question is whether the fact that a majority of buyers enter standard form contracts under 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 minimally enforceable legal protections.[2]

Defenders of freedom of contract have generally rejected intervention by relying on reputational constraints and on the "informed minority" argument. In this paper we focus exclusively on this latter argument, which has perhaps been best articulated in this context by Schwartz and Wilde (1979). Their articulation of the informed minority argument is a specific law and economics application of work on imperfect information by Spence (1977) in the context of product liability and by Salop and Stiglitz (1977) in the context of price dispersion and search.[3] Schwartz and Wilde argue that sellers won't necessarily offer one-sided terms even when the majority of buyers don't read standard form contracts. In their model, non-reading 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 non-reading buyers, sellers will offer the

[2] For a comprehensive review of the factors that might contribute to consumer information problems and subsequent market failures, see Howard Beales, Richard Craswell, & Steven C. Salop, The Efficient Regulation of Consumer Information, J. Law & Econ. 21(3) 491-539 (1981). See also Steven C. Salop, Information and Monopolistic Competition, 66 Am. Econ. Rev. 240 (1976).

[3] Alan Schwartz & Louis Wilde, Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis, 127 U. Pa. L. Rev. 630 (1979); Michael Spence, Consumer Misperceptions, Product Failure and Producer Liability, Review of Economic Studies (1977); Steven Salop & Joseph Stiglitz, Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion, 44 Rev. Econ. Stud. 493, 494 (1977).

terms preferred by all buyers. This argument has been used extensively to resist regulation.[4] For example, it has been argued that imposing a uniform standard would prevent sellers from using contracts to signal their quality or cater to heterogeneous buyer tastes. These scholars advocate rules only to facilitate search by those consumers aspiring to join the informed minority.

Although the informed minority argument is pervasive throughout standard form contract scholarship and argumentation, there has been practically no systematic empirical analysis of its validity. This paper addresses this void by providing the first large-sample evidence on the extent to which buyers actually do read standard form contracts. In other words, while we know that standard form contract readers are in the minority, what we don't know is whether this minority remains large enough to plausibly enforce efficient terms, as continued to be assumed by some theorists and courts, or rather whether it is so miniscule as to be almost surely irrelevant. We need to measure it, and we do so in one specific context in this paper. We also identify some of the factors that move the probability of readership up or down.

Specifically, we examine the extent to which potential buyers of software read End User License Agreements (EULAs). EULAs are contracts that govern the use of software products. For a sample of software companies who offer their products online, we use potential buyers' clickstream information to study their "readership" of the EULA. Our data track the detailed browsing behavior of thousands of Internet users over one month. For each user in the panel, we observe the exact sequence of web pages (URLs) accessed in a particular visit to a site and the time spent on each page. The data also include detailed demographic characteristics of the user, such as age, gender, income, and geographical location. We track the visits of 45,091 households to 66 software companies over a period of one month. Our main finding is that regardless of how strictly we define a "shopper" (or "potential buyer"), only about one or two in one thousand shoppers accesses a product's EULA for at least one second. For comparison, an informed minority of 0.2% is orders of magnitude smaller than the required informed minority size in the theoretical examples of Schwartz and Wilde. Such a small number of contract readers casts doubt on the existence of an informed minority of a size sufficient to police against one-sided

[4] See, e.g., Douglas Baird, The Boilerplate Puzzle, 104 Mich. L. Rev. 933, 933 (2006); Clayton Gillette, Pre-Approved Contracts for Internet Commerce, 42 Hous. L. Rev. 975 (2005). See also Beales, Craswell, & Salop, supra note 2.

terms, at least in the context of software sold online. Furthermore, since online comparison shopping is cheap and easy, the results suggest limits to the mechanism more generally.

Still, giving the theory the maximal benefit of the doubt, we consider whether such a modest fraction of "informed shoppers" could be sufficiently large 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, conservatively assuming each would decline to buy if the given term isn't offered) than to offer this one term. This conclusion would likely persist for a fraction of informed buyers 1-2 orders of magnitude higher than 0.2%.

We then focus on the factors affecting the probability that a EULA will be accessed. We find that shoppers are more likely to access the EULAs of smaller companies or companies that offer ex ante somewhat suspicious products such as freeware. The few shoppers that 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, e.g., 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, although only a tiny fraction of consumers read, the fraction grows a bit when expected benefits are likely to be higher or costs are likely to be lower; thus consumers seem to behave at least directionally in accordance with search theory, consistent with the broader lack of a significant informed minority being due to high search and reading costs of standard form contracts.

What can, and cannot, be concluded from this study? We can plausibly rule out an informed minority mechanism being important in this market. But the absence of this mechanism does not automatically prove that EULA terms will be inefficiently biased in favor of sellers, as sellers could be disciplined by other mechanisms, the most notable being a concern for their reputation. In other words, this particular (theoretically and legally popular) check on one-sided standard form contracts does not appear relevant, at least in this market, meaning that future empirical work can focus on the potential validity of these alternative mechanisms. Furthermore, our findings call into question the effectiveness of policies attempting to prevent potential market failures by requiring increased or mandatory disclosure. Shoppers do not access EULAs

regardless of how accessible they are.[5] As a result, disclosure is unlikely by itself to lead to the emergence of an informed minority of a meaningful size.

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

2. The Informed Minority Hypothesis: Academic and Legal Background

In a typical standard form contract scenario, a buyer purchases 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 over the terms. Fine print is pervasive. Every reader of this article has likely entered into thousands of standard form contracts.

Despite the benefits associated with standard form use, such as a reduction in 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 markets. 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.[6] On the federal front, laws such as the Truth in Lending Act and the Magnuson-Moss Warranty Act seek to decrease reading and search costs by requiring standardized disclosure of mandated terms.[7] More recently, there has been heated debate whether online contracts such as "Terms of Use," privacy policies, and software license

[5] For an analysis of whether increased contract disclosure is associated with increased readership, see Florencia Marotta-Wurgler & Yannis Bakos, Does Disclosure Matter? (work in progress, 2009).

[6] See, e.g., notes 39, 40, infra.

[7] See 15 U.S.C. §§ 2301-2312 (1976); 15 U.S.C. § 1601.

agreements should be enforceable or subject to mandatory disclosure rules or contain mandatory provisions.[8]

2.1. The Informed Minority Hypothesis

The concern that standard form contracts are likely biased towards drafters stems from the view that because many buyers do not read or understand the contract terms, sellers will impose unfair and one-sided terms. Salop and Stiglitz (1977) show that the existence of uninformed consumers need not prevent a competitive outcome. Specifically, they explore the conditions under which a market with consumers heterogeneous in their willingness and ability to become informed about product prices might reach a perfectly competitive price equilibrium. Even when many uninformed consumers exist, a market can yield a competitive equilibrium if enough informed consumers do shop for the competitive price: "[T]here is an informational externality at work between efficient and inefficient information-gatherers. Those agents who become informed give an external economy to the uninformed; the weight of their search keeps prices lower. In fact, if there are enough informed agents, the market price will settle down to the perfectly competitive price."[9]

Schwartz and Wilde extend this argument to a situation in which consumers vary in their ability to become informed about standard form 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. This is because when markets are sufficiently competitive, the cost to the seller of losing a critical mass of informed consumers outweighs the benefits of offering self-serving terms to those uninformed inframarginal consumers.[10] The authors

---

[8] See, e.g., Mark Lemley, Terms of Use, 91 Minn. L. Rev. 459 (2006) (arguing that browsewraps should be enforced only in cases involving sophisticated commercial parties that are repeat players); Hugh Collins, Regulating Contracts 279–86 (Oxford Universtiy Press, 1999) (arguing that mandatory terms would reduce the contractual one-sidedness caused by market imperfections).

[9] Salop & Stiglitz, supra note 3. The authors also mention that "[o]n the other hand, by shopping at high-priced stores, the uninformed inflict an external diseconomy on the informed; these informed must gather costly information to obtain the lower price."

[10] Alan Schwartz and Louis Wilde, supra note 3. In their model, the equilibrium result depends on the fraction of informed consumers. A competitive equilibrium will result if this fraction is "substantial." If the number is too small to support a competitive equilibrium, but it is nonetheless significant, the market will reach an equilibrium with a

conclude that imperfect information alone is not sufficient to warrant market intervention, a conclusion that would become the cornerstone of the law and economics view of standard form contracts.

Schwartz and Wilde state that "[g]enerally, there are a significant number of informed consumers in any given market prior to legal intervention,"[11] although the evidence they cite to support this statement would not appear to be so general or conclusive.[12] The main issue for them is whether these consumers are able to access adequate information at a reasonable cost. The cheaper it is to do so, the larger and thus the more influential this informed minority will be. The determination of the exact proportion of informed consumers necessary to yield a competitive equilibrium is a complicated exercise.[13] Nevertheless, many scholars have relied repeatedly on the informed minority argument to support freedom of contract in mass market transactions.[14]

---

cluster of prices. When there is only a small minority of comparison shoppers, the price distribution will converge towards a monopoly price (or terms). Schwartz and Wilde at 562.

[11] Schwartz and Wilde (1979), id at 636.

[12] The authors cite a study that examines the effect of the Truth in Lending Act on consumers' knowledge of the interest rates charged by their lenders. They find that prior to the passing of the Act, about 14% of the families in the sample estimated accurately the interest rates they paid on their loans and 33% of families did not know their true rates of interest on their loans. After the passing of the Act, the percentage of families with accurate estimates increased to 21% (a statistically significant difference), and the percentage of families with inaccurate estimates of their interest costs decreased to 25%. Lewis Mandell, Consumer Perception of Incurred Interest Rates: An Empirical Test of the Efficacy of the Truth in Lending Law, 26 J. of Finance 1143, 1153 (1971). While the study shows that the Truth-in-Lending Act was modestly effective in helping consumers understand the true rate of interest on their existing loans, it says nothing as to whether consumers were informed about the menu of prices or contract terms available to them when they were shopping among creditors. Furthermore, of all contract terms, price (the interest rate) is likely to be the most salient one; the extent to which understanding fine print of these contracts increased in presumably even more modest.

