LEARNING IN STANDARD FORM CONTRACTS: THEORY AND EVIDENCE

Giuseppe Dari-Mattiacci and Florencia Marotta-Wurgler[1]

ABSTRACT

We explore learning and change in standard form contracts. We hypothesize that drafters (sellers) are more likely to revise the terms they offer when they have an opportunity to learn about their value. These opportunities arise only for those types of terms that allow drafters to experience the relative costs and benefits of offering them. Consider a warranty. Sellers offering a warranty in an initial period will be exposed to claims about malfunction by purchasers and will learn whether it is desirable to offer it going forward. When drafters are unable to learn, either because they fail to offer such learning-enabling terms initially, or because the term in question is one where there is no increased opportunity to learn, we expect that such terms will be revised less frequently. Indeed, a reduced opportunity to learn might create contractual "black holes," where terms that are less likely to be revised and might lose their meaning over time or appear less related to the rest of the contract. Our results support this hypothesis. Using a large sample of changes in consumer standard form contracts over a period of seven years, we find that sellers are more likely to revise terms that offer an opportunity to learn than those that do not. The results suggest that standard form contract terms evolve over time as sellers learn about their benefits, costs, and risks. Our results have normative implications for the design of default rules.

JEL classification: K12.
Keywords: standard form contract, boilerplate, evolution of contracts, learning.

[1] Giuseppe Dari-Mattiacci: University of Amsterdam School of Law and Amsterdam Business School. Florencia Marotta-Wurgler: New York University School of Law. The authors would like to thank Lisa Bernstein, Steven Choi, Kevin Davis, Ron Gilson, Gerard Hertig, David Hoffman, Henrik Lando, Mark Ramseyer, Alan Schwartz, Robert Scott, Jon Zytnick, and the participants in the Conference on Contractual Black Holes (2017) at Duke Law School, the annual meeting of the European Association of Law and Economics (2017) in London, the Conference of Empirical Legal Studies (2017) at Cornell Law School, for helpful comments and suggestions.

TABLE OF CONTENTS

INTRODUCTION 4

I. LEARNING, STICKINESS, AND INNOVATION IN STANDARD FORM CONTRACTS 6

II. A THEORY OF THE EVOLUTION OF STANDARD FORM CONTRACTS 9
A. Framework of analysis 9
B. Experiential learning 12
C. Scope of the theory 14
1. Standard form contracts versus tailored contracts 14
2. Alternative learning channels 15

III. SYMMETRIC VERSUS ASYMMETRIC LEARNING 16
A. Symmetric-learning terms 16
1. Nonlearning terms 16
2. Learning terms 18
B. Asymmetric-learning terms 21
1. Learning-from-default terms 21
2. Learning-from-opt-out terms 25
C. Other determinants of contractual choice 26
1. Opting-out costs 26
2. Switching costs 27
3. Growth prospects 27

IV. EMPIRICAL ANALYSIS 28
A. Empirical implications of the theory 28
B. Data and methodology 30
1. Summary statistics 30
2. Classifying symmetric and asymmetric learning terms 31
C. Analysis 33

V. CONCLUSIONS 37
A. Does experiential learning produce biases pro or against consumers? 38
B. The normative implications of experiential learning 39
C. Open questions for further research 39

VI. APPENDIX 39
A. Theory 39
1. Model setup 39
2. Information and learning 41
3. Determinants of contractual choice 43
4. Nonlearning terms 43
5. Learning terms 44
6. Learning-from-default terms 45
7. Learning-from-opt-out terms 47
8. Summary table for the comparative statics results 48
9. Switching rates 48
B. Tables 50

C. Empirics 60

INTRODUCTION

One of the defining characteristics of standard form contracts is a high degree of standardization. Consumer products tend to be sold with limited warranties, disclaimers of implied warranties, limitations of damages, and dispute resolution clauses, among other terms.[2] Another characteristic of standard form contracts is that their terms tend to be "sticky." In theory, contracting parties should revise their agreements when doing so enhances the value of their transaction. However, the literature has identified a number of factors that might reduce contracting parties' incentives to deviate from the norm or default rules, even when alternative arrangements enhance the value of the transaction.[3]

Yet, stickiness is a complex phenomenon: Some terms seem to be resistant to change while others get updated very quickly. In this paper, we propose a novel account of stickiness and change in standard form contracts based on *experiential* learning by firms. We first outline our theory and then test it on a unique dataset of standard form contracts that tracks the changes in the End User Software License Agreements (EULAs) from 264 firms across 114 different software markets during a period of seven years, from 2003 to 2010.

We begin by observing that contract drafters may be uncertain about the net value of a contract term. As they learn over time, they may drop some terms while adding other terms. Learning might depend on many factors, which include the behavior of competitors, cases litigated in court, technological innovations, and news reports, among others.[4] These channels may depend on the types of term that the firms include in the contract but tend to be largely independent of the specific contractual choices firms make. Firms, however, also learn directly from experience with and feedback from consumers. When learning is experiential, the firm's ability to learn depends on its past contractual choices.

Consider a default implied warranty. The firm may contemplate including a waiver in the standard form contract. If the firm offers the warranty it might be able to charge a higher price for the product but it will also face some costs due to consumers claiming a remedy. The extent to which the warranty is costly and, most importantly, if such costs outweigh the benefits of the warranty to consumers, may be uncertain at the moment the firm makes its choice. Offering the default implied warranty exposes the firm to future financial liability but also offers a possibility to learn the true costs of the warranty and inform future choices.

---

[2] *See generally* Florencia Marotta-Wurgler, *What's in a Standard Form Contract? An Empirical Analysis of Software License Agreements*, 4 J. EMPIRICAL LEGAL STUD. 677 (2007); George Priest, *A Theory of the Consumer Product Warranty*, 90 YALE L.J. 1297 (1981).

[3] *See generally* MITU GULATI & ROBERT E. SCOTT, THE THREE AND A HALF MINUTE TRANSACTION: BOILERPLATE AND THE LIMITS OF CONTRACT DESIGN 33–44 (2013) (exploring theories of what makes contract terms "sticky"); Marcel Kahan & Michael Klausner, *Standardization and Innovation in Corporate Contracting (or "The Economics of Boilerplate")*, 83 VA. L. REV. 713 (1997) (examining how learning benefits and network effects may slow changes in terms); Michael Klausner, *Corporations, Corporate Law, and Networks of Contracts*, 81 VA. L. REV. 757 (1995) (examining how network effects may slow changes in terms).

[4] For a review of the literature on learning and innovation in the standard form contract setting, see Section I.

