1-2014

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

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

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

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

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

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

March 2013

A cornerstone of the law and economics approach to standard form contracts is the "informed minority" hypothesis: In competitive markets, a minority of term-conscious buyers is sufficient to discipline sellers from using unfavorable boilerplate terms. This argument is often invoked to limit intervention or regulate consumer transactions, but there has been little empirical investigation of its validity. We track the Internet browsing behavior of 48,154 monthly visitors to 90 online software companies to study the extent to which potential buyers access the associated standard form contract, the end user license agreement. We find that only one or two out of every thousand retail software shoppers choose to access the license agreement, and most of those that do access it spend too little time to have read more than a small portion of the text. The limiting factor in becoming informed seems to be the cost of reading and comprehending the contracts rather than the cost of finding and accessing the information. We conclude that 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 the cost of becoming informed and comparison-shopping online is low, the results may suggest limits to this mechanism in offline standard form contracting environments as well.

 (JEL Codes: D11; D12; K00; K12)

[*] Stern School of Business, New York University; [**] New York University School of Law; [***] Boalt Law School, University of California at Berkeley.

1.   Introduction

Standard form contracts, often called "fine print" or "boilerplate," are the most common type of economic contract. They apply to untold billions of commercial transactions per year. In a typical scenario, a buyer 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. Examples appear anywhere from safety disclaimers noted on the backs of sporting tickets, warranties packaged with consumer goods, privacy policies and terms of use on websites, and photocopying restrictions appearing in the front matter of this journal. Every reader of this article has entered into thousands of standard form contracts, sometimes unknowingly.

Academics, courts, and policymakers have long debated the degree to which standard form contracts should be enforced, and whether their content or disclosure should be regulated. All sides in this debate realize that in many circumstances, a majority of buyers 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.[1]

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 on this latter argument, which derives from some classic law and economics contributions. The Schwartz and Wilde (1979) articulation of the informed minority argument in this context is a specific 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. 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

---

[1] 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 (1981). See also Steven C. Salop (1976).

discriminate between reading and non-reading buyers, sellers will offer the terms preferred by all buyers. This sort of competitive markets logic has been used extensively to resist regulation.[2] 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 has been influential in the law and economics literature, there has been no systematic empirical analysis of its validity; while it is widely agreed that standard form contract readers are in the minority, the literature offers no evidence whether this minority remains large enough to plausibly enforce efficient terms, as assumed by some theorists and courts, or rather whether it is so miniscule as to be practically irrelevant. We present large-sample evidence on the extent to which buyers actually do read standard form contracts. We also identify some factors that affect the probability of readership, and consider whether this probability is limited primarily by search costs or by reading/comprehension costs.

Specifically, we examine the extent to which potential buyers of software read End User License Agreements (EULAs), which 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. We track 48,154 visitors to 90 software companies over a period of one month, recording their detailed browsing behavior. For each such user we observe the exact sequence of web pages (URLs) accessed in a particular visit and the time spent on each page. The data also include detailed demographic characteristics of each user, such as age, gender, income, and geographical location. Our main finding is that regardless of how strictly we define a "shopper" (or "potential buyer"), only about one or two in one thousand shoppers access a product's EULA for at least one second, yielding an informed minority of 0.2% that is orders of magnitude smaller than the required informed minority size in realistic market settings and theoretical examples suggested in the literature. Such a small number of contract readers would seem to cast doubt on the existence of an informed minority of a size sufficient to police against one-sided terms, at least in the context of software sold online. While we cannot be certain about standard form contract readership in other markets, the fact

---

[2] See, e.g., Douglas Baird (2006), Clayton Gillette (2005), and Beales, Craswell, & Salop (1981).

that the cost of becoming informed online is low relative to other markets suggest limits to the mechanism more generally.

To put our findings in perspective, we perform a simple calibration exercise. We estimate the fraction of informed shoppers needed to induce sellers to offer "good" terms in the software market. We estimate the marginal cost of providing one pro-buyer term, maintenance and support, and find that sellers would find it more cost-effective to lose all informed buyers (that is, 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 unconditionally, 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, suggesting that the lack of a significant informed minority is due to high search and reading costs of standard form contracts.

Our main contribution is the first large-sample study of whether consumers actually inform themselves of important rights and obligations before entering a standard form contract and to empirically test the "informed minority hypothesis." Our specific context is the license agreements for software purchased online. We find that the fraction of consumers that read standard form contracts is so small that we can fairly confidently rule out this hypothesis as a mechanism that shapes standard form contract terms. Our findings, however, call into question the effectiveness of policies to ameliorate potential market failures by requiring increased or mandatory disclosure.[3]

---

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

We stress, however, that the absence of the informed minority mechanism does not immediately imply that EULA terms will be inefficiently biased in favor of sellers. Some sellers, at least very prominent ones, could be disciplined by other mechanisms, for example a concern for their reputation if onerous terms might eventually be discovered.

Furthermore, we cannot conclude whether it would be rational or irrational for shoppers to read the contract; for instance, shoppers may assume that no matter what the EULA terms state, they will be protected by the courts, or by sellers' desire to avoid unhappy customers or reputational damage. Also, we cannot easily determine whether the quality of terms is reflected in the observed prices, due to the many first–order determinants of prices that we are unable to control for. Our tests do not suggest a large effect of terms on prices, but our tests are not powerful. What we are able to demonstrate is that the most widely applied argument for the efficiency of standard form contract terms does not seem compelling in this context. This means that future empirical work should focus on the potential validity of alternative mechanisms.

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

2.   Form Contracts and Competition for Informed Buyers: Background

Despite the transaction-cost reducing benefits associated with the use of form contracts, such as reduced drafting and negotiation costs, academics and policy makers have debated their fairness and the desirability of their enforcement. Concern for consumer welfare has resulted in numerous articles, laws, and initiatives to regulate these contracts. For example, in addition to existing contract law doctrines to protect buyers from abusive terms, such as unconscionability and unfair surprise, several state consumer laws prohibit the use of forum selection clauses and disclaimers of implied warranties in consumer contracts.[4] Federal laws such as the Truth in Lending Act and the Magnuson-Moss Warranty Act seek to decrease reading and search costs by

---

[4] See, e.g., Gatton v. T-Mobile USA, Inc., 152 Cal. App. 4th 571, 585 (Cal. App. 1st Dist. 2007); Fidelity & Deposit Co. v. Gainesville Iron Works, Inc., 125 Ga. App. 829 (Ga. Ct. App. 1972); Idaho Code § 29-110; N.C.G.S. § 22B-3; Mont. Code 36 § 18-1-403.

requiring standardized disclosure of mandated terms.[5] More recently, there has been heated debate whether online contracts such as terms of use, privacy policies, and software license agreements should be enforceable or subject to mandatory disclosure rules or contain mandatory provisions.[6]

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) 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 and find that even in the presence of many uninformed consumers, a market can yield a competitive equilibrium if enough informed consumers shop for the lowest price.[7]

Schwartz and Wilde (1979) extend this argument to our more specific 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, as the cost to the seller of losing a critical mass of informed consumers outweighs the benefits of offering inferior terms to the uninformed inframarginal consumers. They conclude that imperfect information alone is not sufficient to warrant market intervention. This conclusion would become the cornerstone of the law and economics view of standard form contracts.