[13] Schwartz and Wilde provide a numerical example where the presence of one third of informed consumers would generate a competitive equilibrium with respect to price. Changing the relative costs, however, would require a different percentage of informed consumers. The authors expect that there will be fewer consumers informed about terms in a given market because shopping for terms is costlier than shopping for price, but contend that if 33% of those price savvy shoppers are also term savvy, a competitive equilibrium could result. For a critique of this illustration, see. R. Ted Cruz and Jeffrey Hinck, Not My Brother's Keeper: The Inability of an Informed Minority to Correct for Imperfect Information, 47 Hastings L. J. 636 (1995). See also, Oren Bar-Gill and Elizabeth Warren, Making Credit Safer, 157 U. of Pa. L. Rev. 1 (2008). But see, Alan Schwartz, How Much Irrationality Does the Market Permit? 37 J. L. Stud. (2008).

[14] For example, Priest relies on the informed minority argument to address the concern that most buyers will not factor warranty terms in their purchase decisions. George Priest, A Theory of the Consumer Product Warranty, 90(6)

Courts have also relied on the informed minority argument in deciding whether certain standard form contract terms provisions should be enforceable. For example, in *ProCD vs. Zeidenberg*, Judge Easterbrook enforced a restriction in a shrinkwrap license by noting that "[t]erms of use are no less a part of 'the product' than are the size of the database and the speed with which the software compiles listings. Competition among vendors, not judicial revision of a package's contents, is how consumers are protected in a market economy."[15]

2.2. Skeptical Responses

Other scholars are dubious that markets for standard form terms work so well. Some are dismissive based simply on casual observation. Slawson (1975) argues that "[f]or the very reason that these terms are imposed rather than agreed upon, they are almost universally unfair."[16] Rakoff (1983) believes that they should have a presumption of invalidity.[17] Failure to read has been rationalized based on the low probability that the contingencies specified in the contract will materialize as well as individual consumers' inability to alter the terms anyway.[18] Finally, some scholars have explained that consumers' psychological biases and limitations might prevent them from reading or understanding terms, once again making them susceptible to seller manipulation.[19]

Yale L. J (1981) at 1347. Baird a writes that "[the] typical buyer cannot rely on her own expertise or her ability to dicker with her seller. When the market works effectively, however, she benefits from the presence of other, more sophisticated buyers." Douglas Baird, supra note 4. See also Duncan Kennedy, Distributive and Paternalist Motives in Contract and Tort Law, with Special Reference to Compulsory Terms and Unequal Bargaining Power, 41 Md. L. Rev. 563, 614-15 (1982). For examples of references to the informed minority argument outside standard form contracts, see Richard A. Epstein, The Unintended Revolution of Products Liability Law, 10 Cardozo L. Rev. 2193, 2204 (1989); Alan Schwartz, The Case Against Strict Liability, 60 Fordham L. Rev. 819, 826 (1992).

[15] ProCD v. Zeidenberg, 86 F. 3d 1447 (7th Cir. 1996).

[16] W. David Slawson, Standard Form Contracts and the Democratic Control of Lawmaking Power, 84 Harv. L. Rev 529, 530 (1975).

[17] Todd D. Rakoff, Contracts of Adhesion: An Essay of Reconstruction, 96 Harv. L. Rev. 1173, 1226 (1983) (stating that "the ideal adherent who would read, understand, and compare several forms is unheard of in the legal literature, and, I warrant, in life as well.")

[18] See, e.g., Avery Katz, Your Terms or Mine? The Duty to Read Fine Print in Contracts, 21 RAND J. of Econ. 518 (1990); Melvin A. Eisenberg, The Limits of Cognition and the Limits of Contract, 47 Stan. L. Rev. 211, 240-41 (1995); Gillette, supra note 4.

[19] See, e.g., Xavier Gabaix & David Laibson, Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets, 121 (2) Q. J. of Econ. 505 (2006); Oren Bar-Gill, Seduction by Plastic, 98 Nw. U. L. Rev.

Others challenge the informed minority argument on its own terms. Eisenberg (1995) argues that "[t]ypically [a competitive equilibrium] will not occur, because most form takers will find it irrational to engage in search and deliberation on any given form."[20] Recently, Ben-Shahar (2009) has advocated abandoning recent disclosure proposals that seek to increase the "opportunity to read" standard form contracts.[21] He asserts that because nobody reads fine print, regardless of reduced reading costs in environments such as the Internet, rules that focus on increasing contract disclosure are useless, if not dangerous. Lastly, Goldberg (1997) questions the existence of the informed minority as well as whether sellers will indeed find it more profitable to cater to the readers than to take advantage of the non-readers.[22]

Finally, some rely on the informed minority argument to argue against regulation, while at the same time expressing reservations. Gillette (2004) argues that standard form contracts should be enforceable as long as the interests of those uninformed buyers are indeed accurately represented by the informed minority.[23] Hillman and Rachlinski (2002) explore the role of the informed minority argument in standard form contracting online.[24] Following Salop (1976), they conclude that although the low cost of becoming informed on the Internet is likely to increase number of informed consumers, the free-rider problem introduced by those uninformed

---

1373 (2004); Russell Korobkin, Bounded Rationality, Standard Form Contracts, and Unconscionability, 70 U. Chi. L.Rev. 1203 (2003); Ronald J. Mann, "Contracting" for Credit, 104 Mich. L. Rev. 899 (2006).

[20] Melvin Eisenberg, The Limits of Cognition and the Limits of Contract, 47 Stan. L. Rev. 211, 243-44 (1995).

[21] Omri Ben-Shahar, The Myth of the "Opportunity to Read" in Contract Law, European Review of Contract Law, 5(1) (2009). See also Robert Hillman, On-Line Boilerplate: Would Mandatory Website Disclosure of E-Standard Terms Backfire?, 104 Mich. L. Rev (2006).

[22] Victor P. Goldberg, The "Battle of the Forms": Fairness, Efficiency, and the Best-Shot Rule, 76 OR. L. REV. 155, 165 (1997) ("Others will presume that the random buyer they run into will not have read the form and that, by stacking the deck, the seller can perhaps gain more from the nonreaders that it loses to the readers."). See also Shmuel Becher, Asymmetric Information in Consumer Contracts: The Challenge is Yet to be Met, 45 Am. Bus. L. J. 723 (2008). See also Victor P. Goldberg, Institutional Change and the Quasi-Invisible Hand, J. Law & Econ. 17(2), 461-92 (1974).

[23] Clayton Gillette, Rolling Contracts as an Agency Problem, Wis. L. Rev. 679, 690 (2004) (noting that "[t]he question is whether such a group of informed buyers exists and whether it shares contract terms preferences with those non-reading buyers…Indeed, I will suggest that the most difficult issue in finding surrogates for nonreading buyers is that one set of buyers may have very different preferences from another set.").

[24] Robert Hillman and Jeff Rachlinski, Standard Form Contracting in the Electronic Age, 77 NYU L. Rev. 429 (2002).

consumers is likely to result in an under-production of knowledgeable buyers.[25] Others have pointed out that many consumers are unlikely to behave rationally in regards to boilerplate, thus making the "informed" minority less effective. Korobkin (2003) notes that even when choosing to become informed, boundedly-rational buyers are unlikely to consider *all* the contract terms of product attributes in making a purchase decision.[26] He predicts that sellers will offer one-sided *non-salient* terms.

2.3. Prior Evidence

Despite the theoretical importance of the informed minority, there has been little empirical investigation of its validity, presumably because observing "readership" is difficult. There is some related survey evidence. In a study of reading practices of online standard form contracts, Hillman (2006) surveys 92 contracts students at Cornell Law School and finds that only four percent of those who purchased products online claim to read standard form contracts "as a general matter."[27] Almost sixty percent of respondents, however, claim that they would be prompted to read the contract depending on the type of vendor, the price of the product purchased, and the term. Becher and Unger-Aviram (2009) survey 147 students in law and other areas and ask about the likelihood that respondents will read standard terms in different scenarios (presented as vignettes in the survey).[28] Sixty percent of respondents claim they skim or read parts of a standard form contract before entering a transaction. Similarly, in a recent study, Bartlett and Plaut (2009) survey 182 undergraduate students to find out the extent to which respondents read and understand standard form contracts and, to the extent that respondents

[25] See Salop, supra note 2. The authors also note that "[r]ational calculation alone cannot explain consumers' nearly universal failure to read standard-forms. In some circumstances, the market should produce a sufficient number of consumers who attend to the unlikely contingencies covered by the standard form such that businesses feel disciplined." Id. at 447.

[26] Korobkin, supra note 19. For an account of how consumers' focus on only those observable attributes of product quality distorts sellers' incentives to produce optimal quality products, see also Beales, Craswell, & Salop, supra note 2, at 510-11. For other behavioral accounts of why the existence of informed consumers might not result in competitive outcomes, see Laibson and Gabaix, and Bar-Gill, supra note 19.

[27] Robert Hillman, Online Consumer Standard Form Contracting Practices: A Survey and Discussion of Legal Implications, in Is Consumer Protection an Anachronism in the Information Economy? (Jane K. Winn, Ed. 2006).

[28] Shmuel I. Becher & Esther Unger-Aviram, Myth and Reality in Consumer Contracting Behavior, available at SSRN: http://ssrn.com/abstract=1117422.

report non-reading, to find out the reasons for that failure to read.[29] They find that about 80% of respondents claim not to read contracts and much of the remainder claims to "skim" them.