Opting for the waiver saves costs in the short run but also prevents the firm from learning. That is, the firms' choice as to the *modality* of this particular term (opting into the warranty default versus opting out) affects its ability to learn about the term's net value.[5]

Different terms are characterized by different probabilities of providing feedback about their value in the future. We distinguish between two broad categories of terms. What we name "symmetric-learning terms" are such that future information about the net value of such terms to the drafter does not depend on the initial contract. That is, the firm either learns or fails to learn about the term's value regardless of its contractual choice with regards to this term at an initial period. In "asymmetric learning" terms, instead, the firm may learn depending on whether it has adopted the default term or has opted out of it, as in the example illustrated above. Adoption of the term that guarantees learning carries with it a real option value: the firm may effectively invest in information gathering when selecting which contract terms to offer. The ex ante contractual choices of firms are thus likely to be affected by the information-generating capabilities of a particular term. Ex post, the firm can revise the contract and switch to (or away from) the default option if it has learned that it has low (or, respectively, high) costs. The prevalence of ex post switches will necessarily depend on the firm's ex ante choices and on whether those choices make the firm learn. Consequently, asymmetric learning terms that are offered in their learning modality in an initial period allow firms to learn and are thus more likely to be revised at a later stage than similar terms who are adopted in their non-learning modality.

Consider again the example of a default implied warranty. The firm learns the costs of offering the warranty only if it adopts the default term in the standard form contract. Better information about costs will allow the firm to revise the term later. If the firm opts out of the default by including a waiver of implied warranties in its contract, the firm protects itself against future liability but also forgoes the option to learn and hence will be less likely to revise the term at a later stage. In addition, the fact that the default offers an option to learn, which is absent when opting out, should increase the firm's propensity to adopt the default.

The learning that results from adoption of the warranty is likely to increase the probability that firms who adopted it will revise it at a later stage relative to those firms who decided to waive it. That is, we expect that asymmetric learning terms that are adopted in their "learning" modality (in this case, keeping the implied warranty default) at an initial stage will be revised more frequently at a later stage than when such terms are adopted in their "nonlearning" modality (in this case, including a warranty waiver). This same reasoning, appropriately modified, applies to cases in which the opt-out option provides learning. In contrast, for those terms where learning is symmetric—that is, a firm learns or not irrespective of the modality of the term—, the probability of revision at a later stage is

[5] Offering a warranty to consumers could also be a way to obtain feedback on the quality of the firm's suppliers. If the product breaks down frequently, the firm learns that its supplier delivers low-quality products. The interesting implication is that, in this case, the firm's response to learning is a change of supplier rather than a change of term. The firm may want to keep offering the warranty in order to learn about the new supplier. We do not elaborate on this alternative learning motive but we stress that this is also a form of experiential learning.

uncorrelated with the firms' contractual choice in an initial period.

Stickiness, in our framework of analysis, is the result of the inability to learn. The dynamic of contractual innovation that results from experiential learning produces stickiness for non-learning terms and for asymmetric-learning terms. However, in the latter case, stickiness occurs only when the initial balance of costs and benefits induced firms to opt for the non-learning modality and hence is the result of a conscious choice rather than a factual impossibility. Default terms, which are typically very attractive for firms and consumers alike, might for this very reason also be more prone to become sticky over time.

After reviewing the literature on standard form contracts and contractual innovation in Section I, in Section II we propose a simple theory of experiential learning by firms. In a stylized model, we describe the behavior of a firm that is face with the choice between adopting a default contractual term or opting out of it. Later, through its experience with consumers the firm may or may not learn the true costs associated with this term and, consequently, revise its initial choices and amend the contracts that regulate future transactions. We emphasize that a firm's decision to revise the terms of its standard form contract may crucially depend on the terms that the firm chose to start with. In Section III, we explore learning under different categories of terms. Since some terms allow the firm to learn asymmetrically, choosing the default or opting out of it has an effect on the firm's ability to revise the contract based on new information. Initial contractual choices generate a degree of path-dependency: firms that choose nonlearning terms at the initial stage are less likely to revise them. In addition, we investigate the extent to which stickiness depends on the authoritative power of defaults or can be explained by lack of new information due to previous contractual choices and suggest that, in our context, the latter may be more important than the former. In Section IV we investigate and test the empirical predictions deriving from our theory, using a unique dataset of standard form contracts. In Section V, we conclude with additional considerations on the normative implications of our theory and open questions for further research. The Appendix contains all the details of the theoretical model and the empirical analysis.

I. LEARNING, STICKINESS, AND INNOVATION IN STANDARD FORM CONTRACTS

The benefits of standardization are well understood and expand beyond the consumer setting, and have been explored extensively in the literature. Terms that become well known are easy for contracting parties and courts to interpret. Moreover, the use of known, similar terms confers various spillover effects, such as lower reading costs, increased certainty of legal interpretation, and reduced litigation risk.[6] The benefits created by standardization, such as

[6] Kahan & Klausner, *supra* note 3(discussing learning benefits and innovation); *see* Klausner, *supra* note 3 (discussing learning benefits, network benefits, and innovation); Avery Wiener Katz, *Standard Form Contracts*, [3] THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW 502 (Peter Newman ed., 1998) (discussing network effects); Stephen J. Choi & Mitu Gulati, *Innovation in Boilerplate Contracts: An Empirical Examination of Sovereign Bonds*, 53 EMORY L.J. 929 (2004) [hereinafter Choi & Gulati, *Innovation in*

learning and network benefits, may stand in the way of change, reducing contracting parties' incentives to revise familiar terms.[7] Markets that experience higher network benefits might also encounter stronger resistance to change and higher degrees of stickiness. Other factors also contribute to stickiness. Law firms, which are usually involved in drafting and creating new terms, but which are also organized in hierarchical manners and likely benefit from re-using their old forms, are likely to experience switching costs.[8] Weak property rights in contractual innovations are likely to further reduce incentives to innovate.[9]

Default rules can also contribute to contractual stickiness. Status quo bias can create inertia that makes switching difficult.[10] When states enact particular defaults, parties might refrain from deviating from them because customizing a term outside of the default might prove too costly.[11] Contracting parties might also be reluctant to deviate when they perceive that opting out of the default, even if value generating, might signal negative information.[12] Reluctance to change in light of a superior alternative might give rise to contractual "black holes," where parties enter agreements with terms that no longer serve the contracting goals of the parties, either because they no longer reflect the optimal allocation of rights and risks between them, or because they might be interpreted unfavourably by a court, among others.

Despite the obstacles, change and innovation can still happen. Large repeat players,

---

Boilerplate Contracts] (reviewing literature on innovation in contract terms); Clayton P. Gillette, *Lock-In Effects in Law and Norms*, 78 B.U. L. REV. 813, 819 (1998) (noting lock-in effects generated through extensive interpretation of a term).

[7] Kahan & Klausner, *supra* note 3, at 724–29 (finding that learning benefits may discourage switching).

[8] *See* GULATI & SCOTT, *supra* note 3, at 139–40 (positing that law firm structure and existing agency costs within firms further dilute incentives to innovate); Claire A. Hill, *Why Contracts Are Written in "Legalese,"* 77 CHI.-KENT L. REV. 59, 60, 80–81 (2001) (arguing that fear of mistakes may discourage attorneys from changing terms).