---

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

[6] See, e.g., Mark Lemley (2006) (arguing that browsewraps should be enforced only in cases involving sophisticated commercial parties that are repeat players).

[7] Salop & Stiglitz (1977) note: "[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."

Schwartz and Wilde (1979) state that "[g]enerally, there are a significant number of informed consumers in any given market prior to legal intervention,"[8] although the evidence they cite to support this statement would not appear to be so general or conclusive.[9] 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.[10] Nevertheless, many scholars have relied repeatedly on the "informed minority" argument to support freedom of contract.[11]

This assumption that a critical number of readers may lead to more efficient seller terms has guided courts on whether to enforce consumer agreements involving standard terms. Courts rely on several contract law doctrines that stress, among other things, that consumers must be

---

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

[9] 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 (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.

[10] 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 (1995). See also, Oren Bar-Gill and Elizabeth Warren (2008).

[11] For example, Priest (1981) relies on the informed minority argument to address the concern that most buyers will not factor warranty terms in their purchase decisions. 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." The terms of standard form contracts can also be seen as a non-price add-on in the spirit of Ellison (2005). While it had been often argued that seller profits from add-ons will be competed away through lower prices, e.g. see Lal and Matutes (1994). Ellison shows that when informed and uninformed consumers are sufficiently different, obfuscated add-ons soften competition and transfer surplus to sellers. This could be the case, for instance, if price cuts disproportionally attract consumers likely to become informed about standard contract terms. As more consumers become informed about contract terms, the outcome would approach the non-obfuscated equilibrium; thus the informed consumers could be seen as supporting an informed minority type of mechanism to discipline sellers.

given a meaningful "opportunity to read" terms.[12] To many commentators, a meaningful opportunity to read preserves individual autonomy and reduces the cost of becoming informed.[13].

Courts currently disagree on how much opportunity must be given. For example, some courts allow enforcement of shrinkwraps and "pay now, terms later" contracts (where terms are only available post-purchase), whereas other courts do not.[14] In addition, courts 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.[15]

There have been recent proposals to increase consumer protection in mass market transactions, especially those that take place over the Internet, the context of this study, by seeking to facilitate contract access and increase the opportunity to read. 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.[16] The drafters posit that reducing the cost of accessing the contract might increase the number of informed consumers to a number sufficient to constitute an informed minority of buyers.

This view is not universal, as some have expressed doubt that markets for standard form terms work well by relying on the creation or the existence of an informed minority. Katz (1990) shows how failure to read can be rationalized based on the low probability that the contingencies specified in the contract will materialize as well as individual consumers' inability to alter the

---

[12] See Specht v. Netscape Communications Corp., 306 F.3d 17 (2nd Cir. 2002) (refusing to enforce an arbitration agreement made available via a browsewrap a few screens down because it didn't provide sufficient notice). But see Hubbert v. Dell Corp., 359 Ill.App.3d 976, 835 N.E.2d 113 (5th Dist. 2005) (enforcing terms in browsewrap contract because the hyperlink was conspicuous enough to provide sufficient notic); Ticketmaster Corp. v. Tickets.com, Inc. 2000 WL 1887522 (C.D Cal. 2000) (enforcing browsewrap agreement because the defendant had knowledge of the terms after repeated visits to the site)

[13] See Omri Ben-Shahar (2009) and references cited therein. See also Robert Hillman (2006).

[14] See, e.g., ProCD v. Zeidenberg 86 F.3d 1447 (7th Cir., 1996) (enforcing a shrinkwrap contract); Step-Saver Data Systems, Inc. v. Wyse Technology,939 F.2d 91 (3rd Cir. 1991) (refusing to enforce a shrinkwrap contract). For a detailed discussion, see Florencia Marotta-Wurgler (2009).

[15] See Ben-Shahar (2009) and Restatement (Second) of Contract § 211. Courts also rely on the doctrine of unconscionability to refuse enforcement of onerous terms. Uniform Commercial Code (U.C.C.) § 2-302.

[16] ALI Principles of the Law of Software Contracts [hereinafter *ALI Principles*] at 117.

terms anyway.[17] Eisenberg (1995), Goldberg (1997), and Ben-Shahar (2009) challenge the informed minority argument on its own terms, noting that consumers will find it irrational to engage in search and deliberation of fine print terms and that under reasonable conditions sellers will find it more profitable to take advantage of non-readers than catering to readers. Ben-Shahar 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.

Bar-Gill (2004) and Gabaix and Laibson (2006) posit that consumers' psychological biases and limitations might prevent them from reading or understanding terms, again making them susceptible to seller manipulation. Gabaix and Laibson develop a model where sellers have no incentive to help buyers become informed because of the curse of debiasing: teaching "myopic" consumers to become "sophisticates" does not pay off whether these consumers are own customers or customers of a competitor; thus "informational shrouding flourishes even in highly competitive markets, even in markets with costless advertising, and even when the shrouding generates allocational inefficiencies." Thus an informed minority mechanism never arises.

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. Hillman and Rachlinski (2002) explore the role of the informed minority argument in standard form contracting online. 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 consumers is likely to result in an under-production of knowledgeable buyers.[18]

---

[17] See also W. David Slawson (1975) (noting that "[f]or the very reason that these terms are imposed rather than agreed upon, they are almost universally unfair."); Todd D. Rakoff, (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] 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.

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.[19] 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.

2.2.   *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.

Hall (1997) surveyed 100 purchasers of HP inkjet printers and found that only 3% reported that they were informed about replacement cartridge ink costs at the time of purchase of the printer (and none of these 3 offered an actual estimate close to the correct cost, or reported being aware of ink costs for other manufacturers). While ink costs typically account for as much as 90% of the total cost of ownership for inkjet printers, consistent with Gabaix and Leibson's model other printer manufacturers do not attempt to inform these consumers and Hall found no indication of an informed minority able to discipline sellers.