Although such surveys provide suggestive evidence, they are based on self-reported behavior or hypothetical commercial scenarios. The survey subjects are highly unrepresentative students, sometimes law students who someday will be writing boilerplate themselves, and who are, in a sense, being put on the spot. Our guess is that this leads to a large upward bias in the estimate of readership rates, but in any case evidence on actual behavior is required to answer the question with any confidence. A recent study by Hillman and Barakat (2009) provides some indirect evidence of failure to read; it reports that most of the 100 software publishers with top-selling software on Amazon.com offer express warranties on their website and disclaim such warranties in their EULAs.[30]

Three other studies show that standard form contract terms are less one-sided in favor of sellers than might be possible if buyers were completely uninformed, and thus provide indirect evidence for the existence of the informed minority. In a study of 62 warranties of an array of consumer durable goods, Priest (1981) argues that warranties are not biased towards sellers, but rather reflect by the relative ability of buyers and sellers to prevent and insure against loss.[31] 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.[32] In a study of contracting practices by online retailers, Mann and Siebeneicher (2008) find that few sellers offer excessively one-sided terms (and many present their contracts in ways that would be deemed unenforceable).[33]

---

[29] Robert P. Bartlett, III & Victoria C. Plaut, Blind Consent? A Social Psychological Investigation of the Non-Readership of Contracts (mimeo, 2009).

[30] Robert A. Hillman & Ibrahim Baraka, Warranties and Disclaimers in the Electronic Age, 11 Yale J. L. & Tech. 1 (2009).

[31] Priest, supra note 14.

[32] Florencia Marotta-Wurgler, What's in a Standard Form Contract? An Empirical Analysis of Software License Agreement, 4 J. Emp. L. Stud. 677 (2007); Florencia Marotta-Wurgler, Competition and the Quality of Standard Form Contracts: The Case of Software License Agreements 5(3) J. Emp. L. Stud. 447 (2008).

[33] Ronald Mann & Travis Siebeneicher, Just One Click: The Reality of Internet Contracting, 108 Colum. L. Rev. 984 (2008).

2.4. *Other Perspectives*

While this paper limits its scope to the validity of the informed minority hypothesis, it is important to note that law and economics scholars have also proposed other mechanisms that would induce markets to behave competitively when consumers are imperfectly informed. For instance, when sellers are constrained by reputation or the threat of litigation, they will find it in their best interest to offer terms preferred by buyers to protect their reputational investment. Such a mechanism could substitute for the informed minority. Sellers might also offer one-sided terms to all consumers, only to later relax them to accommodate reasonable buyer complaints.[34] Similarly, in the case of experience goods or repeat purchases, buyers who do not read terms might ultimately become familiar with the contents of the sellers' boilerplate. Our data are not conducive to assessing these mechanisms as opposed to the informed minority mechanism.

2.5. Current Legal Landscape

The law governing standard form contracts affords courts great flexibility in deciding whether to enforce consumer agreements involving standard terms. Those that conclude that sufficient market pressure (either from an informed minority or because sellers are constrained by reputation) exists to ensure competitive terms will enforce boilerplate.[35] For example, in *ProCD v. Zeidenberg*, mentioned above, Judge Easterbrook enforced a commercial-use restriction clause in a shrinkwrapped software license agreement, rejecting the buyer's argument that he couldn't have agreed to "hidden" terms. He reasoned that the burden is on the buyer to protect his own interests: "ours is not a case in which a consumer opens a package to find an insert saying 'you owe us an extra $10,000' and the seller files suit to collect. Any buyer finding such a demand can prevent formation of the contract by returning the package, as can any consumer who concludes that the terms of the license make the software worth less than the purchase price. Nothing in the UCC requires a seller to maximize the buyer's net gains."

Many courts, like many academics, are less willing to assume that buyers are careful shoppers. Many rely on the doctrine of unconscionability to invalidate apparently one-sided

[34] See Gillette, supra note 4; Lucian A. Bebchuk & Richard A. Posner, One-Sided Contracts in Competitive Consumer Markets, 104 Mich. L. Rev. 827 (2006).

[35] Enforcement will occur absent fraud, duress, or misrepresentation.

contract terms.[36] A term will be struck down under this doctrine if it is found to be procedurally and substantively unconscionable. The procedural aspect of the inquiry focuses on whether the buyer is deprived of an opportunity to meaningfully assent to the terms because the contract is hard to find, in miniscule print, hard to understand, or because the buyer lacks meaningful choice. The substantive aspect asks whether a particular term is so one-sided that it would "shock the conscience" of an informed buyer. Although this doctrine gives courts little guidance on how to effectively distinguish mutually beneficial clauses from exploitative ones, courts that believe that failure to read might result in seller abuse are somewhat eager to strike down terms.'[37] Similarly, they rely on the "reasonable expectations" doctrine to deny enforcement of clauses that are hard to understand or read, enforcing instead rights and obligations a reasonable consumer would anticipate.[38]

Furthermore, some state legislatures have mandated particular terms in consumer contracts. For example, Idaho,[39] North Carolina, and Montana[40] have statutes prohibiting the enforcement of forum selection clauses in consumer transactions.[41] California's Arbitration Act mandates a waiver of arbitration fees for low income consumers and requires arbitration

---

[36] Uniform Commercial Code (U.C.C.) § 2-302. Comment 1 states that "The principle is one of the prevention of oppression and unfair surprise…and not of disturbance of allocation of risks because of superior bargaining power."

[37] For example, the California Court of Appeals recently struck down as unconscionable an arbitration clause with a class action waiver that was visibly included in the fine print of a cell phone service agreement. After noting that contracts of adhesion should be carefully examined, the court stated that "[t]he possibility of overreaching is even greater in ordinary consumer transactions involving relatively inexpensive goods or services because consumers have little incentive to carefully scrutinize the contract terms or to research whether there are adequate alternatives with different terms, and companies have every business incentive to craft the terms carefully and to their advantage. The unconscionability doctrine ensures that companies are not permitted to exploit this dynamic by imposing overly one-sided and onerous terms." Gatton v. T-Mobile USA, Inc., 152 Cal. App. 4th 571, 585 (Cal. App. 1st Dist. 2007).

[38] Robert Keeton, Insurance Law Rights at Variance with Policy Provisions, 83 Harv. L. Rev 961 (1970); See also Ben-Shahar, supra note 21. See also Restatement (Second) of Contract § 211.

[39] Idaho Code § 29-110.

[40] N.C.G.S. § 22B-3; Mont. Code 36 § 18-1-403.

[41] Several courts also refuse to enforce forum selection clauses. See, e.g., Fidelity & Deposit Co. v. Gainesville Iron Works, Inc., 125 Ga. App. 829 (Ga. Ct. App. 1972). For a more detailed review of prohibitions on dispute resolution clauses, see Florencia Marotta-Wurgler, "Unfair" Dispute Resolution Clauses: Much Ado About Nothing? in (Boilerplate: Foundations of Market Contracts (Omri Ben-Shahar, ed., Cambridge University Press, 2007).

organizations to make all consumer arbitration decisions publicly available.[42] There are also several federal laws that, conceptually, aim to increase the size of the informed minority by decreasing reading and comparison shopping costs. The most famous is the Magnuson-Moss Warranty Act, a disclosure law enacted in 1975 to regulate the form and content of consumer product warranties.[43] It requires that sellers who provide warranties draft them in clear language and present them in a standardized fashion.[44] It also seeks to reduce shopping costs by requiring that warranties be available for inspection prior to purchase.[45] Yet despite these protective measures, the law is that buyers are under a duty to read standard form contracts and are thus deemed to have given "blanket assent" to reasonable terms whether they have read them or not.[46]

There are additional, recent proposals to protect consumers in mass market transactions, especially those that take place over the Internet, the context of this study, by seeking to facilitate contract access. Some of the most contentious proposals involve software and other information goods. For example, the American Law Institute (ALI) has approved new Principles of Software Contracts to harmonize and increase certainty of the laws governing software transactions online. One of its main goals is to "promote reading and the opportunity to read terms" as a way of alleviating market failures.[47] For example, the rules require software vendors to post the terms of their license agreements in their corporate website, thus effectively eliminating pure "pay now, terms later" contracts in Internet transactions (where buyers cannot access the contract until after

---

[42] Cal. Civ. Proc. Code § 1284.3, 1281.96.

[43] 15 U.S.C. §§ 2301-2312 (1976).

[44] 15 U.S.C. § 2302(a).

[45] 15 U.S.C. § 2302(b)(1)(a). Another prominent example is the Truth in Lending Act (TILA), a disclosure law that seeks to protect consumers in credit transactions by requiring a clear and standardized disclosure of the essential terms and costs associated with the deal. 15 U.S.C. § 1601.

[46] Karl Llewellyn, The Common Law Tradition: Deciding Appeals, 370-1 (Boston & Toronto: Little, Brown and Company,) (1960).

[47] ALI Principles of the Law of Software Contracts [hereinafter *ALI Principles*] at 117. The reporter explains that "because case reports and the websites of watchdog groups already evidence vendors' use of unsavory terms, [the] Principles assume that market pressure is insufficient in software retail markets to assure the production of reasonable terms, both in presentation and substantive content… The preferred strategy of [the] Principles is to draft rules that promote reading terms before committing to a transfer, which, in turn, should decrease the instances of market failure." ALI Principles at 115. For a detailed account of these proposals and an analysis rejecting the "opportunity to read" approach, see Omri Ben-Shahar, supra note 21.

they had paid for the product).[48] In addition, the ALI rules would deem many browsewraps—contracts referenced by hyperlinks that don't have to be expressly agreed to, such as websites' "Terms of Use"—unenforceable.[49]

Another well-known effort to create a uniform and cohesive body of law for computer information including software is the Uniform Computer Information Transactions Act (UCITA). This is a model act drafted by the National Conference of Commissioners on Uniform State Laws. The act has been met with strong opposition academics and consumer advocates, and it has so far been enacted only in Maryland and Virginia.[50] One of the reasons UCITA has encountered resistance is that, in contrast to the ALI's *Principles*, it allows enforcement of "pay now, terms later" contracts as long as buyers are made to assent to the terms (such as a EULA) and are able to return the software after having had an opportunity to review them. But ultimately, whether the approach endorsed by the drafters of the *Principles* or the drafters of UCITA is the most appropriate depends on whether there is market failure and whether increased disclosure would ameliorate it at acceptable costs; this study sheds empirical light on these questions.