[9] *See* Kevin E. Davis, *The Role of Nonprofits in the Production of Boilerplate*, 104 MICH. L. REV. 1075, 1086 (2006) (arguing that contractual innovations are forms of technological progress that can generate economic growth and examining the process of contractual innovation more generally); Charles J. Goetz & Robert E. Scott, *The Limits of Expanded Choice: An Analysis of the Interactions Between Express and Implied Contract Terms*, 73 CALIF. L. REV. 261, 286, 289–305 (1985) (noting public goods aspect of standard terms); Katz, *supra* note 6, at 503 (arguing that because innovations in standard terms are public goods, the absence of intellectual property rights diminishes the incentive to innovate).

[10] Russell Korobkin, *The Status Quo Bias and Contract Default Rules*, 83 CORNELL L. REV. 608 (1998) (identifying various behavioral biases that might deter parties from moving away from default rules or established terms).

[11] Goetz & Scott, *supra* note 9 (discussing how state regulation of contract terms creates barriers to innovation).

[12] Kathryn E. Spier, *Incomplete Contracts and Signaling*, 23 RAND J. ECON. 432 (1992) (showing that if opting out signals some private information, parties might be reluctant to opt-out); Jason Scott Johnston, *Strategic Bargaining and the Economic Theory of Contract Default Rules*, 100 YALE L.J. 615 (1990) (suggesting that it will be easier for parties to bargain around expansive default rules than around restrictive or penalty default rules); Omri Ben-Shahar & John A.E. Pottow, *On the Stickiness of Default Rules*, 33 FLA. ST. U. L. REV. 651, 655–60 (2006) (arguing that deviations from known terms might raise suspicions and scare away potential counterparties). Others have identified additional sources of stickiness. Lisa Bernstein, *Social Norms and Default Rules Analysis*, 3 S. CAL. INTERDISC. L.J. 59 (1993) (explaining how social norms and negotiation strategy might lead parties to stick to default rules). In addition, the endowment effect might be playing a role in the decision of whether to deviate from existing contractual terms. Kahneman, D., J. Knetsch, and R. Thaler, *Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias*. The Journal of Economic Perspectives 5(1), 193–206 (1991). See also Korobkin, R. *Status Quo Bias and Contract Default Rules*. Cornell L. Rev., 83, p.608 (1997); Isabel Marcin and Andreas Nicklisch, *Testing the Endowment Effect for Default Rules* (2014). MPI Collective Goods Preprint, No. 2014/1. (2014).

such as law firms and investment banks, might find it profitable to invest in innovation—even in the absence of strong property rights—through their ability to spread costs among clients.[13] In-house counsel in legal departments of firms engaged in mass-market commerce work closely with management and understand changes in technology that might give rise to new terms. In addition, in-house counsel are more likely to receive feedback from offering or refraining to offer particular types of terms, allowing them to revise the agreements to adapt to new legal and market environments.[14] There are some accounts that posit that the opportunity to experiment can result in learning and change.[15] Change and innovation can also be spurred by "shocks," such as new laws, changes in legal interpretations of terms, or technological advances.

Most of the empirical evidence on contract change and innovation comes from studies of bond covenants and financial products. Marcel Kahan and Michael Klausner, among others, found evidence of switching and learning costs in the corporate bond covenant context.[16] Stephen Choi, Mitu Gulati, and Eric Posner studied the evolution of sovereign debt covenants and found an S-shaped innovation pattern, where parties slowly move from the old standard to a new one in response to various exogenous shocks.[17] There is also evidence of switching costs in law firms. Mitu Gulati and Robert Scott found that lawyers in law firms failed to revise terms even after those terms had acquired ambiguous meanings that increased litigation risk. In the handful of cases where terms were revised, this was often achieved by including additional terms and not by correcting the perceived errors in existing ones.[18] In a recent study of change and innovation in a large sample of merger agreements, John Coates found significant changes over time, finding that such contracts have doubled in size, and that

[13] *See* Kahan & Klausner, *supra* note 3; GULATI & SCOTT, *supra* note 3.

[14] *See* Stewart Macaulay, *Private Legislation and the Duty to Read—Business by IBM Machine, the Law of Contracts and Credit Cards*, 19 VAND. L. REV. 1051 (1966) (observing in 1966 that in-house counsel drafted the fine print of contracts used by large corporations, while the fine print in small firms' contracts had come from trade associations or by copying the terms used by other firms). *See also* George G. Triantis, Collaborative Contract Innovation (April 30, 2010) (unpublished manuscript) (on file with the New York University Law Review). For a discussion of modular integration more generally, see YOCHAI BENKLER, THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM 1–2 (2006) (noting the "greater scope for individual and cooperative nonmarket production" in the modern information economy).

[15] Patrick Bolton and Christopher Harris, *Strategic Experimentation*, 67 ECONOMETRICA 349 (1999) (providing the first model of strategic experimentation among many agents who can free ride on the results obtained by others). See also Godfrey Keller, Sven Rady, & Martin Cripps, *Strategic Experimentation with Exponential Bandits,* 73 ECONOMETRICA 39 (2005) for a tractable model of experimentation.

[16] *See* Kahan & Klausner, *supra* note 3, 743–53 (finding evidence of switching and learning costs in a study of the emergence and adoption of event risk covenants—terms designed to protect bondholders in the event of a leveraged acquisition); *see also* Stephen J. Choi & G. Mitu Gulati, *An Empirical Study of Securities Disclosure Practice*, 80 TUL. L. REV. 1023, 1062–66 (2006) (finding that terms were slow to change after courts interpreted a term in a new and unfavorable way, and that when change occurred, high-volume issuers' counsel spurred it).

[17] Stephen J. Choi, Mitu Gulati & Eric A. Posner, *The Dynamics of Contract Evolution*, 88 N.Y.U. L. REV. 1 (2013) (finding that innovation in business-to-business boilerplate occurs in three stages roughly similar to product innovation). *See also* Choi & Gulati, *Innovation in Boilerplate Contracts*, *supra* note 5 (examining boilerplate innovation in the context of reinterpretation of terms); Stephen J. Choi, G. Mitu Gulati & Eric A. Posner, *Pricing Terms in Sovereign Debt Contracts: A Greek Case Study with Implications for the European Crisis Resolution Mechanism*, 6 CAPITAL MARKETS L.J. 163 (2011).

[18] GULATI & SCOTT, *supra* note 3, at 10–11; *see also* Hill, *supra* note 8, at 80–81 (arguing that fear of mistakes may discourage attorneys from changing terms).

about 20% of such change can be attributed to new terms.[19]

To summarize, there have been numerous accounts to explain and document both stickiness and change in standard form contracts. In this paper, we propose a new mechanism that can account for contract change: learning from experience. To the best of our knowledge, this is the first paper to explore this mechanism in the consumer standard form contract setting. We offer some evidence in support of our hypothesis by examining a large sample of consumer EULAs over a period of time.