In the legal literature, 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." Almost 60% 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

---

[19] See Gillette (2004); Lucian A. Bebchuk & Richard A. Posner (2006).

the likelihood that respondents will read standard terms in different scenarios. Sixty percent of respondents claim they skim or read parts of a standard form contract before entering a transaction. Similarly, 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 report non-reading, to find out the reasons for that failure to read. They find that about 80% of respondents claim not to read contracts and much of the remainder claims to skim them. These surveys are suggestive but somewhat limited in what they can demonstrate. They are based on self-reported behavior or hypothetical commercial scenarios, and the survey subjects are not representative, e.g. in some cases they include law students who will write boilerplate for a living.

Other studies show that standard form contract terms are less one-sided in favor of sellers than might be possible if buyers were completely uninformed, and thus are consistent with 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. Marotta-Wurgler (2007, 2008) analyzes the terms of 647 online EULAs and shows that while almost all of them are more restrictive than the relevant default rules, they do not all converge to the legal minimum. In a study of contracting practices by online retailers, Mann and Siebeneicher (2008) find that few sellers offer excessively one-sided contracts.

3.   Research Framework

In this paper we investigate whether an informed minority of buyers capable of disciplining the market actually exists by studying the browsing and shopping behavior of online consumers. In particular, we track the behavior of Internet visitors to 90 software companies and we examine the rate at which shoppers choose to become informed about the EULAs that govern the featured software.

Online software purchases provide an apt setting 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.[20] For instance, as end-users increasingly rely on software to perform a variety of routine tasks and critical functions, damages from software failure can be significant. Third, shopping for competing goods and the terms that govern them is cheap and easy online relative to most commercial settings. Search costs are low (Bakos 2001). 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, several recent debates on legal reform in standard form contracts focus on electronic contracts in general and software contracts in particular. Our study of the informed minority in online software markets places us at the center of these debates.

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. 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.[21]

---

[20] 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).

[21] 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.

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

We can break down readers into $se_1 b_1$ 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 that constitute the informed minority, by writing this fraction as

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

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. This dataset was made available to us by a major online research company, which has recruited a representative panel of U.S. households that have agreed to install on their computers a data collection plug-in that records the URL address of each 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 dataset of significant size.[22]

---

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

The panel of households was selected to be demographically and geographically balanced and representative of the population of U.S. households with Internet access.[23] 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 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

---

[23] 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).

wishing to download it. Examples include browser toolbars, plug-ins, and browsers.[24] 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.

Some EULAs were presented as browsewraps, i.e., they were posted as a hyperlink somewhere on the seller's website. We included all these companies because we can easily measure whether users voluntarily clicked on the EULA hyperlink.[25] A minority of companies presented their EULAs as clickwraps. This mode of presentation *requires* consumers to click on an "I agree" button acknowledging the EULA terms before they can purchase a product. There are two types of clickwrap sites. One type of clickwrap sites present the EULA terms via a hyperlink adjacent to the "I agree" button and thus require an additional click to access the

---

[24] 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.

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

EULA; in this case while all buyers are forced to acknowledge the EULAs, we can measure what fraction takes the extra step and actually clicks on the contract link, which is a necessary step in order to become informed about the terms. The other type of clickwrap sites present the terms in a scrollable text-box above the "I agree" button; we removed companies that use this type of clickwrap from our data set because we do not have a way to measure whether buyers read the terms. While scrolling through and reading EULA terms would likely increase the total time spent on the corresponding checkout pages, which we can observe, there are several other actions users are typically required to take on the same pages. As a result measures based on the total time spent on the checkout page where the terms are presented were too noisy to be useful for this type of companies. Finally, we removed from our data set companies that did not make their EULAs available online.

After excluding companies for which we do not have enough data or otherwise are inappropriate for our tests, we arrive at a final sample of 78 retail and 12 freeware companies. 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 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 being equal, consumers may feel less need to scrutinize the terms in EULAs from companies that are large or old because they assume that such companies are more trustworthy and fair. To test this hypothesis, we obtain information about each company's annual revenue, year of incorporation, and public or private status. These data were obtained from Hoovers.com, Yahoo! Finance, or via 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 $1.56 billion with a standard deviation of $6.98 billion, a number obviously driven by a few large firms. Median revenue for this category is $5.2 million. The mean age of these companies, measured by subtracting their year of incorporation from 2009, is 15.8 years (median is 14.5). Twenty seven

percent of the companies in this category are publicly traded. In contrast, the mean age of freeware firms is 9.08 years (median is 6.5), and only 8% of these companies are publicly traded. For each company we note the number of products it offers (counting each distinctly named product as a separate product), which allows us to calculate the average revenue per product. The average revenue per product is $99.3, with a standard deviation of $406 (median is $0.99).

We also collect several product characteristics and we record one "flagship" product. Many small and medium size companies market one main product, in which case we select that product as flagship. For larger companies, we select the product accounting for the largest fraction of sales or, when this information was not available, the product most prominently featured on the website, as these might be the products most consumers care about.[26] Of course, consumers might be less inclined to read the EULAs of the most popular (or flagship) products, as they are likely to have established stronger reputations. Still, the choice of flagship product should not affect our overall findings, as the readership and visit metrics we report are for all products, not just flagship products.

It is possible that users are more inclined to become informed about the EULA terms of 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.[27] We also note whether the product is oriented toward business users or the general public. Finally, we classify each product into one of 150 software product categories, e.g. antivirus or word processing, based on the characterizations of software products available at Amazon.com.

---

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

[27] 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.

Panel B reports summary statistics for the flagship products' characteristics. The average number of products per company is 15 (median is 6). The average product price for retail companies is $408 and the median is $65. For each company we compute the median price of all software products listed on the website, and the mean of those medians is $365; the median of the median is $49. A fraction of 68% of retail products and 83% of freeware appears targeted to consumers rather than businesses. Finally, 83% of retail sites offer a trial version of their featured product or of the product we selected as flagship, and 78% 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 collected *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 sample companies.