3. Research Framework

Given the various theoretical arguments and significant practical consequences for courts and consumers, it seems critical to investigate whether an informed minority of buyers capable of disciplining the market actually exists. Our approach is to study the browsing and shopping behavior of online consumers. We track the behavior of Internet visitors to 66 software

[48] For a study of whether sellers who use "pay now, terms later" contracts offer more one-sided terms than sellers who disclose their contract prior to purchase, see Florencia Marotta-Wurgler, Are "Pay Now, Terms Later" Contracts Worse For Buyers? Evidence From Software License Agreements, 38(2) J. Leg. Stud. (2009).

[49] *ALI Principles* §2.02(b). The comments to this section state that "[f]or example, mere reference to standard terms found on another page (browsewrap) may be insufficient under the reasonable-transferor test unless the transferee is already well-acquainted with the terms, for example, from previous notices and transactions", at 124. In some circumstances, even terms that appear in a hyperlink next to a box where the consumer must click on "I agree" could also be unenforceable if the consumer was unaware of the terms because of a lack of previous dealings with the seller. ALI Principles at 125.

[50] See Va. Code Ann. §§ 59.1-501.1 to 59.1-509.2 (Michie. 2001); Md. Code Ann. Comm. Law §§ 22-101 to 22-816 (Supp. 2002).

companies. 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 to look for the informed minority. First, while non-price features such as the associated contractual rights and restrictions are important for all types of products, it is 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 heavily litigated in the past decade.[51] 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. To the extent 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, as mentioned above, several recent debates on legal reform in standard form contracts focus on electronic contracts in general and software contracts in particular. A study of the informed minority in online software markets places us at the center of these debates.

In order to empirically investigate the existence and size of the informed minority, we classify visitors to the websites of the companies in our sample, described below, into potential buyers and those visiting for other reasons. Inspection of a sample of URLs clicked shows these reasons to include the following: to look for online instruction manuals, perhaps for a product they already own; to search user forums for troubleshooting information; to watch entertaining advertising such as the "Mac vs. PC" commercials; and so on. We denote by s the fraction of potential buyers ("shoppers"); non-shoppers make up the remaining fraction $1-s$. We denote by e_1 the fraction of shoppers and by e_2 the fraction of non-shoppers that read the online EULAs. Finally, we denote by b_1 the fraction that purchase the product ("buyers") among shoppers that read the EULA and by b_2 the fraction of buyers among shoppers that do not read the EULA.

[51] See, e.g., M.A. Mortenson Co. v. Timberline Software Corp., 998 P.2d 305 (Wash. 2000); Davidson & Assoc. v. Internet Gateway, 344 F. Supp. 2d 1164, 1178 (D. Mo. 2004); Altera Corp. v. Clear Logic, Inc. 424 F.3d 1079 (9th Cir. 2005)); Blizzard v. BDNet, 422 F.3d 630 (8th Cir. 2005).

This framework is depicted in Figure 1. In this setting, the informed minority corresponds to the fraction e_1 of shoppers that reads the online EULA.[52]

In what follows we estimate the number of visitors in our sample for each of the six categories shown in Figure 1. That is, we estimate the number of readers and nonreaders among visitors classified as buyers, shoppers and non-shoppers. We use access to a EULA page for more than 1 second to identify readers; this method is likely to bias upward our estimate of the informed minority in that some EULA accesses are accidental, are inconsequential to the buying decision, are accessed so briefly that little content could have been grasped, or are read but the legal jargon is not understood. We use initiation of a secure checkout process to identify buyers. We use other contextual information to distinguish shoppers from non-shoppers.

We can break down readers into se_1b_1 readers that buy and $se_1(1-b_1)$ readers that 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 non-shoppers read EULAs and thus we expect $(1-s)e_2$ to be small. Finally, the fraction of non-shoppers that 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 in the informed minority as $\frac{se_1b_1 + se_1(1-b_1)}{se_1b_1 + se_1(1-b_1) + s(1-e_1)b_2 + s(1-e_1)(1-b_2)}$. 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 clickstream data set represents the browsing behavior of 92,411 U.S. households for January 2007. These data were made available to us by a major online research company which

[52] It is possible that for some shoppers accessing the EULA will not affect their probability of buying the product. For instance, some shoppers may not know what a EULA is, or discover after accessing the EULA that they are not capable to comprehend its language, or may access it 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 informed minority fraction of shoppers. We can explore the significance of this to some degree by studying time spent on the EULA page by those that access it. On the other hand, because we don't consider other ways in which shoppers might become informed about the terms (e.g., 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.

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 webpage 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 very large dataset.[53]

The panel of households was selected to be demographically and geographically balanced and representative of the population of U.S. households with Internet access.[54] The information captured for each web page visited by a panelist in the raw data is coded with 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 page visited, the time that webpage was accessed, the time spent on that page, whether that page was within a secure (i.e., 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 comprise the bulk of the data, but we were provided with useful additional files included non-personally identifiable demographic information about the panelists, and a corporate hierarchy identifying the parents, if any, for the divisions or companies owning the web servers that appear in the data (e.g., office.microsoft.com and mail.hotmail.com are properly identified as companies or divisions having the same corporate parent, Microsoft).

4.1. Sample Construction

Within these data we consider one important market and one important contract within that market. Specifically, we study user visits to software companies that sell or distribute their products through their corporate websites and make their EULAs available on their site for users to peruse at their option, prior to any purchase decision. We use the data provider's classification

[53] Information was captured for 6,355,922 user sessions in January 2007, with 461,027,284 corresponding web page views.

[54] This data provider's panel is considered one of the largest representative media research samples in existence. The sample of participants is defined using Random Digit Dialing principles: the company selects a random set of phone numbers from all available residential numbers in the U.S. and attempts to recruit each at most 15 times at different times of the day and on different days. The panel also includes university students and individuals in the workplace. The company updates its demographic information regularly, has implemented various procedures to keep the panel updated, and ensures that tracking is unobtrusive to prevent any distortions in behavior. Finally, each January, the panel is compared to the U.S. Census Bureau data to maintain its representativeness, so we choose January for our analysis (this choice, however, is quite inconsequential to our results).

of markets to identify visits only to software companies. 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 website. Freeware providers offer their software for free to anyone wishing to download it. Examples include browser toolbars, plug-ins, and browsers.[55] We are interested in observing users' propensities to become informed about the terms of these two types of software.

For the purpose of 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 that viewed at least two pages during their visit; our interest is in users with intent or potential intent to purchase, or "shoppers," and users that 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 web page addresses (URLs) of all EULAs available on the company's website. To find these we visited each company's website and used manual browsing, Google searches within the website and, if available, searches of the website provided by the company. In addition, we searched all page views in the clickstream data corresponding to these companies to identify possible EULA pages (e.g., pages whose web address contained "EULA" or "legal" or "terms"), which we then investigated manually. Finally, we remove companies that did not make their EULAs available online as well as companies that *required* users to agree to their EULAs during their checkout process (by making them click "I agree" below a text box with the EULA or next to a hyperlink with the EULA). After excluding all these companies that do not provide enough data or otherwise are inappropriate for our tests, we arrive at a final sample of 56 retail and 10 freeware companies. We mention here that we see no reason to believe that our basic results or conclusions would change significantly were we to enlarge the number of companies in the sample or the time window that panelists were followed. Also, the size of our sample is probably more usefully characterized in terms of the tens of

[55] We classify a company as retail if it offers its core or much of its software for sale, even if it also offers software for free. For example, Adobe offers several free plug-ins, such as Shockwave and a PDF reader, but we classify it as retail.

thousands of company visits that we track, described shortly, 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 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 are from Hoovers.com, Yahoo! Finance, or direct communications with the companies in the sample.

Panel A in Table 1 reports summary statistics for the company characteristics for the two types of companies analyzed. For retail companies, average revenue is $2.14 billion with a standard deviation of $8.19 billion, a number obviously driven by a few large firms. Median revenue for this category is $13.8 million. The mean age of these companies, measured as 2009 minus the year of incorporation, is 17.3 years old (median is 15). Thirty percent of the companies in this category are publicly traded. In contrast, the mean age of freeware firms is 10.2 (median is 7), and only ten percent of these companies are publicly traded.

We collect several product characteristics. For each company, we record one "flagship" product per company. Many small and medium size companies market one main product, in which case we select that product. For larger companies, we select the product accounting for the largest fraction of sales or, when this information was not available, we selected the product most prominently featured on the website. A reasonable hypothesis is that users are more inclined to become informed about the EULA terms of higher priced products, so we record the price for the flagship product as well as median price of all products available on the website for that firm. We record whether the product is a single or multi-use license, because multi-seat 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.[56] We also note

[56] Trial versions are generally offered with limited functionalities over a limited time period. Marotta-Wurgler (2007) found that the majority of trial licenses are noticeably different from the product licenses (e.g., the trial license reads "Trial License" and is generally shorter than the product license), such that a user would not consider them substitutes.

whether the product is oriented toward business users or the general public. Finally, we classify each product into one of 150 software product categories, e.g. antivirus or word processing, based on the characterizations of software products available at Amazon.com.

Panel B reports summary statistics for the flagship products' characteristics. The average product price for retail companies is $484 and the median is $64. For each company we compute the median price of all software products listed on the website, and the mean of those medians is $389; the median of the median is $47. A fraction of 68% of retail products and 90% of freeware appears targeted to the general public (or very small businesses) as opposed to large businesses. Finally, 86% of retail sales sites offer a trial version of their featured product or of the product in our sample, and 80% offer trial versions for most of their products.

4.3. Contract Characteristics

We want to measure the fraction of shoppers that become informed about EULA terms. We thus collect *all* the EULA URLs that are available on a company's website. As noted above, many firms only sell one product and thus they only make available online 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 website, 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 56 retail companies and 34 unique URLs corresponding to EULAs for for our 10 freeware companies.