II. A THEORY OF THE EVOLUTION OF STANDARD FORM CONTRACTS

We propose a simple theory of contractual choice.[20] The theory captures the determinants of a firm's choice of contractual terms. In particular, we will zero in on the reasons for the firm to update the contractual terms it offers to consumers over time. To do so in a simple way, we will focus on a stylized interaction between firms and their consumers, which we introduce in in Section **Error! Reference source not found.**. After having illustrated our framework of analysis, we will discuss learning in our theory, in Section II.B. Sections II.C and **Error! Reference source not found.** discuss how our framework relates to alternative theories.

A. Framework of analysis

We postulate that firms and consumers interact at two discrete moments in time, which we call "time 0" and "time 1." (This choice also reflects the structure of the data that we will analyze in Section IV.) At time 0 a firm offers a standard form contract to consumers. The contract may decide to include a standard default term in the contract or to opt out of it.[21] In the interim period, the firm may have an opportunity to learn the costs associated with offering the term. If it learns new information, the firm may, at time 1, revise the contract.[22]

---

[19] John C. Coates, IV, *Why Have M&A Contracts Grown? Evidence from Twenty Years of Deals* (Harvard Law School John M. Olin Center Discussion Paper No. 889, European Corporate Governance Institute (ECGI) - Law Working Paper No. 333/2016 2017).

[20] The model presented here is a simplified version of the extended model introduced in the Appendix. Readers interested in the formal details of the analysis are invited to refer to the Appendix.

[21] The model only considers one alternative to the default option, while in reality there may be many. With many alternatives to the default term, the firm not only faces a decision of whether to opt out but it also has to choose among many possible terms, each of them with possibly different feedback mechanisms. Learning in this context becomes more complex and may bring about interesting interactions. After learning that, say, alternative I has high costs, the firm may decide to switch back to the default or to start experimenting with, say, alternative II, and so forth. Moreover, while we allow only one learning period, in reality the firm might learn continuously and be able to switch between one term and the other at several time periods, possibly going back to terms it had discarded in the past. This more general approach would be close to a version of the well-known "multi-armed bandit problem" in probability theory. For a classic formulation of the problem and fundamental results, see T.L. Lai & Herbert Robbins, *Asymptotically Efficient Adaptive Allocation Rules*, 6 ADVANCES IN APPLIED MATHEMATICS 4 (1985).

[22] In the real world, learning occurs continuously at several moments in time, so that it may be difficult to identify a time 0 when the firm does not know its costs and a time 1 when the firm knows. Most realistically, time 0 is preceded by an earlier stage and time 1 is followed by a later stage, both of

The default term allows the firm to capture a value v (for instance, through an increase in price due to the fact that consumers are willing to pay more for the default term) but exposes the firm to either high costs, $c_H > v$, or low cost, $c_L < v$.[23] (Opting out does not produce any value and does not impose any costs.[24]) In theory, firms would want offer the default if costs are low and opt out if costs are high, but this information is not available at time 0 and hence firms cannot tailor the contract to the costs they face.[25] Yet, at time 0 the firm knows that it will face high costs with probability p and low cost with the complementary probability $1 - p$. This setup is summarized in Table 1.[26]

Table 1. (Normalized) values and costs of default and opt-out

Term	Value with default	Value with opt-out	Cost with default	Cost with opt out	Probability
High costs	v	0	$c_H > v$	0	p
Low costs	v	0	$c_L < v$	0	$1 - p$

For example, consider firms selling tax preparation software packages to consumers. Firms may offer a warranty that they will be liable for any penalty that the user has to pay for miscalculations resulting from software bugs or other inherent malfunctions (the default term) or they may disclaim all liability (the opt out term). The decision whether to offer the

which are also characterized by learning. Considering learning through multiple periods would make our model more realistic but also substantially more complex without adding much in term of basic intuitions. One could interpret our distinction between time 0 and time 1 as focusing attention to the *incremental* learning that occurs between these two periods.

[23] Note that the value v is known to consumers and firms and that it does not necessarily captures an accurate valuation of the value of a particular term on the part of consumers. The variable v may simply capture a consumer's preference for the default term. Note also that both v and the cost variables could in theory be negative, allowing the opt out to be more valuable or costlier to implement than the default.

[24] This is a normalization that does not affect the results. See the discussion provided below in the text. In the Appendix, we present a model where both the default term and the opt out produce value for consumers and yield costs for the firm.

[25] Under a different interpretation of the model, we consider a firm that tries to infer the type of consumers it faces, of type H or of type L, which in turn is correlated with costs. Under this interpretation, our problem is similar in structure to one that has been studied widely in the past few decades: the insurance market, where insurance companies ignore the risk-characteristics of those who purchase coverage. This literature—which has been extended to cover a wider range of issues—was pioneered by Michael Rothschild & Joseph E. Stiglitz, *Equilibrium in Insurance Markets: An Essay on the Economics of Imperfect Information*, 90 Q. J. OF ECON. 629 (1976). There is an important difference between our setup and the traditional insurance analysis. The latter is a problem of asymmetric information where one party is informed and the other is uninformed. In our setup, both parties are uninformed about the costs of different terms. As we will see in Section **Error! Reference source not found.**, this difference makes the traditional solutions to this problem unworkable in our setting. There also is an equally extensive literature on the opposite problem, the one that consumers face when they are unable to distinguish between "good" and "bad" products, while firms are informed. For early, ground-breaking analyses of this problem and its legal implications see Alan Schwartz & Louis Wilde, *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, 127 UNIV. OF PENNSYLVANIA L. REV. 630 (1979); George L. Priest, *A Theory of the Consumer Product Warranty*, 90 YALE L. J. 1297 (1981). We do not deal with these issues in the present article.

[26] We are not interested in the absolute costs and benefits of offering the default, but rather in the costs and benefits *relative* to the opt out. We can therefore interpret the variables in the text as the relative value and costs of the default over the opt out, so that v measures how much *more* consumers are willing to pay for the default and similarly for the costs. Readers may consult the Appendix for a mathematical foundation of this approach.

default term or to opt out depends on the costs of offering the default warranty and consequently being exposed to liability compared to opting out and facing no costs. At the outset, firms may not have enough information to assess the costs associated with offering the warranty with absolute accuracy, yet the available information may be enough to assess the probability to face high or low costs ex post.

Firms may vary in the specifications of the products they offer, their location, the law applicable to them, the courts in which they litigate, their customers' use patterns and other factors that affect the costs of offering a particular default term. To capture this heterogeneity, we think of firms as differing in the ex ante probability p of facing high costs. The probability p can therefore be interpreted as a synthesis of the firm-specific factors mentioned above. Loosely speaking, high-p firms will face high expected costs if they offer the default, and will consequently tend to disclaim liability. Conversely, low-p firms will tend to offer the warranty because they face low expected costs.

Given a firm's characteristics summarized by p, we focus on the firm's decision whether to adopt the default term or to opt out of it at time 0 and, subsequently, whether to revise this choice at time 1. In terms of the simple model sketched above, this question requires us to identify a cutoff level of p such that firms with a p below the cutoff find it advantageous to adopt the default at time 0—these are the low-p firms facing low expected costs when offering the default—and firms with a p above the cutoff opt out—the high-p firms.