4.4.   Defining Shoppers and Shopping Visits

Among panelists in our data that visit a given company's website we need to define "shoppers," i.e., visitors with some potential to purchase, since a potentially large fraction of visitors may be browsing without any intent to purchase. This is more important for 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 visitors with unlikely intent to purchase by excluding 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

17

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

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 for the 90 companies in our sample the web page addresses that would be accessed only during the checkout and payment process and subsequently recognizing visits that access such pages.[28] While knowing that a user started a checkout or payment process provides no guarantee that the transaction was completed, it indicates a high likelihood that a transaction was at least contemplated. This definition of shopping visit is likely to be overly restrictive, as it excludes visits that do not result in the initiation of a checkout process.[29]

A possible concern with the third 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. Given that our sample consists of most or all major software vendors and several smaller ones, and that browsing behavior of our panelists is followed for a month, it is reasonable to assume that shoppers with true intent to purchase may visit several merchants, but in the end will purchase

---

[28] 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.

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

from *some* merchant (that 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 read EULAs) are equally likely to do so in any of the merchants they consider. In that case, as explained in Section 5.2, the behavior of the visitors that initiate a checkout session will be representative of the shoppers as a whole, and thus the likelihood to access EULAs by shoppers that initiate a checkout session then provides an appropriate estimate for the size of the informed minority.

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*

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, as summarized in Moe & Fader (2004). 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 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 a store on the web has a low cost, users are more inclined to visit a company's website several times over an extended period before finally deciding whether or what to purchase. Thus, users accrue

19

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) who 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 company visits with intent to purchase.[30]

4.6. Demographic and Geographic Data

We utilize personal information about our panelists to identify characteristics of shoppers and shopping households that affect their likelihood of becoming informed about standard terms. Our dataset includes the age and sex of the head of the household, household income, household size, and whether there are children present in the household. In Table 2 we report the corresponding summary statistics.

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 48,154 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 $60,502 with a standard deviation of $39,704. 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 33,655 unique household user-visitors. The characteristics of these users are similar to those above. Panel C reports summary statistics for unique visitors that have selected a product for purchase and have begun the checkout process. Here there are 2,831 unique user-visits. As online conversion rates are typically low (in the order

---

[30] 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.

of a few percentage points across all goods), this set of users is significantly smaller compared to all shopping visits, but the users in this subsample are similar to the rest.

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*

We want 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 131,729 such visits to software retailers and 28,663 to freeware providers, including repeat visitors. For retail companies, an average visit consisted of

12.1 page views over 303 seconds (5.05 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 101 seconds (1.68 minutes).

The data indicate that EULAs were accessed in only 63 of the 131,729 visits to software retailers (0.05% of all such visits) and in 44 visits to freeware companies (0.15%). Users that accessed EULAs in retail companies visited an average of 19.1 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 59.4 seconds and the median time was 34 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.) Forty-six percent of these accesses were less than 30 seconds, and 92% were less than 2 minutes. The average number of words of EULAs for retail products in the sample (unreported) is 2,277 with a median of 2,187 words and a standard deviation of 1,148 words. The time spent on the EULAs relative to their length indicates that most readers did not read terms in their entirety, especially as they are generally written in complex legalese and since consumers are unlikely to be aware of the default rules, even if EULAs do spell out some terms in clear language, there is still a likelihood for misunderstanding. Bailey & Bailey (1999) find that 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 to 10 minutes, rather than under one minute. 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. On the other hand, the small number of people that read EULAs may not be representative of the average reader and may have developed the ability to quickly "skim" the essential information.

Visits to freeware providers have a mean of 13.4 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 0.15% of these visits. The median time spent on EULAs is a similar 33.5 seconds with 50% under 30 seconds and 86% under 2 minutes, against a median length of these EULAs of 1,754 words, so an overcount of effective readers is likely here as well.

When a visit is defined to require five or more pages accessed at the company visited, there are 72,282 uninterrupted session/visits to software retailers and 13,715 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 183 seconds (3.05 minutes). Distributions of page views and duration are again skewed. EULAs were accessed at a slightly higher rate in these visits, 57 times among software retailers (0.08%) 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. Times spent on EULAs are similar as before, with about half the accesses under 30 seconds and 90% under 2 minutes.

Finally, limiting our consideration to visits to software retailers that included initiation of a secure checkout session, the number of visits falls to 4,866, with 5 median page views per visit, but longer mean and median durations. This is expected since purchases require more time to process the transaction. In this restricted sample, there are 7 voluntary accesses of a EULA in the course of purchase, constituting 0.14% of all visits. All accesses are at least 30 seconds, and the median time spent in the EULA almost doubles for users in this group to 60 seconds. Interestingly, out of *all* sessions with EULA visits, 3.7% (if we use the two page visit definition) or 6.7% (if we use the five page visit definition) resulted in initiating a checkout session. If all of the initiated checkout sessions were completed leading to a purchase, the resulting conversions would be significantly higher than the typical 2% conversion rate in Internet purchases.

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

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). 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 0.65%, meaning that there are about six 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 $se_1 b_1$ readers that buy and $se_1(1-b_1)$ readers that do not buy are 9 and 47, respectively. There are $s(1-e_1)b_2$ or 2,982 buyers that are not readers, and $s(1-e_1)(1-b_2)$ or 40,670 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 5 (out of 25,664 visits). Finally, the large majority of non-shoppers do not read EULAs; this number is $(1-s)(1-e_2)$ or 25,661 based on the above proxies. We thus arrive at an estimate for e_1, the fraction of shoppers that constitutes

the informed minority of $\dfrac{se_1 b_1 + se_1(1-b_1)}{se_1 b_1 + se_1(1-b_1) + s(1-e_1)b_2 + s(1-e_1)(1-b_2)} = \dfrac{56}{43,708} = 0.13\%.$

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 9/56 = 16.1% (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. 9/2991 = 0.30%.[31,32]

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.[33]

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

---

[31] 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 56/(2991+47) = 1.84% of the total number of shoppers.

[32] 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 (0.30%) provides an estimate for the size of the informed minority among these most determined shoppers.

[33] 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 7 readers that buy, 50 readers that do not buy, 4,859 buyers that are not readers, and 67,366 shoppers that neither read nor buy, and 59,447 non-shoppers that include only 6 readers. The fraction e_1 of shoppers in the informed minority would be 57/72282 = 0.079%. The fraction of readers that initiate checkout sessions would be 7/57 = 12.3%, and assuming the same conversion ratio for non-readers would give 7/4866 = 0.14% as the informed minority. A conversion ratio of 100% for non-readers would give an upper bound for the informed minority of 57/(4866+50) = 1.16%.