4.4. Defining Shoppers and Shopping Visits

Panelists in our data unfortunately don't come labeled as "shoppers" when they visit a given company's website. We must therefore define shoppers, i.e. visitors with some potential to purchase. Our data provider reports *all* the Internet browsing activity of its users and a large fraction of visitors may be browsing without any intent to purchase; potential motivations behind such browsing were mentioned above. This is more of an issue with larger diversified websites that provide significant non-product oriented content such as Microsoft, Adobe, or Symantec. Smaller companies generally have sparser websites focused on supporting the purchase process. As noted earlier, our data provider tracks the particular web servers that host the URLs accessed, which allows us to exclude visits clearly unrelated to shopping.

To restrict our analysis to visitors with potential intent to purchase, we exclude visits that do not access servers dedicated to shopping or purchasing activities. For example, Microsoft, Adobe, and Symantec locate their user forums, software patches and rebate pages on separate servers, such as adobeforums.com and symantecrebates.com. We then adopt a variety of additional approaches to identify shopping-oriented visits more precisely.

We define a "user visit" as all page views (URL accesses) from a company's website within a single user "session." One way of identifying shoppers is by examining the intensity of a company visit. A user with intent to purchase is likely to view several pages in the retail side of the company's website. Our broadest definition of a shopping visit is a visit with at least two page views, as a minimal definition.[57] A second, more restrictive definition includes all visits by users who accessed at least five pages in a given company's website. This is progressively more likely to exclude casual browsers.[58]

At the other extreme, a visitor that 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 and subsequently recognizing for the 66 companies in our sample the web page addresses that would be utilized only during the checkout and payment process.[59] While knowing that a user started a checkout or payment process provides no guarantee that the transaction was completed, it indicates an extremely high likelihood that a transaction was at

[57] The literature that studies Internet shopping and browsing behavior generally excludes visits to companies with only one page view, as they have been found to be mostly mistakes or the result of redirects. Instead, the two-URL definition has been used extensively in the literature to define a valid company visit. See Randolph E. Bucklin & Catarina Sismeiro, A Model of Web Site Browsing Behavior Estimated on Clickstream Data, J. of Marketing Res. Vol. XL, 249-267 (2003); Wendy W. Moe & Peter S. Fader, Dynamic Conversion Behavior at e-Commerce Sites, Mgmt. Sci. 50 (2004); Catledge, Lara D. and James E. Pitkow, Characterizing Browsing Behaviors on the World Wide Web, Computer Networks and ISDN Systems, 27 (6) (1995), 1065–73.

[58] Several studies have shown that visits with higher page views are more likely to result in purchases. Bucklin & Sismeiro, supra note 57; Wendy Moe, Hugh Chipman, Ed George, & George McCulloch, A Bayesian Tree Model of Online Purchasing Behavior Using In-Store Navigational Clickstream Data (working paper 2002). Moe et al identify different browsing behaviors by shoppers to a particular website and find that shoppers view an average of 4.01 pages in a given session.

[59] Some companies have integrated handling of part or all of the checkout and payment process within their own website, while other companies outsource parts of the checkout and/or payment processes. For instance, when a visitor proceeds to checkout, they may be redirected to a company like Digital River that will process the transaction. This mode of checkout is common. Sometimes, in addition to the checkout company, a payment company is involved in completing the transaction such as PayPal. We account for these situations.

least contemplated. This definition of shopping visit, relative to the other two, is overly restrictive, as it excludes those that do not result in the initiation of a checkout process.[60]

A possible concern with the last definition above is that it would exclude shoppers who access the EULA and decide not to purchase, which would bias the estimated size of the informed minority if the conversion rates for readers and non-readers differ. If we are willing to assume that shoppers with true intent to purchase may visit several merchants, but in the end will purchase from *some* merchant (which may or may not be included in our sample), and that among these shoppers the ones that constitute the informed minority (i.e., those that access EULAs) are equally likely to do so in any of the merchants they visit, then the behavior of the visitors that initiate a checkout session will be representative of the shoppers as a whole. Their likelihood to access EULAs then provides an appropriate estimate for the size of the informed minority. These assumptions seem reasonable, given that our sample consists of most or all major software vendors and several smaller ones, and that browsing behavior is followed for a reasonably extensive period of time. We also address these points in Section 5.2.

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

A last issue with defining a shopping visit is ascertaining the length of time of a particular visit. Our data provider and the industry in general define user sessions as periods of web browsing activity separated by at least 30 minutes of inactivity. Under this definition, a user can have multiple visits to a given company in a day, a week, or a month. We adopt this definition to be consistent with prior literature.[61] We refer to all page views from a unique company's website within a single user session as a "company visit" by that user. For example, consider an uninterrupted session where a particular user first visits Symantec (a company in our sample), then Banana Republic, and then McAfee (another company in our sample). This session yields

[60] Given the low conversion rates in electronic commerce, such visits are likely to represent the majority of shopping visits.

[61] See Moe & Fader, supra note 57.

two unique "company visits" in our sample: one to Symantec and one to McAfee. If, after visiting McAfee, the user goes back to Symantec within the same uninterrupted session, we aggregate that second visit to Symantec with the first visit to that company. If this user is deciding whether to buy from McAfee or Symantec, this aggregation method still allows us to see whether the user accessed EULAs when deciding which product to purchase.

Despite its popularity in this literature, the uninterrupted session measure might be too narrow. It is conceivable that a user's shopping activity on a given company spans several days or even weeks. Research on Internet shopping behavior reveals that because visiting and "traveling" to a store on the web is free and simple, users are more inclined to visit the company several times over an extended period before finally deciding whether or what to purchase.[62] Thus, users accrue information about a product over time and across several visits. If this is the case, then the uninterrupted session measure will overestimate the number of visits with intent to purchase. To account for this, we adopt the methodology of Johnson et al (2004).[63] They conclude that repeated visits to a company within a month typically correspond to the same shopping cycle. We thus aggregate visits to a unique company in a given month and present these aggregated sessions as an alternative measure of a company visit with intent to purchase.[64]

4.6. Demographic and Geographic Data

To identify characteristics of shoppers and shopping households that affect their likelihood of becoming informed about standard terms, we also utilize personal information about them. We have the age and sex of the head of the household, household income, household size, and whether there are children present in the household. Table 2 reports summary statistics.

---

[62] Id.

[63] Eric J. Johnson, Wendy W. Moe, Peter S. Fader, Steven Bellman, & Gerald L. Lohse, On the Depth and Dynamics of Online Search Behavior, 50 Mgmt. Sci. 3, 299-308 (2004) (finding that less than 1% of all month-long sessions in their sample contained more than one purchasing transaction in a given company).

[64] A possibility here is that, for some users, we might compound multiple shopping visits into a single one, thus undercounting the number of shopping visits (and, conversely, overcounting the fraction with EULA visits). On the other hand, this measure may allow us to include repeated short visits to a given company that would be excluded under the alternative measures of visit. As noted earlier, we expect that actual shopping visits will lie somewhere in the middle. In any case, the results for the various definitions of visits are similar.

Panel A includes visitors who accessed a minimum of two pages in at least one of the companies in the sample during a single uninterrupted session. The sample is comprised of 45,091 unique visitors. The average age of the users in this group is 46, and the range is reportedly from 18 to 99. Average income for heads of households is $60,487 with a standard deviation of $39,666. Income (and perhaps age) is topcoded; median income ($37,500) better describes the sample. About half of the heads of households are male. The average number of household members is 2.8. There are children in 41% of these households.

Panel B reports summary statistics for the sample of visitors who accessed a minimum of five page visits in at least one company. Given that this is a more restrictive measure of a visit with intent to purchase, the sample drops to 31,969 unique household user-visitors. The characteristics of these users are similar to those above. Panel C reports summary statistics for user-visitors that have selected a product for purchase and have begun the checkout process. Here there are 1,653 unique user-visits. Since online conversion rates are generally less than two percent across all goods, the size of this sample relative to all shopping visits is a little larger. The users in this subsample are also similar to those in the others.

5. Results

Our analysis here is based on company/shopping visits 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 of company visits and EULA accesses under 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 EULA Accesses

Recall that we wish to measure the fraction of buyers that *seeks* to become informed about EULA terms in deciding whether to purchase, and as noted we have defined the sample to include only those company websites where EULA access is possible but optional. Tables 3 and 4 summarize the characteristics of visits to such companies, measured either as uninterrupted sessions (Table 3) or 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 company/shopping visit. 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 the number of company visits under each definition, the left halves of these tables show the number of pages viewed during such visits, and the duration of the visits in seconds. In the right halves, we tabulate the subset of these visits that included a EULA access, the number of pages viewed before the first EULA access, and the length of time spent viewing EULAs in visits where a one was accessed. These last two measures give us some indication of shoppers' level of care or intent in accessing EULA pages.

Looking at uninterrupted session/visits (Table 3), under the least strict definition of a visit (2 or more pages accessed), there are 120,545 such visits to software retailers and 28,007 to freeware providers, including repeat visitors. For retail companies, an average visit consisted of 12.4 page views over 311 seconds (5.2 minutes). These numbers, however, are driven by extreme values. The median number of pages visited in any given company is 5 and the median time spent is 105 seconds (1.75 minutes).

The data indicate that EULAs were accessed in only 55 of the 120,545 visits to software retailers (0.05% of all such visits) and in 40 visits to freeware companies (0.14%). Users that accessed EULAs visited an average of 12.2 pages (median of 7 pages) in that company's site prior to the EULA page. These figures are already telling, but another consideration is whether shoppers who access the EULA actually read it. For users in this group, the average time on the EULA page was 47.7 seconds and the median time was 29 seconds. (Note that we are defining "access" as a EULA visit of at least one second, for purposes of obtaining a conservatively high number of EULA accesses.) The average number of words of EULAs for retail products in the sample (unreported) is 2,277 with a standard deviation of 1,148 words. The time spent on the EULAs relative to their length indicates that most readers did not read terms in their entirety. The average reading rate of American adults is 250 to 300 words per minute, so a complete read of the typical EULA would require 8 or 10 minutes, not less than one minute.[65] In other words, even the small number of EULAs accessed in our sample is still likely to be an overestimate, probably a substantial overestimate, of the number of *effectively* informed readers.