In expectation, offering the default term has a positive net expected value if the firm's expected costs, given the firm-specific p, are less than the value, that is, if $pc_H + (1 - p)c_L < v$.[27] Vice versa, if expected costs are larger than value, the default term has a negative net expected value and the firm should opt out (and earn zero). Rearranging the latter inequality allows us to define the following cut-off level for p:

$$p^* \equiv \frac{v - c_L}{c_H - c_L} \tag{1}$$

Firms characterized by $p < p^*$ have lower ex ante costs than the value of the term and hence, in expectation, can enhance the net contract surplus by adopting the default term. In

[27] The firm chooses the contract terms that maximize the net value of the contract. One reason why the firm may do that is because the firm is a monopolist and has all the bargaining power, that is, it can set the price at the consumers' willingness to pay given the specific combination of terms offered in the contract. Considering firms with less than full market power would not qualitatively alter the results. First, (some degree of) competition would reduce the firm's ability to capture consumer surplus thereby requiring us to distinguish between the value consumers attach to the default, v, and the price increase the firm is able to sustain when offering the default, which could be less than v when firms compete. This, however, would not alter our analysis in any way, as we allow v to vary. Second, competing firms might learn from each other, which both boosts learning—because it magnifies the effects of any individual firm's experimentation with new clauses—and hinders it—because it creates a free-riding problem that reduces a firm's incentives to experiment. While this aspect of the problem would add a layer of complexity to the analysis, it would not affect our basic distinction among contract terms based on their learning characteristics and hence would not qualitatively alter our results. Third, competitive forces might induce firms to follow what most of their competitors do because, for instance, consumers might be unwilling to buy a product that is offered together with an unfamiliar set of clauses. This trend might hinder learning and change but they would not alter the gist of our results. A formal model of standard form contracts in competitive markets might unveil additional implications.

contrast, firms with $p > p^*$ would detract from the contract surplus in expectation if they adopted the default term, because the default imposes larger ex ante costs than its value. (If $p = p^*$, expected costs are exactly equal to the value of the term; for ease of notation we disregard this possibility.) To fix intuitions, consider the following example.

Example 1. Firm 1's consumers value the default term at $100. Offering the default costs $150 (high costs) in 40% of the contracts and $50 (low costs) in 60% of the contracts.

In this example, we have v = \$100, c_L = \$50, c_H = \$150 and p = .4. Should the firm offer the default? Doing so costs Firm 1 an expected amount equal to .4*\$150 + .6*\$50 = \$90. Consumers value the term at \$100; hence, the firm can make a positive profit of \$10 by offering the default term. Consider a similar case, but now a firm faces high costs in a larger fraction of the contracts it offers.

Example 2. Firm 1's consumers value the default term at $100. Offering the default costs $150 (high costs) in 60% of the contracts and $50 (low costs) in 40% of the contracts.

The parameters of the example are all the same but for p which is now equal to .6. Should Firm 2 offer the default? Doing so costs Firm 2 an expected amount equal to .6*\$150 + .4*\$50 = \$110. Now the costs exceed the value of the term for the consumers. Consequently, Firm 2 would make a loss of \$10 if it offered the default term and should thus opt out.

By plugging in our parameters in Expression (1), we can easily see that p^* = .5 in the case described in Example 1 and Example 2. To see more clearly why this is the case, consider that a firm with p = .5 would face expected costs equal to .5*\$150 + .5*\$50 = \$100; that is, its expected costs from offering the default term would be exactly equal to the value of the default. Hence, this firm would be indifferent between offering the default term and opting out. Absent learning, firms with $p < .5$ (like Firm 1) find it advantageous to offer the default, while all other firms, with $p > .5$ (like Firm 2), offer the opt out. We will now examine how the firm's choice at time 0 may be affected by the possibility to learn and hence be able to revise the contract optimally at time 1.

B. Experiential learning

We are now ready to introduce a firm's learning options in the framework described above. From the contractual relationships with consumers at time 0, the firm may or may not learn useful information about the actual costs of a certain contract term. If the firm learns from the time-0 contracts, then the firm has an opportunity to revise the standard form contract. It is evident that if the firm has no new information it will have no reason to revise the contract. The revised form will apply to all subsequent transactions. For now, we assume that there are no switching or opting-out costs. We also assume that the volume of purchases does not change between time 0 and time 1 and that there is no discounting. (We will introduce switching costs and opting-out costs and consider the effect of firm growth on contractual

choice in Section III.C.)

Our theory of learning rests on two fundamental building blocks. First, we postulate that firms learn a great deal of important information from direct experience with consumers and we show that this specific learning channel has important implications for contractual choice. To focus on this issue, we assume that the only source of learning is experience with consumers who have purchased the product and hence are in a contractual relation with the firm and that there are no other sources of information. (Alternative forms of learning—such as learning from competitors, news reports, court cases and other sources of information—occur irrespective of the distinctions we make here and hence their existence does not affect our main results. We provide a detailed discussion in Section II.C.)

The second fundamental building block of our model is the recognition that the contractual choices made at the outset affect whether or not a firm learns through experience. To capture this aspect of the problem we introduce the notion of information type of a term, which can take four possible values labeled *N*, *L*, *D*, and *O*. The information type relates to whether and how the firm learns about the cost of the default term after time 0. We will see that learning is important not only for the decision whether to revise a term at time 1 but it may also affect the firm's weighting of costs and benefits at time 0.

We first consider two types of symmetric-learning terms. Terms of type *N* are "nonlearning" terms and are such that the firm receives no new information after time 0. Terms of type *L* are "learning" terms and are such that the firm receives new information at time 1 irrespective of adoption at time 0. In particular, between time 0 and time 1, the firm learns whether costs are high or low. In this case, we firm has an opportunity to revise the contract at time 1 irrespective of the choice made at time 0.

The last two types of terms involve asymmetric learning. Terms of type *D* are learning-from-default terms: between time 0 and time 1, the firm learns the value of *c* only if it has adopted the default term at time 0. Conversely, terms of type *O* are learning-from-opt-out terms: between type 0 and time 1, the firm learns the value of *c* only if it has opted out of the default at time 0.

Table 2. Information-types and modalities of contract terms

Information-type		Default	Opt-out
Symmetric learning terms	Nonlearning terms (*N*)	Nonlearning	Nonlearning
	Learning terms (*L*)	Learning	Learning
Asymmetric learning terms	Learning-from-default terms (*D*)	Learning	Nonlearning
	Learning-from-opt-out terms (*O*)	Nonlearning	Learning

Table 2 illustrates the information-type of terms that we consider in the analysis and emphasizes when each term is in a learning or nonlearning modality. The symmetric-learning terms are always in the same modality: *N*-terms are always in nonlearning modality and *L*-terms are always in learning modality, irrespective of whether the firm adopts the default

contract term or opts out of it. In contrast, asymmetric-learning terms can be in either learning or nonlearning modality depending on the contractual choice at time 0. We will analyze adoption decisions at time 0 and at time 1 by the firm for each one of the four types of terms. Section III will contain the analysis of experiential learning by a firm offering one of the four term types illustrated above. The main questions Section III will address are: How does the information-type of a term affect contractual choices at time 0 and at time 1? The remainder of the present section further elaborates on the scope of our theory.