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 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:

$$r \geq \frac{c_g - c_b}{\left(1 - \dfrac{b_1}{b_2}\right)(c_g - c_b) + \dfrac{b_1 - b_3}{b_2}(p - c_b)}.\ ^{34}$$

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

---

[34] 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.

marginal cost of "good" terms, and put our observed fraction of readers into perspective. M&S is a key term in software EULAs,[35] and thus the cost of supplying M&S should be an order-of-magnitude approximation of the cost of offering good EULA terms.[36]

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[37] 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.

---

[35] 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.

[36] 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.

[37] 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.

All of the above factors would result in a M&S-to-product price ratio that is higher than the cost of providing M&S. On the other hand, M&S is only one of 23 key EULA terms, which include several other types of warranties and permissions to copy or distribute the software that can impose opportunity costs. Furthermore, M&S pricing is similar in enterprise software markets, where significant competition exists from third-party M&S providers and purchase of M&S contracts is almost universal.

On balance, it is reasonable to assume that the cost of this level of M&S is around 20% of the product price, or $0.2p$ 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

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

If $(b_1 - b_3)/b_2 \leq 1$, which would be the case if readers purchase with at least the same probability as nonreaders when they discover good terms 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 $r \geq 0.2$. In other words, if the cost of providing good terms is $0.2p$ and the three reasonable assumptions in the previous sentence hold, as is the case in our data, then one would require $r \geq 0.2$ in order to support an informed minority equilibrium. 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.[38] In either case, our data are unlikely to be consistent with an informed minority equilibrium.

A similar and perhaps simpler approach to whether we might be observing an informed minority equilibrium 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

---

[38] 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.

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 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. 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. Thus the observed consumer behavior could represent a reputational equilibrium as analyzed in Shapiro (1982). In that type of equilibrium reputation alone will not ensure an efficient provision of terms because of lags in reputational adjustment, however reputational learning could in thoery have similar effects as the informed minority mechanism in disciplining sellers to increase the quality of provided 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 reviewed 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

131,729 sessions with at least two pages accessed, only three shoppers accessed pages with EULA information in consumer review sites.[39] At the monthly level of aggregation, 11,657 (16.8%) of visits 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.[40] A total of 84 shoppers 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 behavior suggests that it is highly unlikely that shoppers are 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 for the possibility that we are observing a reputational equilibrium.[41]

It is also possible that the quality of terms is reflected in the seller's price, and thus shoppers do not need to become informed about these terms. As is the tenet of the informed minority hypothesis, however, for terms to be reflected in the observed prices, a sufficient number of consumers need to be informed about terms when they comparison shop (unless all sellers offer the worst enforceable terms and consumers know this and therefore have no incentive to become informed about terms). Marotta-Wurgler (2007) finds that most sellers offer worse terms than those provided by default rules, but not so bad that no information could be gained by becoming informed about these terms. Our results are consistent with this,, although we are not able to draw any conclusions about the economic efficiency of this outcome.

---

[39] 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.

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

[41] There are still other ways for consumers to become "informed" without reading, such as word of mouth or through the mass media. For instance, a particularly oppressive term might gain notoriety as 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 outlets and ultimately Facebook rescinded some of the objectionable changes. This reputational mechanism deserves further study in that it can work with very few readers and it may be able to deter sellers from providing inefficiently unfavorable terms, however it is distinct from the informed minority mechanism that is the focus of this paper. Furthermore it is likely to be less relevant for less-known software sellers whose standard form terms would be of little interest to the popular media.

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

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

Another possibility is that the degree of disclosure of the EULA terms (e.g., the prominence with which they are displayed) signals their quality. For instance, an equilibrium may arise where good terms are prominently displayed while poor terms are made harder to find or inaccessible. Our results show that such an equilibrium, if it exists in our setting, is not maintained based on an informed minority. Furthermore, Marotta-Wurgler (2009) found that prominently disclosed clickwrap contracts had roughly the same one-sidedness as non-clickwrap contracts that required more effort to be discovered. This suggests firms are not using the form of disclosure to signal the quality of terms. They do not seem to use clickwraps to increase awareness of good terms and similarly they do not seem to be using browsewarps to hide bad terms.

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 shown 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 not significant indicating that the lack of propensity to read is present irrespective of the product price.

Consumers may be less likely to access EULAs for products they know about and thus may trust, and this trust may be proxied for by average product revenues as they are likely to indicate a higher degree of familiarity among potential buyers. We thus consider as a regressor the natural log of a company's revenues (as a proxy for size) divided by the number of products; coefficient for product revenue is significantly negative as would be expected, supporting the "product 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, we explore the role of demographic characteristics (income, age and gender), as, for instance, 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 may have a higher opportunity cost, and thus 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), but these variables mostly lacked statistical significance significant.

In unreported regressions we examined 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. For instance such users may have concluded that they will not need support from the vendor, or may have already researched the availability and terms of such support. We find that 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. We also study whether shoppers are less likely to read the EULAs of products that are more likely to be used over long periods of time and involve repeat purchases (e.g., because of upgrades). Users that become familiar with a product that is continuously updated, like Microsoft Office, may feel less need to concern themselves with the EULA as they may be familiar with the terms from previous use. Other products, such as test preparation software, are less likely to be purchased repeatedly. On the other hand, products characterized by long-term use may be more important to their users, and may thus induce them to read the EULA terms. We created a dummy variable that captures whether the company markets products that in our judgment are likely to be repeat purchases. The relationship between this variable and users' propensity to access EULAs is generally positive and under certain definitions of shoppers it is statistically significant.[42]

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.

5.5.   Why do so few consumers read?

---

[42] Since the low probability of EULA access could bias the parameter estimates obtained by logit regressions, we attempted to correct these estimates using the relogit methodology described in Gary King & Langche Zeng, Logistic Regression in Rare Events Data, 9 Pol. Analysis 137 (2001). This provided corrected parameter estimates for models (1), (2), (4) and (5) in Table 5, but the corrected values were essentially identical to the ones in the Table. Furthermore, to see whether systematic differences in companies that had relatively few visits in our sample biased the predicted probability of EULA access, we also estimated logit regressions with frequency weights; these uniformly resulted in lower predicted probabilities, showing that any frequency bias in our results will result in overestimating the probability of EULA access.