[65] See Bailey, R.W & Bailey, L.M, Reading speeds using RSVP, User Interface Update (1999).

Visits to freeware providers have fewer page views (the median is 4 pages) and are of shorter duration (median time spent is 43 seconds). This is expected, as freeware sites tend to be sparser. EULAs are accessed in 0.14% of these visits. The median time spent on EULAs is also 29 seconds, against a median length of these EULAs of 1,754 words, so an overcount of effective readers is likely here also.

When a visit is defined to require five or more pages accessed at the company visited, there are 67,769 uninterrupted session/visits to software retailers and 13,520 to freeware companies. The median number of pages viewed in a given visit to a retailer is now 10 pages and the median length is 185 seconds (3.1 minutes). Distributions of page views and duration are again skewed. EULAs were accessed at a slightly higher rate in these visits, 50 times among software retailers (0.07%) and 30 visits among freeware companies (0.22%). The median number of pages seen before accessing a EULA was 8 for retailers and 4 for freeware providers.

Finally, limiting our consideration to visits to software retailers that included initiation of a secure checkout session, the number of visits falls to 5,509, with similar median page views per visit, but about twice as long mean and median durations. This is expected since purchases require more time to process the transaction. In this restricted sample, there are 5 voluntary accesses of a EULA in the course of purchase, constituting 0.11% of all visits. The median number of pages accessed is 20 for users in this group, suggesting even more intense shopping within the site. The median time spent in the EULA also doubles for users in this group. Interestingly, out of *all* sessions with EULA visits, 4% (if we use the two page visit definition) or 8% (if we use the five page visit definition) resulted in purchases. These numbers are significantly higher than the typical 2% conversion rate in Internet purchases.

Aggregating all monthly sessions of an individual user into a monthly visit (shown in Table 4) leads to similar results. In most cases the total number of visits is reduced as multiple visits by individual users are combined. The average number of sessions per user is 3.2 (unreported). Not surprisingly, the results for the most inclusive definition of a visit, at least two page views, do indicate that this category captures a nontrivial number of casual browsers with little intent to shop. An exception is software retailer visits that included initiation of a secure checkout session. Visits with secure checkout increase, albeit moderately, because combining visits for certain users on a monthly basis resulted in a qualifying monthly visit replacing two or more non-qualifying uninterrupted session visits. The overall results of Table 4, 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 0.22%, meaning that there are about two readers per every one thousand shoppers.

5.2. Interpreting the Results: Can this be an Informed Minority Equilibrium?

Coming back to the empirical framework of Figure 1, visitors to the websites of the companies in our sample can be classified into potential buyers or users visiting for other reasons, such as looking for online instruction manuals for a product they already own, to search user forums for troubleshooting information, or for entertainment—e.g., to watch the "Mac vs. PC" commercials. We measure the total number of page views during each visit, as well as whether a EULA was accessed and whether a secure checkout session was initiated. This data, reported for individual sessions in Table 3 and for monthly visitors in Table 4, allows us to estimate the number of readers, buyers and shoppers by using access to a EULA page as a proxy for reading, initiating the checkout process as a proxy for buying, visits with 5 or more page views as a proxy for identifying shoppers and visits between 2 and 5 page views as a proxy for identifying non-shoppers. Based on 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.

The se_1b_1 readers that buy and $se_1(1-b_1)$ readers that do not buy are 6 and 43, respectively. There are $s(1-e_1)b_2$ or 3,528 buyers that are not readers, and $s(1-e_1)(1-b_2)$ or 37,120 shoppers that neither read nor buy. Few non-shoppers would be expected to read EULAs, so it is not surprising that $(1-s)e_2$ is small; in our sample it equals 4 (out of 22,575 visits). Finally, the large majority of non-shoppers do not read EULAs; this number is $(1-s)(1-e_2)$ or 22,571 based on the above proxies. We thus arrive at an estimate for the fraction e_1 of shoppers that are in the informed minority of $\dfrac{se_1b_1 + se_1(1-b_1)}{se_1b_1 + se_1(1-b_1) + s(1-e_1)b_2 + s(1-e_1)(1-b_2)} = \dfrac{49}{40{,}697} = 0.12\%$.

It is possible that considering all visitors with 5 or more page views as shoppers will overestimate 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 non-readers is the same as that for readers at 6/49 = 12.2% (which is higher than purchase conversion ratios of 2-5% 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 for all shoppers would be the same as the fraction for buyers, i.e. 6/3534 = 0.17%.[66,67]

The bottom line is that the fraction of visitors that access EULAs is very small, on the order of 0.1%. While a number of alternative estimates can be calculated, these estimates point to that fraction being well under 1%. 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 offer an example where the informed minority needs to be 20% to 30% to be effective. Our estimates here are imperfect, but they are two orders of magnitude smaller.[68]

Theoretically, the size of informed minority required to induce sellers to provide good terms depends on the tradeoff 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. Specifically, consider a seller that may offer standard contract terms that are more or less favorable to the buyers, which we will call respectively "good" and "bad" terms. The fraction r of buyers that 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 based on the

[66] Most models of the informed minority predict that the conversion ratio for non-readers would be the same or higher as the conversion ratio for readers, as the latter will be 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 $b_2 = 100\%$, i.e., that 100% of non-readers proceed to purchase the product. In that case the informed minority would be 49/(3534+43) = 1.37% of the total number of shoppers.

[67] As mentioned in section 4.4, if we assume that "real" shoppers will purchase from some merchant (while they may visit many), and that among these shoppers the ones that constitute the informed minority are equally likely to access a EULA in any of the merchants they visit, then the behavior of visitors in our sample that initiate a checkout session is representative of shoppers as a whole, and their likelihood to access EULAs (0.17%) provides an estimate for the size of the informed minority among these most determined shoppers.

[68] The estimates presented above are based on monthly visits as reported in Table 4. This is conservative in the sense that using visits defined as individual sessions would result in lower estimates for the size of the small minority. Specifically, using session data from Table 3 would result in 5 readers that buy, 45 readers that do not buy, 5,504 buyers that are not readers, and 62,215 shoppers that neither read nor buy, and 52,776 non-shoppers that include only 5 readers. The fraction e_1 of shoppers in the informed minority would be 50/67769 = 0.074%. The fraction of readers that initiate checkout sessions would be 5/50 = 10%, and assuming the same conversion ratio for non-readers would give 5/5509 = 0.09% as the informed minority. A conversion ratio of 100% for non-readers would give an upper bound for the informed minority of 50/(5509+45) = 0.91%.

characteristics of the setting (e.g., based on buyer search strategies as in (Schwartz and Wilde 1979)). In our sample, this corresponds to the fraction e_1 of shoppers that are in the informed minority. Buyers value the rights and restrictions incorporated in the standard form contract (e.g., 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 terms, resulting in corresponding product costs of c_g and c_b, with $c_g > c_b \geq 0$. Amending our earlier notation, 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 based on 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 terms is:

$$r \geq \frac{c_g - c_b}{\left(1 - \frac{b_1}{b_2}\right)(c_g - c_b) + \frac{b_1 - b_3}{b_2}(p - c_b)}.$$ [69]

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

This general theoretical conclusion is rather unhelpful by itself, since 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 put our observed fraction of readers into perspective. M&S is

[69] An outcome where 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 to provide them.

a key term in software EULAs,[70] and thus the cost of supplying M&S should be an order-of-magnitude approximation of the cost of offering good EULA terms.[71]

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 (i.e., did not charge per incident). On average, M&S were priced at 26% of the product price (exclusive of the M&S). The median is 20% and the standard deviation is 22%. Since there was high inter-company correlation, we focused on company means. Figure 2 shows the distribution for the 40 companies remaining after dropping two outliers with too-high ratios, which has a mean of 0.29, median of 0.24, and standard deviation of 0.16.

Thus a year of M&S for software products[72] is on average priced at 25-30% of the product price. 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% of product price may be too high an estimate: 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 upfront purchase (the software product) and a higher price for the subsequent purchase (M&S); consumers that purchase M&S are likely to have higher M&S costs due to adverse selection and/or moral hazard; software companies may have substantial market power in providing M&S due to 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 the above factors would result in a M&S-to-product price ratio that is higher than the cost

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

[71] The analysis of 647 EULAs in Marotta-Wurgler (2007) gives M&S an average value of 0.68 on a scale from 0 to +1 where 0 indicates the default terms in the absence of any EULA provisions (no M&S) and +1 indicates free M&S for 31 days or more. About three quarters of the sample companies commit to a free (i.e., included in the base price of the software) M&S period in their EULAs with these periods ranging from 60 days to 2 years, with a mean of 292 days and a mode and median of 1 year. Thus M&S provisions in EULAs are significantly more favorable to consumers than default, and likely to constitute an important fraction of the cost of offering pro-consumer EULA terms.

[72] As mentioned above, one year was the most common as well as the median duration of free M&S for the companies that provided such a period of free M&S.

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% of the product price, or 0.2*p* in the notation of section 3.1. Dividing numerator and denominator of the fraction of informed buyers necessary to induce the seller to offer good terms by *p* we get $r \geq \frac{(c_g - c_b)/p}{\left(1-\frac{b_1}{b_2}\right)\frac{c_g - c_b}{p} + \frac{b_1 - b_3}{b_2}\left(1-\frac{c_b}{p}\right)}$. If $(b_1 - b_3)/b_2 \leq 1$, which would be the case if readers purchase with the same probability as nonreaders when they discover good terms but 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 $r \geq (c_g - c_b)/p$, or $r \geq 0.2$. This is 200 times larger than the value of $r \approx 0.001$ (or 0.1%) that we observe in the data. Alternatively, our data suggest that in order for the informed buyers to induce the seller to offer good terms, the incremental cost of these terms would have to be almost negligible at under 0.1% of the selling price.[73] In either case, our data are unlikely to be consistent with an informed minority equilibrium.