C. Scope of the theory

The theoretical framework presented above focuses on learning direct costs, but the choice of terms can generate other forms of learning that affect contractual choices at a later stage. Consider, for example, a retailer that sells products manufactured by a number of suppliers and is uncertain about the quality of the products of each supplier. Offering a secondary warranty for the products sold exposes the retailer to claims made by consumers under the warranty, allowing it to learn the varying degrees of quality of each supplier. In this case, the term allows the seller to learn about supplier types rather than consumer types. Similarly, a choice of law clause may or may not be desirable depending on whether it lowers or raises the costs of litigating a case in court for the firm depending on unknown factors, determining whether the firm faces high or low costs. Our analysis applies also to these cases.

More generally, our framework applies to all cases in which firms offer standard form contracts to consumers and learn from interacting with them. Yet, firms routinely attempt to tailor their contracts to the specific characteristics of individual or groups of consumers. Moreover, firms invest relevant resources in learning through other channels. In the following two subsections, we show defend the generality of our analysis in the face of these two sets of possible objections.

1. Standard form contracts versus tailored contracts

Firms in our model offer standard for contracts to consumers. Consumer characteristics may affect the costs that offering a specific term generates and hence firms might find it advantageous to tailor contracts to specific consumers or consumer groups rather than offering all of their consumers the same contract.

A widely-studied way for a firm to tailor a contract to the specific characteristics of its consumers is to "screen" consumers by offering different contracts at different prices and letting consumers choose their preferred contract.[28] However, firms in our model cannot do so

[28] The theory of screening has been pioneered by Joseph E. Stiglitz, *The Theory of Screening, Education and the Distribution of Income*, 65 AM. ECON. REV. 283 (1975) and is a contractual response to the problems arising in markets with asymmetric information, as identified by George A. Akerlof, *The Market for 'Lemons': Quality Uncertainty and the Market Mechanism*, 84 Q. J. OF ECON. 488 (1970). In a screening model, the uninformed party makes an offer to the informed party. The offer allows the informed party to choose between two different contracts, say, one with a warranty and one without, at two different prices. If appropriately designed, such offer results in a separation between the two (or more) types of consumers; in our context, this means that consumers

because consumers are uniform with respect to the value, *v*, that they attach to the default term and ignorant about the costs that alternative terms impose on the firm.[29] Hence, if offered different contracts, consumers in our model would all choose the same contract, defeating the firms' attempt to screen among them.[30]

There are two reasons why our framework of analysis plausibly captures the reality of standard form contracts. First, although consumers are typically not perfectly homogenous in their evaluation of different contract terms and might know something about the costs that those terms impose on firms, consumers often know *less*, rather than more, than firms do about their own future use patterns, exposure to risk, probability of accidents and other important factors that determine both the value and the costs of different terms.[31]

Second, contract standardization offers numerous advantages to firms, which would be lost if the firm were to tailor the contract to individual consumers, making standardization advantageous even in those cases in which tailoring would be theoretically possible. There is abundant empirical evidence suggesting that retail chains in the United States do not adapt their prices to local conditions. That is, those firms do not only fail to tailor their contracts to the characteristics of specific consumers but also fail to adapt their contracts to easily identifiable groups of consumers leaving in different states and hence with markedly different preferences, wealth, education and, ultimately, willingness to pay for certain products. Price terms are allegedly the easiest terms to vary in a contract and hence evidence that retail chains do not tailor prices is strong evidence of relevant fixed costs in tailoring contracts also with respect to other, more difficult to individualize, terms.[32]

2. Alternative learning channels

While our model focuses on experiential learning, under certain circumstances, sellers can also learn from various alternative mechanisms, including litigated cases, the behavior of other firms operating in the same market, industry studies, and feedback from legal counsel.

For instance, firms can learn by being aware of technological changes and the ways consumers use products, rather than from experience with the terms themselves. Consider

---

of type *L* would choose the contract with the default warranty and consumers of type *H* would choose the (cheaper) contract without warranty.

[29] In many cases the value that a consumer attaches to a contract term varies with the consumer's use patterns, which in turn determine the costs borne by the firm. A warranty may provide the typical example: high-risk consumers both value the warranty more and impose larger costs to the firm. Under these conditions, heterogeneous valuations are correlated with consumer types and allow the firm to adopt a screening strategy. We do not consider these cases for reasons explained below in the text.

[30] Note that screening presupposes that consumers know their type. This is not the case in our model.

[31] For an early recognition of this problem see Alan Schwartz & Louis Wilde, *Imperfect Information in Markets for Contract Terms: The Examples of Warranties and Security Interests*, 69 VIRGINIA L. REV. 1387 (1983). For a recent detailed analysis of various ways in which firms exploit consumers' lack of information, see Oren Bar-Gill, SEDUCTION BY CONTRACT: LAW, ECONOMICS, AND PSYCHOLOGY IN CONSUMER MARKETS (2012).

[32] Stefano Della Vigna & Matthew Gentzkow, *Uniform Pricing in US Retail Chains*, Working Paper (documenting price uniformity across the stores of consumer chain and finding support for the that varying prices is costly for the chain).

again a software firm that includes a term in its contract allowing the buyer to modify the software. We classify such clauses as nonlearning because there is no learning that occurs from experience. Rather, the term either allows the consumer to customize aspects of the software to make it compatible with other software or not. This is a product attribute that, while affecting demand for the product, is unlikely to generate learning by including it in the contract. Yet, firms can learn about consumer preferences and uses of software by looking at purchasing patterns or examining the offerings of competitors in their own market and adapt terms accordingly. Change occurs, but it is driven by a different mechanism as compared with the one we examine.

Similarly, firms can learn from litigated cases about possible contractual choices as well as about the enforceability of particular clauses (e.g., a "change of terms" clause that allows firms to modify standard agreements unilaterally), both of which affect the decision to adopt terms in an initial period as well as the decision to revise terms at a later time. The literature on contractual innovation has also pointed out that firms can learn from other firms by studying the latter's standard form contracts, creating a well-known free rider problem.[33] In addition, as discussed in more detail in Section I, law firms can also transmit knowledge to their seller-clients, who can then revise their terms accordingly. In sum, the type of learning that might affect a firm's contractual choices can occur in many ways.

Other means of learning, while possible, are generally orthogonal to the experiential learning capability of each term and hence do not affect contractual choice in the same way as we examine here. We leave it to further research to explore the possible interactions between experiential learning and other learning channel.

III. SYMMETRIC VERSUS ASYMMETRIC LEARNING

In this section, we analyze the contractual choices of firms in light of the taxonomy of terms presented in the previous section.