The small fraction of consumers accessing EULAs suggests a high total cost of finding the EULA and reading the terms, resulting in too few consumers reading to support an equilibrium consistent with the informed minority hypothesis. If the primary cost is in locating and accessing the standard form contract, then mandating disclosure would reduce this cost, and thus increase the fraction of consumers becoming informed. However, Marotta-Wurgler (2012) analyzes this same data set and finds that increased contract disclosure is not associated with increased readership. This suggests that the primary cost facing consumers is in reading and assessing contract terms; in this case, knowing that reading EULA terms is likely to be prohibitively costly, it is possible that consumers decide not even to access these terms, even if they could do so with only the click of the mouse, which would be an interpretation consistent with our results. An important implication is that mandating disclosure is unlikely to have a major impact on the fraction of consumers becoming informed, and thus it would be unlikely to support an informed minority equilibrium. If it is desirable to promote this type of equilibrium, or any equilibrium where a significant fraction of consumers become informed, measures that reduce the cost of reading and assessing the contract terms are likely to be more successful. Thus a regulatory approach focusing on shortening and simplifying online contracts, standardizing their terms, and providing a standardized summary is more likely to increase readership than an approach focusing on mandating disclosure. This suggests that an important benefit of regulations that mandate disclosure of basic credit terms in a standardized manner and large fonts, such as the "Schumer box" in the U.S. and the "Summary Box" in the UK, comes from reducing the cost of reading and comprehending these terms. Simple and plain language requirements can be seen in the same light.

6.   Discussion and Implications

Consumer access to the terms of standard form contracts has been at the center of a legal and policy debate, and a major question has been whether disclosure of terms in standard form contracts that govern consumer transactions should be regulated. A related debate has focused on the enforceability of terms and possible need to regulate disclosure for software in general, and software purchased online in particular. A central issue in these debates is the validity of the informed minority hypothesis: the view that comparison shoppers for standard terms help sustain

efficient 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 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 consequential in more general contexts where comparison shopping is harder and costlier. The low readership of contract terms, even for those EULAs where consumers are required to acknowledge reading and agreeing to these terms before purchasing a product and the terms are prominently accessible, suggests that it is the cost of reading and assessing these terms, rather than the cost of finding them, that discourages consumers. This is confirmed by related work showing that ease of access to the EULA, based on its specific location on the seller's website, does not meaningfully increase readership. A potential implication of these findings is that regulation aspiring to promote the emergence of an informed fraction of consumers that can subsequently affect the terms offered by sellers may be too ambitious. In particular, mandating increased disclosure of contract terms in this setting is likely to be ineffective; if the goal is at all achievable, it will require making contract terms easier to read, assess and compare.

References

Bailey, R.W. and Bailey, L.M., 1999. "Reading speeds using RSVP." *User Interface Update*.

Douglas Baird, 2006, The Boilerplate Puzzle, 104 Michigan Law Review 933, 933.

Bakos, Yannis, 2001. "The emerging Landscape for Retail E-Commerce." *Journal of Economic Perspectives* 15(1): 69-80.

Bar-Gill, Oren, 2004. "Seduction by Plastic." *Northwestern University Law Review* 98: 1373.

Oren Bar-Gill and Elizabeth Warren, 2008. Making Credit Safer, 157 U. of Pa. L. Rev. 1.

Beales, Howard, Richard Craswell and Steven C. Salop, 1981. "The Efficient Regulation of Consumer Information." *Journal of Law & Economics* 21(3): 491-539.

Bartlett, Robert P. III and Victoria C. Plaut, 2011. Blind Consent? A Social Psychological Investigation of the Non-Readership of Contracts, *Law and Human Behavior*: 1-23.

Lucian A. Bebchuk & Richard A. Posner, 2003. One-Sided Contracts in Competitive Consumer Markets, 104 Mich. L. Rev. 827.

Becher, Shmuel I. and Esther Unger-Aviram, 2009. "The Law of Standard Form Contracts: Misguided Intuitions and Suggestions for Reconstruction." *DePaul Business & Commercial Law Journal*.

Ben-Shahar, Omri, 2009. "The Myth of the "Opportunity to Read in Contract Law." *European Review of Contract Law* 5(1).

Omri Ben-Shahar and Carl Schneider, 2011. The Failure of Mandated Disclosures, 159 University of Pennsylvania Law Review 3.

Bucklin, Randolph E. and Catarina Sismeiro, 2003. "A Model of Web Site Browsing Behavior Estimated on Clickstream Data." *Journal of Marketing Research* 40: 249-267.

Catledge, Lara D. and James E. Pitkow, 1995. "Characterizing Browsing Behaviors on the World Wide Web." *Computer Networks and ISDN Systems* 27 (6): 1065–1067.

R. Ted Cruz and Jeffrey Hinck, Not My Brother's Keeper: The Inability of an Informed Minority to Correct for Imperfect Information, 1995. 47 Hastings L. J. 636.

Eisenberg, Melvin A., 1995. "The Limits of Cognition and the Limits of Contract." *Stanford Law Review* 47 (211): 240-41.

Ellison, Glenn, 2005. "A Model of Add-On Pricing." *Quarterly Journal of Economics*, 120 (2): 585-637.

Gabaix, Xavier and David Laibson, 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics* 121 (2): 505.

Clayton Gillette, 2005. Pre-Approved Contracts for Internet Commerce, 42 Houston Law Review, 975.

Clayton Gillette, Rolling Contracts as an Agency Problem, 2004. Wis. L. Rev. 679, 690 .

Goldberg, Victor P., 1997. "The 'Battle of the Forms': Fairness, Efficiency, and the Best-Shot Rule. *Oregon Law Review* 76 (155): 165.

Hall, Robert, 1997. The Inkjet Aftermarket: An Economic Analysis. Working Paper, Stanford University.

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

Hillman, Robert, 2006. "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.)

Hillman, Robert, 2006. "On-Line Boilerplate: Would Mandatory Website Disclosure of E-Standard Terms Backfire?" *Michigan Law Review* 104: 837-856.

Johnson, Eric J., Wendy W. Moe, Peter S. Fader, Steven Bellman, and Gerald L. Lohse, 2004. "On the Depth and Dynamics of Online Search Behavior." *Management Science* 50 (3): 299-308.

Katz, Avery, 1990. "Your Terms or Mine? The Duty to Read Fine Print in Contracts." *RAND Journal of Economics* 21: 518

Russell Korobkin, 2003. "Bounded Rationality, Standard Form Contracts, and Unconscionability," 70 U. Chi. L.Rev. 1203.

Lal, Rajiv and Carmen Matutes, 1994. "Retail Pricing and Advertising Strategies." *Journal of Business* 67: 345-370.

Mark Lemley, 2006. "Terms of Use," 91 Minn. L. Rev. 459.

Lewis Mandell, 1971. Consumer Perception of Incurred Interest Rates: An Empirical Test of the Efficacy of the Truth in Lending Law, 26 J. of Finance 1143, 1153.