A similar and perhaps much simpler approach to whether we might be observing an informed minority equilibrium here is to look at the seller calculus for visitors that initiate a checkout session. In an informed minority equilibrium sellers would be offering good terms and according to Table 4, EULAs are accessed 6 out 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% of these readers. Thus, if sellers are trading off the net revenue from the sales to the 6 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 0.17% of the selling price. This

[73] Of course, one could argue that offering a pro-buyer dispute resolution clause might cost the seller 0.1% of price. Although it is theoretically possible, we find it unlikely that sellers will change this clause to in response to the threat of losing one in one thousand buyers. More important, the terms that matter the most to consumers (such as M&S and warranties) are among the costliest ones.

argument relies on fewer assumptions than that developed above and once again leads to the same conclusion, namely that our data seem inconsistent with an informed minority equilibrium.

5.3. Becoming Informed Without Reading?

Some consumers may become informed about EULAs by consulting other websites instead of reading the contract.[74] The Internet contains several consumer product review sites, blog posts with rants about product quality, and specialized news outlets that could perhaps discuss the content of standard form contract terms. While casual observation suggests that this behavior is unlikely to be widespread, our data allow us to investigate it more directly.

From the website www.alexa.com 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 accessed these sites. Very few sites focus exclusively on software and EULA terms. Other sites that occasionally discuss EULA terms are technology-related news sites, such as Wired News and Ars Techica, sites that offer general consumer protection news and rants about abusive practices by sellers such as The Consumerist and BoingBoing, and general technology news such as PC Magazine. Finally, there are general consumer reports that do not focus on software but contain software product reviews, such as Consumer Reports.

We review the particular pages accessed by shoppers in each URL from these 25 sites to make sure that the pages accessed contain information about EULAs. We find that out of the 148,552 sessions with at least two pages accessed, only three shoppers accessed pages with EULA information in consumer review sites.[75] At the monthly level, 11,657 (16%) shoppers accessed at least one of the 25 consumer sites, but not a single aspect of that activity was related to EULAs. Of this group, only 69 shoppers accessed pages with particular software product reviews or information such as tax software review pages in that month.[76] A total of 84 shoppers

[74] See Shmuel I. Becher & Tal Zarsky, "E-Contract Doctrine 2.0: Standard Form Contracting in the Age of Online User Participation," Mich. Telecomm & Tech. L. Rev. 4(2) (2008).

[75] 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.

[76] A majority of these shoppers accessed reviews on epinions, Consumer Research, and Consumer Reports on accounting, tax preparation, and photo editing software.

accessed news or general information pages about software, such as how to obtain certain free plug-ins. The remaining visits were to pages unrelated to software or EULA terms.

This investigation confirms anecdotal evidence. It is highly unlikely that shoppers are, to an important extent, becoming informed about EULA terms by consulting other online sources, suggesting we do not need to qualify our previous conclusions about the existence of an informed minority equilibrium.[77]

5.4. Determinants of EULA Visits

Although few potential shoppers in our sample actually click on the EULA, it is interesting to examine what characteristics of the company, product, user, and website that distinguish the readers (or, more precisely, the "clickers") from the non-readers. We therefore estimate logit regressions with the dependent variable being a dummy whether a EULA was accessed during a particular company visit.

The results are in Table 5. The freeware dummy is positive and generally significant, as expected given earlier results. Some consumers may fear that there is a "catch" in products offered for free. The coefficient for the median product price is positive and significant in several samples, as expected in the sense that the terms associated with expensive products involve higher costs to the consumer. However, this is not the case for those visits where a checkout process is initiated. It is possible that buyers of expensive products are committed to the product already, or else lack useful alternatives, and so are less interested in the standard terms.

Consumers may be less likely to access EULAs from companies they trust, where trust may be proxied for by company size or degree of familiarity. We consider as regressors the natural log of a company's revenues (as a proxy for size) and a public company dummy. The coefficient for company revenue is significantly negative, while the public companies dummy is

[77] There are still a few other ways to become "informed" without reading. One is word of mouth, although it is farfetched that this works on significant scale. Another is the mass media. A particularly oppressive term might gain notoriety. This actually happened in February of 2009 when a consumer protection site noted that Facebook changed its terms of use in a way detrimental to consumers. The story was picked up by major news outlet and ultimately Facebook changed its terms. Although this reputational mechanism is interesting in that it can work with very few readers, it is not the informed minority mechanism that is the focus of this paper. And of course it is not relevant for the hundreds of less-known software sellers whose standard form terms would be of little interest to the popular media.

actually positive; this provides mixed evidence on the "company familiarity" or reputation hypothesis.

Extensive visits with a large number of page views may have increased likelihood of including a EULA access, as, for instance, they may be more likely to represent serious shoppers that are likely to access the EULA as part of their due diligence on their prospective purchase, so we include the number of (non-EULA) pages accessed as a regressor. This has a generally positive effect on the probability of accessing a EULA page.

Finally, high-income visitors may be better able to understand the language (and importance) of EULAs and thus more likely to access them. Alternatively, high-income visitors have a higher opportunity cost, and actually be less likely to spend time reading a EULA. We include the natural log of income, age, and a dummy for male gender (of the head of household). Women are insignificantly more likely to access EULAs than men, and the effects of income and age are positive and occasionally significant.

In unreported regressions we examine whether the option to download the trial version of a product affects user's propensity to access the EULA. As noted in Table 1, a large fraction of sellers offer trial versions. A plausible hypothesis is that users that become familiar with the trial version of a product might be less inclined to read the EULA of the retail version, perhaps due increased familiarity and comfort with the product. There is virtually no relationship between the presence of trial versions and users' likelihood of accessing EULAs for the broadest definitions of visits, but when visits are defined as beginning the checkout process, the coefficient on this variable is negative and highly significant.

Finally, we study whether shoppers are less likely to read the EULAs of products that are more likely to be purchased repeatedly. Users that become familiar with a product that is continuously updated, like Microsoft Office, may feel less need to concern themselves with the EULA. Other products, such as test preparation software, are less likely to be purchased repeatedly. We create a dummy variable that equals one if the company markets products that are in our judgment likely to be repeat purchases. However, we find no relationship between the nature of the use of the software and users' propensity to access EULAs.

These regressions show some determinants of the probability of EULA access consistent with search theory: this probability is higher when the benefits of access are higher, e.g., for relatively obscure companies less likely to care about their reputation and that may be felt to be

more likely to attempt to force bad terms on an unsuspecting consumer, or when the costs of finding EULAs and reading contract terms are likely to be lower, e.g., because consumers have a lower opportunity cost for their time or because they are more educated and thus find it easier to read contract terms. The impact of these factors, even collectively, is small, however. The main result of this section is that the most important term in the regression is the constant term: EULAs are rarely accessed, and thus rarely read, by anyone.

The small fraction of consumers accessing EULAs suggests a high total cost of finding the EULA and reading the terms. If the primary cost is in locating and accessing it, then mandating disclosure would reduce this cost, and thus increase the fraction of consumers becoming informed. If, on the other hand, the primary cost is in reading and assessing contract terms, mandating disclosure is unlikely to have a major impact on the fraction of consumers becoming informed.[78]

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 equilibria in the provision of those terms. In this paper we investigate the extent to which consumers actually do access the terms of certain standard form online contracts: we attempt to measure directly the "informed minority." Our clickstream data allow us to measure this aspect of consumer behavior with reasonable precision.

We find that very few consumers choose to become informed about standard form online contracts. In particular, we estimate the fraction of retail software shoppers that accesses EULAs at between 0.05% and 0.22%, and the very few shoppers that do access it do not, on average, spend enough time on it to have digested more than a fraction of its content. We also document

[78] In this case, knowing that reading EULA terms is likely to be prohibitively costly, consumers are likely not even to access these terms, even if they could do so with only the click of the mouse.

that shoppers rarely access other substitute information sources, such as consumer product review or relevant news sites, to learn about EULA terms. Even under generous assumptions, it is hard to envision the probability that EULAs are read, and understood, growing even to 1%. 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 simple theoretical calculations further indicate that it is too small to sustain an informed minority equilibrium.

While our results apply directly only to one standard form context, the fact that online comparison shopping is so cheap and easy raises questions about whether informed minority mechanisms could be active in more general contexts in which comparison shopping is harder and costlier. A potential implication of our results is that, given the uniform low readership, regulation mandating increase accessibility of contract terms may be ineffective. In work in progress we are investigating the extent to with the ease of access to the EULA, based on its specific location on the seller's website, shifts readership to a higher level.[79]

[79] Marotta-Wurgler & Bakos, supra note 5.