A. Symmetric-learning terms

First, we focus on the two symmetric learning terms: nonlearning terms and learning terms. While these terms appear radically different at the outset, the analysis will reveal important similarities.

1. Nonlearning terms

Nonlearning terms (*N*) have the feature that no new information resulting from the firm's

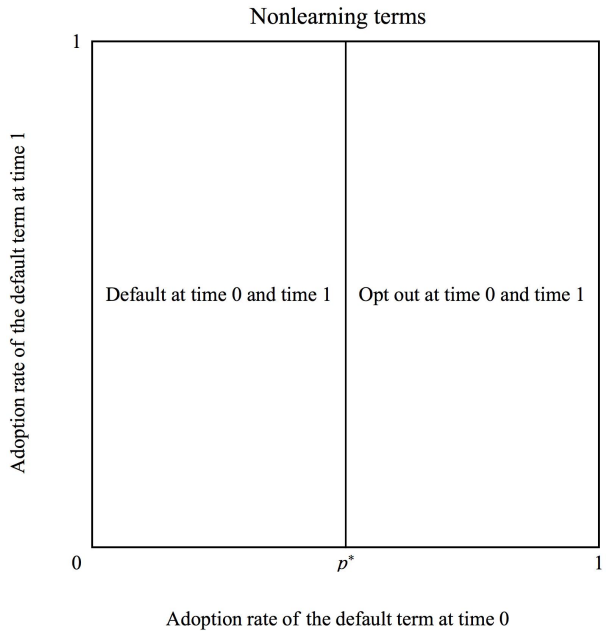
[33] *See* Goetz & Scott, supra note 9; Kevin E. Davis, The Role of Nonprofits in the Production of Boilerplate, 104 Mich. L. Rev. 1075 (2006) (arguing that non-profits can produce boilerplate which might otherwise be lacking due to the free rider problem associated with the creation of boilerplate). Relatedly, learning can also occur by examining the publications of trade associations.

experience in adopting a particular modality of the term is available at time 1, when the firm has the option to revise the contract. These terms are likely to reflect product attributes or pure information, such as a term limiting the number of devices to which a user can download a software program or licensed song. Product attributes define what product the firm sells to consumers. In this case, the firm could sell a license for one devise or a license for five devices. The firm is unlikely to learn anything about its costs from offering either version of the term. The value of these terms is more likely to be affected (and ascertained) by changes in technology or the way consumers listen to music, which the firm will not learn directly through dealing with its consumers but rather, most likely, through external channels, which are not the focus of this model.

Consider again Example 1. Firm 1 knows that 40% of its contracts routinely result in high costs (the remaining 60% result in low costs). The firm will not learn any new information after interacting with the time-0 consumers. Some of the time-0 contracts will then have resulted in high costs, but the probability that any new contract entered into at time 1 will generate high costs is the same as it was at time 0. Since the firm-specific value of p is known ex ante and does not change between time 0 and time 1, Firm 1 will offer the default both at time 0 and at time 1. In Example 2, instead, Firm 2 will do the opposite and opt out both at time 0 and at time 1 because the firm's expected costs are higher than those of Firm 1. Those firms do not switch and contracts remain unchanged over time.

More generally, at both times, it is advantageous for the firm to adopt the default term if the expected costs of the term are lower than its value and to opt out of the default otherwise. Firms characterized by low expected costs, such as Firm 1—that is, those with $p < p^*$, as defined in Expression (1)—will offer the default, while the remaining firms will opt out.

Figure 1. Nonlearning terms: adoption decisions at time 0 and time 1



We will now visualize this result graphically. To do so, we will focus on the special case in which v falls exactly half-way between c_L and c_H. This case includes our examples but applies more generally.[34] Figure 1 shows graphically this result, displaying the adoption rates of the default term at time 0 and time 1 across firms under the parameter values used in the examples above, which imply that $p^* = .5$, as observed above.

Given that different firms are characterized by different levels of p, the adoption rate of the default term at time 0 is equal to the fraction of firms with $p < p^*$ (which in turn is equal to p^* if we assume a uniform distribution of firms.)[35] The adoption rate at time 0 is displayed on the horizontal axis of the graph. At time 1, there are no switches and hence the adoption rate of the default term for time-0 adopters—among which is Firm 1—is equal to 1, while the adoption rate for time-0 non-adopters—including Firm 2—is equal to 0. All the action takes place at time 0, where 50% of all firms choose the default and 50% the opt out.

2. Learning terms

Consider now a term L, which is characterized by learning after time 0. That is, while the real cost of the term is unknown to the firm at time 0, it is known at time 1 due to the firm's experience with consumers. Examples of such terms are choice of law clauses and fee-shifting clauses. In these cases, the firm learns information from the outcome of litigation, which follows directly from its relationship with consumers, but choosing the default—which implies choosing the law of the forum and adopting a fee-shifting rule that makes a losing firm reimburse the attorney's fee of the consumer—or opting out does not seem to affect learning. Consider a firm that disclaimed the fee-shifting rule at time 0. This provision is valuable if the firm is exposed to frequent disputes because it might lower the firm's litigation costs, encourage consumers to abandon weak cases early on and, consequently, lower the amount or possible settlements. The firm will learn its exposure to consumer disputes from interacting with consumers after time 0. Yet, the firm would learn the same information even if it adopted the default fee-shifting rule.

Very differently from the nonlearning term examined above, now learning occurs from experience with consumers. However, similarly to nonlearning terms, also in this case the term chosen does not affect whether the firm learns or not. Therefore, the optimal choice at time 0, when c is still unknown, is again to adopt the default term if $p < p^*$ and not to do so if $p > p^*$. The reason is that what is at stake is again only the expected net value of the

[34] More precisely, we assume $v = \frac{c_L + c_H}{2}$; this is the case in Example 1 and Example 2, where $v = \frac{50+150}{2} = 100$. This drives the result $p^* = .5$. In general, v could be closer to c_L than to c_H or vice versa and hence p^* could fall anywhere between 0 and 1, moving the vertical line in the graph to the left or to the right, respectively. The same would occur with the other three figures. Yet, the interpretation given in the text would not change qualitatively.

[35] The uniform distribution implies that a fraction p^* of the firms is characterized by $p < p^*$. For instance, if $p^* = .5$, then 50% of the firms have $p < .5$ choose the default; the remaining 50% of the firms have $p > .5$ and choose the opt out, as depicted in Figure 1. Using the uniform distribution makes it easier to display and discuss the results.

contract. Learning from experience occurs irrespective of the firm choice at time 0 and hence it does not affect this choice. Therefore, even if here the firm learns, the time-0 choice is the same as with nonlearning terms. To fix intuitions, consider the following example, which slightly modifies Example 1.

Example 3. Firm 3's consumers value the default at $100. Offering the default costs either $150 (high costs) in all of the contracts, with probability .4, or $50 (low costs) in all of the contracts, with probability .6.