Marotta-Wurgler, Florencia, 2007. "What's in a Standard Form Contract? An Empirical Analysis of Software License Agreements." *Journal of Empirical Legal Studies* 4: 677.

Marotta-Wurgler, Florencia, 2008. "Competition and the Quality of Standard Form Contracts: The Case of Software License Agreements." *Journal of Empirical Legal Studies* 5(3): 447.

Florencia Marotta-Wurgler, 2009. Are "Pay Now, Terms Later" Contracts Worse for Buyers? Evidence From Software License Agreements, 38(2) J. L. Stud. 309..

Marotta-Wurgler, Florencia, 2012. Does Contract Disclosure Matter? 168 *Journal of Institutional and Theoretical Economics* 120..

Moe, Wendy W. and Peter S. Fader, 2004. "Dynamic Conversion Behavior at e-Commerce Sites." *Management Science* 50.

Mann, Ronald and Travis Siebeneicher, 2008. "Just One Click: The Reality of Internet Contracting." *Columbia Law Review* 108: 984.

Priest, George, 1981. "A Theory of the Consumer Product Warranty." *Yale Law Journal* 90(6): 1347.

Todd D. Rakoff, 1983. Contracts of Adhesion: An Essay of Reconstruction, 96 Harv. L. Rev. 1173, 1226.

Salop, Steven, 1976. "Information and Monopolistic Competition." *American Economic Review* 66 (2): 240-245.

Salop, Steven and Joseph Stiglitz, 1977. "Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion." *Review of Economic Studies* 44 (3): 493-510.

Schwartz, Alan and Louis Wilde, 1979. "Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis." University of Pennsylvania Law Review 127: 630.

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

Shapiro, Carl, 1982. "Consumer Information, Product Quality, and Seller Reputation." *The Bell Journal of Economics* 13 (1).

Spence, Michael, 1977. "Consumer Misperceptions, Product Failure and Producer Liability." *Review of Economic Studies* 44 (3): 561-572.

Table 1. Company and Product Characteristics

	N	Mean (s.d)	Minimum	Median	Maximum
Panel A. Company Characteristics					
Retail					
Revenue (Millions $)	78	1,560 (6,980)	0.1	5.2	51,100
Average revenue per product (Millions $)	78	99.3 (406)	0.004	0.99	2,550
Age (years)	78	15.84 (10.10)	3	14.5	56
Public Company	78	0.27 (0.45)	0	0	1
Freeware					
Revenue (Millions $)	12	1,160 (4,010)	0.1	0.1	13,900
Number of products offered	12	9.25 (28.58)	1	1	100
Average revenue per product (Millions $)	12	12.5 (39.8)	0.1	0.1	139
Age (years)	12	9.08 (6.72)	3	6.5	27
Public Company	12	0.08 (0.29)	0	0	1
Panel B. Product Characteristics					
Retail					
Consumer Product	78	0.68 (0.47)	0	1	1
Number of products offered	78	14.78 (25.02)	1	6	150
Price ($)	78	407.95 (1,042.92)	9.97	65.12	5,295
Median Price ($)	78	364.84 (1,032.23)	1	49	5,000
Trial (featured product)	78	0.83 (0.38)	0	1	1
Trial (most products)	78	0. 78 (0.42)	0	1	1
Freeware					
Consumer Product	12	0.83 (0.39)	0	1	1

Table 2. User Characteristics

Panel A. Users accessing at least 2 pages in at least one sample company

	N	Mean (s.d)	Minimum	Median	Maximum
Age (years)	48,154	46.22 (13.78)	18	46	99
Gender (1= Male)	48,154	0.50 (0.50)	0	0	1
Income ($)	48,154	60,502 (39,704)	12,500	37,500	150,000+
Household Size	48,154	2.78 (1.27)	1	3	5+
Presence of Children (1= Yes)	48,154	0.41 (0.49)	0	0	1

Panel B. Users accessing at least 5 pages in at least one sample company

	N	Mean (s.d)	Minimum	Median	Maximum
Age (years)	33,655	46.37 (13.70)	18	46	99
Gender (1= Male)	33,655	0.50 (0.50)	0	1	1
Income ($)	33,655	60,612 (39,782)	12,500	37,500	150,000+
Household Size	33,655	2.79 (1.27)	1	3	5+
Presence of Children (1= Yes)	33,655	0.41 (0.49)	0	0	1

Panel C. Users initiating checkout in at least one sample company

	N	Mean (s.d)	Minimum	Median	Maximum
Age (years)	2,831	47.39 (14.03)	18	47	99
Gender (1= Male)	2,831	0.52 (0.50)	0	1	1
Income ($)	2,831	63,008 (41,373)	12,500	75,000	150,000+
Household Size	2,831	2.77 (1.23)	1	3	5+
Presence of Children (1= Yes)	2,831	0.40 (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; Median length of EULA access in seconds (s.d.)	Percent of EULA accesses lasting less than 30 seconds; 2 minutes
Panel A. At Least 2 Pages Accessed During Visit										
Retail	131,729	12.1 (26.7)	5	303.4 (698.4)	101	63 (0.05%)	19.1 (62.3)	7	59.4; 34 (91.7)	46%; 92%
Freeware	28,663	13.4 (36.5)	4	164.3 (616.3)	43	44 (0.15%)	6.9 (14.8)	2	99.1; 33.5 (227.3)	50%; 86%
Panel B. At Least 5 Pages Accessed During Visit										
Retail	72,282	19.8 (34)	10	436.8 (891.6)	183	57 (0.08%)	20.9 (65.3)	8	59.6; 30 (96.3)	49%; 91%
Freeware	13,715	25 (50.3)	12	241 (855.2)	68	30 (0.22%)	9.5 (17.4)	4	60.6; 20.5 (104.5)	57%; 93%
Panel C. At Least 1 Secure Checkout Page Accessed During Visit										
Retail	4,866	13.1 (30.7)	5	568.9 (1,861.8)	207.5	7 (0.14%)	15.7 (15.9)	7	166.3; 60 (234.6)	0%; 71%