Figure 1. Empirical Framework

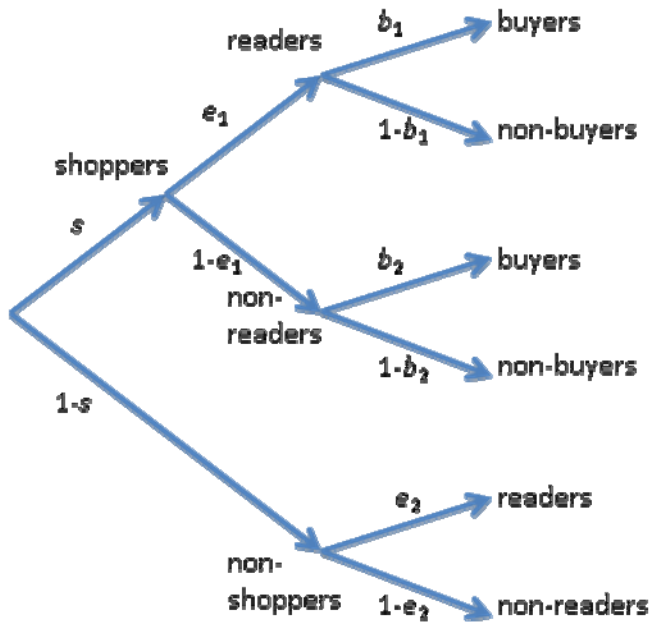
Visitor types	Fraction of visitors
	se_1b_1
	$se_1(1-b_1)$
	$s(1-e_1)b_2$
	$s(1-e_1)(1-b_2)$
	$(1-s)e_2$
	$(1-s)(1-e_2)$

Figure 2. Annual maintenance to product price ratio, forty companies

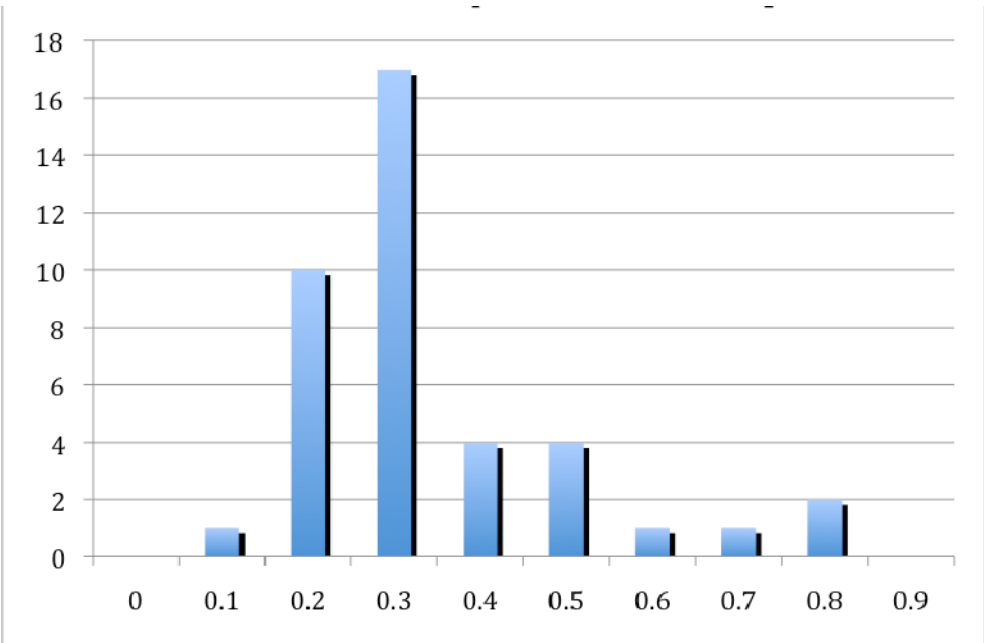


Table 1. Company and Product Characteristics

	N	Mean (s.d)	Minimum	Median	Maximum
Panel A. Company Characteristics					
Retail					
Revenue (Millions $)	56	2,140 (8,190)	0.1	13.8	51,100
Age (years)	56	17.30 (11.03)	3	15	56
Public Company	56	0.30 (0.46)	0	0	1
Freeware					
Revenue (Millions $)	10	1,390 (4,390)	0.1	0.1	13,900
Age (years)	10	10.20 (6.84)	5	7	27
Public Company	10	0.10 (0.32)	0	0	1
Panel B. Product Characteristics					
Retail					
Consumer Product	56	0.68 (0.47)	0	1	1
Price ($)	56	483.97 (1,203.9)	9.97	63.50	5,295
Median Price ($)	56	388.72 (1,032.53)	1	47	4,995
Trial (featured product)	56	0.86 (0.35)	0	1	1
Trial (most products)	56	0.80 (0.40)	0	1	1
Freeware					
Consumer Product	10	0.90 (0.47)	0	1	1

Table 2. User Characteristics

	N	Mean (s.d)	Minimum	Median	Maximum
Panel A. Users accessing at least 2 pages in at least one sample company					
Age (years)	45,091	46.23 (13.78)	18	46	99
Gender (1= Male)	45,091	0.50 (0.50)	0	0	1
Income ($)	45,091	60,487 (39,666)	12,500	37,500	150,000+
Household Size	45,091	2.78 (1.27)	1	3	5+
Presence of Children (1= Yes)	45,091	0.41 (0.49)	0	0	1
Panel B. Users accessing at least 5 pages in at least one sample company					
Age (years)	31,969	46.37 (13.70)	18	46	99
Gender (1= Male)	31,969	0.50 (0.50)	0	1	1
Income ($)	31,969	60,566 (39,719)	12,500	37,500	150,000+
Household Size	31,969	2.79 (1.27)	1	3	5+
Presence of Children (1= Yes)	31,969	0.41 (0.49)	0	0	1
Panel C. Users initiating checkout in at least one sample company					
Age (years)	1,653	47.29 (13.90)	18	47	99
Gender (1= Male)	1,653	0.51 (0.50)	0	1	1
Income ($)	1,653	64,027 (42,064)	12,500	75,000	150,000+
Household Size	1,653	2.77 (1.23)	1	3	5+
Presence of Children (1= Yes)	1,653	0.39 (0.49)	0	0	1

Table 3. Company and EULA visits. Visits measured as uninterrupted sessions.

	N of company visits	Mean N of pg. clicks per company visit (s.d.)	Median N of pg. clicks per company visit	Mean length of company visit in seconds (s.d.)	Median length of company visit in seconds	N of EULA visits (% of company visits)	Mean N of pg. viewed before EULA access (s.d)	Median N of pg. viewed before EULA access	Mean length of EULA access in seconds (s.d.)	Median length of EULA access in seconds
Panel A. At Least 2 Pages Accessed During Visit										
Retail	120,545	12.4 (26.9)	5	310.9 (713.6)	105	55 (0.05%)	12.15 (16.6)	7	47.7 (45.4)	29
Freeware	28,007	13.6 (36.9)	4	163.51 (618.8)	43	40 (0.14%)	7.45 (15.4)	3	99.6 (237.8)	29
Panel B. At Least 5 Pages Accessed During Visit										
Retail	67,769	19.9 (34)	10	439 (903.1)	185	50 (0.07%)	13.2 (17)	8	46.9 (47.3)	27
Freeware	13,520	25.2 (50.6)	12	239.2 (856)	67	30 (0.22%)	9.5 (17.4)	4	60.6 (104.5)	20.5
Panel C. At Least 1 Secure Checkout Page Accessed During Visit										
Retail	5,509	13.1 (30.7)	5	586.5 (1,939)	218	5 (0.11%)	20.4 (16.8)	20	84 (60.5)	60

Table 4. Company and EULA visits. Visits measures as monthly aggregates of uninterrupted sessions.

Company	N of company visits	Mean N of pg. clicks per company visit (s.d.)	Median N of pg. clicks per company visit	Mean length of company visit in seconds (s.d.)	Median length of company visit in seconds	N of EULA visits (% of company visits)	Mean N of pg. viewed before EULA access (s.d)	Median N of pg. viewed before EULA access	Mean length of EULA access in seconds (s.d.)	Median length of EULA access in seconds
Panel A. At Least 2 Pages Accessed During Visit										
Retail	63,272	23.67 (79.2)	7	592.4 (2,880)	161	53 (0.08%)	21.5 (43.04)	9	51.9 (47.5)	30
Freeware	11,010	35.8 (292.8)	4	415.9 (2,811)	75	42 (0.38%)	10.9 (22.1)	3.5	102.5 (234.9)	29
Panel B. At Least 5 Pages Accessed During Visit										
Retail	40,697	35.3 (96.9)	14	837 (2,562)	292	49 (0.12%)	23.1 (44.4)	10	50.16 (47.82)	29
Freeware	5,370	70.6 (416.5)	11	741.5 (3,993)	148	34 (0.63%)	13.2 (24.1)	4	104.9 (251.2)	25
Panel C. At Least 1 Secure Checkout Page Accessed During Visit										
Retail	3,534	34 (79.1)	10	1,419 (4,681.5)	450	6 (0.17%)	24.2 (23.3)	16	99 (60.1)	94

Table 5. Logit regressions of determinants of EULA visits.

Dependent Variable: EULA access						
	Uninterrupted sessions			Monthly aggregate sessions		
	At Least 2 Pages	At Least 5 Pages	At Least 1 Secure CP	At Least 2 Pages	At Least 5 Pages	At Least 1 Secure CP
	(1)	(2)	(3)	(4)	(5)	(6)
Freeware dummy	1.31*	0.99	-	2.34***	2.17**	-
	(0.71)	(0.84)		(0.62)	(0.70)	
Ln Med. Price	0.29**	0.29**	-1.38***	0.31***	0.30**	-1.34***
	(0.10)	(0.12)	(0.11)	(0.09)	(0.10)	(0.11)
Ln Revenue	-0.42***	-0.46***	-0.13*	-0.42***	-0.44***	-0.14*
	(0.03)	(0.04)	(0.08)	(0.03)	(0.04)	(0.07)
Public	1.98***	2.23***	5.84***	1.93***	2.17***	5.34***
	(0.47)	(0.55)	(0.51)	(0.41)	(0.47)	(0.43)
N Pgs. Viewed	0.05***	0.02	0.04***	0.06***	0.03**	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Gender	-0.34	-0.24	-0.06	-0.32	-0.33	-0.10
	(0.23)	(0.25)	(0.19)	(0.21)	(0.23)	(0.19)
Ln Income	0.15	0.21	0.29**	0.18	0.22	0.20
	(0.16)	(0.18)	(0.14)	(0.15)	(0.16)	(0.13)
Ln Age	0.55	0.33	0.70**	0.45	0.35	0.43
	(0.37)	(0.40)	(0.30)	(0.35)	(0.37)	(0.31)
N	148,552	81,289	5,509	74,282	46,067	3,534
R^2	0.10	0.09	0.63	0.12	0.12	0.64

Note: * = denotes significance at 0.1 level, ** at 0.05 and *** at 0.01. Standard errors are clustered by visitor.