41

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; Median length of EULA access in seconds (s.d.)	Percent of EULA accesses lasting less than 30 seconds; 2 minutes
Panel A. At Least 2 Pages Accessed During Visit										
Retail	69,372	23 (77.6)	7	576.1 (2,278)	152	61 (0.09%)	30.1 (73.8)	10	63.4; 42 (93.2)	44%; 90%
Freeware	11,323	35.1 (288.8)	4	416 (2,781.7)	76	46 (0.41%)	10.2 (21.2)	3	101.7; 33.5 (225)	49%; 85%
Panel B. At Least 5 Pages Accessed During Visit										
Retail	43,708	34.9 (95.8)	14	829.2 (3,470.8)	286	56 (0.13%)	32.7 (76.6)	11	62.7; 32 (96.8)	46%; 91%
Freeware	5,509	69.4 (411.4)	11	741.3 (3,956.7)	150	36 (0.65%)	12.6 (23.5)	4	107.8; 33.5 (244.2)	49%; 84%
Panel C. At Least 1 Secure Checkout Page Accessed During Visit										
Retail	2,991	37.1 (88.6)	11	1,456.8 (4,657.5)	457	9 (0.30%)	20.7 (20.9)	12	160.4; 106 (203.2)	0%; 67%

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	1.06	0.71	-	2.05***	1.95**	-
	(0.75)	(0.81)		(0.75)	(0.79)	
Ln Med. Price	-0.004	-0.02	0.38	0.07	0.06	0.07
	(0.13)	(0.14)	(0.39)	(0.14)	(0.14)	(0.41)
Ln Revenue per product	-0.52***	-0.55***	-0.32	-0.52***	-0.54***	-0.36**
	(0.05)	(0.06)	(0.21)	(0.05)	(0.06)	(0.14)
Public	2.28***	2.57***	0.29	1.85***	2.13***	1. 08
	(0.35)	(0.43)	(1.16)	(0.35)	(0.39)	(0.89)
N Pgs. Viewed	0.06***	0.03**	0.13***	0.07***	0.04***	0.12***
	(0.01)	(0.01)	(0.04)	(0.01)	(0.01)	(0.03)
Gender	-0.39*	-0.30	-1.23	-0.35*	-0.36*	-0.48
	(0.22)	(0.24)	(0.94)	(0.20)	(0.22)	(0.71)
Ln Income	0.11	0.21	-0.23	0.13	0.21	-0.21
	(0.15)	(0.17)	(0.43)	(0.14)	(0.15)	(0.42)
Ln Age	0.40	0.14	1.45	0.30	0.19	-0.49
	(0.35)	(0.39)	(1.31)	(0.33)	(0.36)	(1.38)
N	160,392	85,997	4,866	80,695	49,217	2,991
Pseudo-R^2	0.11	0.11	0.20	0.13	0.13	0.15

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

Figure 1. Empirical Framework

Visitor types	Fraction of visitors

readers $\xrightarrow{b_1}$ buyers $\qquad se_1b_1$

e_1

shoppers $\xrightarrow{1-b_1}$ non-buyers $\qquad se_1(1-b_1)$

s

$1-e_1$ non-readers $\xrightarrow{b_2}$ buyers $\qquad s(1-e_1)b_2$

$\xrightarrow{1-b_2}$ non-buyers $\qquad s(1-e_1)(1-b_2)$

$1-s$

$\xrightarrow{e_2}$ readers $\qquad (1-s)e_2$

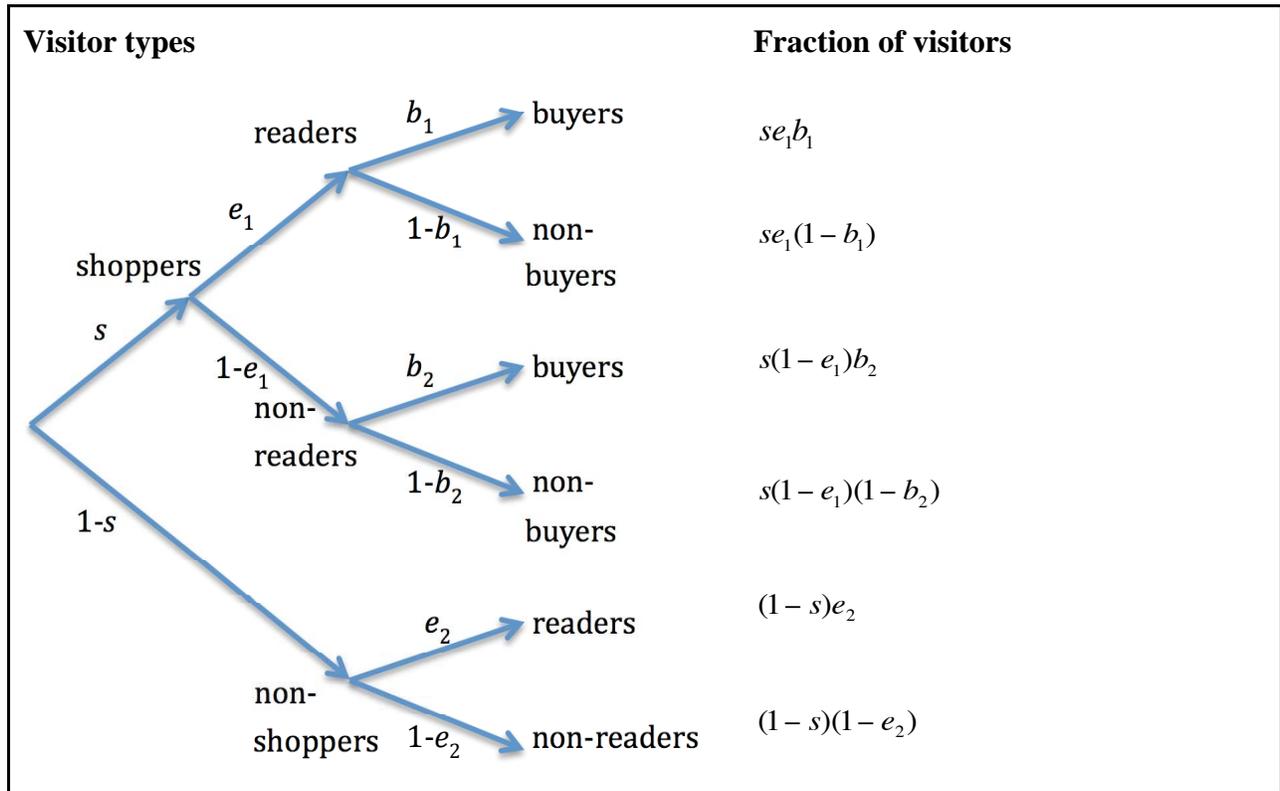non-shoppers $\xrightarrow{1-e_2}$ non-readers $\qquad (1-s)(1-e_2)$

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