In this example, we have again v = \$100, c_L = \$50 and c_H = \$150; also the value of p is the same as in Example 1 (p = .4 = 40%) and hence Firm 3 should offer the default at time 0, as in the case of nonlearning terms for Firm 1. Consider now a similar case, but now the firm faces a larger probability of high-cost consumers.

Example 4. Firm 4's consumers value the default at $100. Offering the default costs either $150 (high costs) in all of the contracts, with probability .6, or $50 (low costs) in all of the contracts, with probability .4.

We know that Firm 4 will opt out at time 0. As before, a firm characterized by a customer base p = .5 faces costs equal to .5*\$150 + .5*\$50 = \$100, which is exactly equal to the value of the term. So far, the results are the same as in the nonlearning case. Example 3, however, differs from Example 1 in one important dimension (and so does Example 4 from Example 2). In Example 1 40% of the contracts result in high costs, while in Example 3 there is a probability equal to .4 that *all* contracts will result in high costs. Although the firm does not know at time 0 whether it is facing a low-cost or a high-cost situation, this ambiguity will be fully resolved after the firm observes the cost yielded by the time-0 contracts.

Therefore, after interacting with the consumers at time 0, Firm 3 learns whether costs are high or low, while Firm 1 did not learn.[36] Since the firm has new information about its costs, the optimal decision at time 1 is adoption of the default term if costs are low, c_L, and opting out if costs are high, c_H.

Figure 2. Learning terms: adoption decisions at time 0 and time 1

[36] In the Appendix, we will formalize this intuition by explaining that if costs are correlated across different contracts, as they are in Example 3, then the firm learns something important about the costs it will face at time 1 by interacting with consumers at time 0. Instead, if costs are uncorrelated across different contracts, as they are in Example 1, then the firm does not learn.

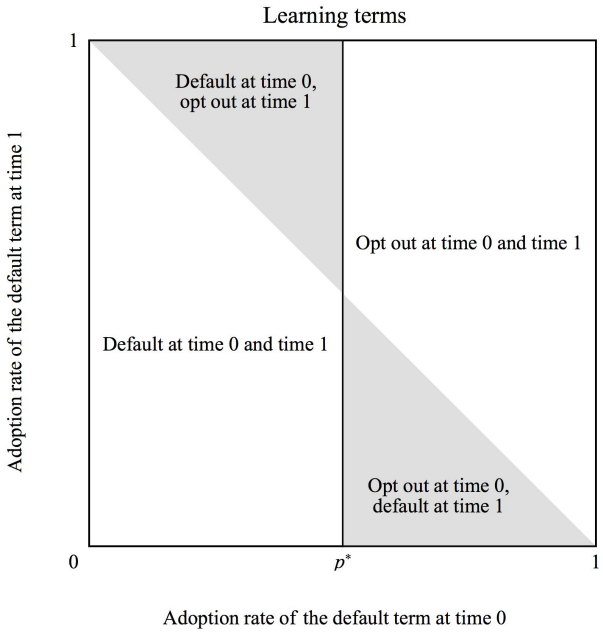


Figure 2 shows the adoption rates of the default term at time 0 and time 1 for learning terms.[37] At time 0 firms that adopt the default are those with $p < p^*$, like Firm 3. At time 1, however, there is new information available and with probability p an adopter discovers that costs are high and decides to switch and opt out of the default (with the complementary probability $1 - p$, the adopter discovers that costs are low and keeps adopting the default term). The area of the grey triangle depicts the ex ante probability mass of switches from adoption of the default term to opt-out. Note that the lower a firm's p, the lower the probability that that firm switches from default to opt out at time 1, because it is unlikely that costs are high. In the example, Firm 3 will switch at time 1 to the opt out with probability .4 and will retain the default with probability .6.

Conversely, Firms with $p > p^*$, like Firm 4, decide not to adopt the default term at time 0. At time 1, with probability p, such firms discover that the cost is in fact high and confirm the opt-out decision, while they discover that costs are low with probability $1 - p$ and switch. The grey triangle depicts again the ex ante probability of switches, now from opt out to default. Firm 4 in the example will switch at time 1 to the default with probability .4 and will retain the opt out with probability .6. Considering all firms in the case described in Figure 2, 50% of the firm choose the default at time 0 and 25% of them (that is 12.5% of the total number of firms)[38] switch to the opt out at time 1. The remaining 50% of the firm choose the opt out at time 0 and again 25% of them switch to the default at time 1. In total, 25% of all firms in switch; 12.5% from default to opt out and 12.5% in the other direction. Note that while the time-0 choices are the same as with nonlearning terms, the situation at time 1 is very different.

[37] Recall that in all figures we focus on the special case in which v falls exactly half-way between c_L and c_H. See note 34.

[38] Note that the area of the gray triangle is equal to 1/8 or 12.5% of the square of area equal to 1, which represents the totality of the firms on the market.

B. Asymmetric-learning terms

We now consider asymmetric learning terms, starting for the learning-from-default type, *D*, and then moving to the learning-from-opt-out terms, *O*. Since each of these two cases is the mirror image of the other, all the important results will be the same and we will analyze in depth only the learning-from-default terms. We will simply state the results for the learning-from-opt-out terms to stress what changes. Most importantly, from the analysis it will emerge that the prospect of asymmetric learning induces firms to alter their time-0 contractual choices, which did not occur with symmetric learning terms.

1. Learning-from-default terms

With learning-from-default terms, *D*, the firm learns its costs only if it has adopted the default contract term at time 0. Adoption of the default gives the firm the option to learn and revise its decision at a later time. In contrast, the opt-out alternative does not imply any learning and hence the optimal decision for the firm at time 1 is to confirm the decision taken at time 0.

Consider a term that disclaims consequential damages. Firms adopting the default (no limitation) will face future claims with some probability and learn the amount of those damages, which will be related to the product, the way consumers use it, the activities consumers are involved in, interaction with other products, local conditions, and other factors that will typically be unknown at the outset. Faced with specific claims, the firm will learn the cost of being exposed to liability for consequential damages. Firms opting out of the default and disclaiming consequential damages will not face claims (or claims will not result in a judicial assessment of damages) and hence will not learn.

The asymmetric learning value of the alternatives available to the firm at time 0 gives rise to what is known as a "real option."[39] By choosing the learning modality (in this case, the default term), the firm purchases an option to reverse its decision later on, when new information becomes available. In turn, this option confers additional value to the choice of the default at time 0, which goes beyond the mere expected value of the clause in terms of its

[39] The literature on real options starts with Robert McDonald & Daniel Siegel, *The Value of Waiting to Invest*, 101 Q.J. ECON. 707 (1986) (showing that irreversible decisions to invest can be understood using the framework developed in finance for the study of option contracts). Real option theory has been applied to the study of diverse topics. *See* Alexander J. Triantis & George G. Triantis, *Timing Problems in Contract Breach Decisions*, 41 J. LEGAL STUD. 163 (1998) (applying option theory to the study of contract breach); Douglas G. Baird & Edward R. Morrison, *Bankruptcy Decision Making*, 17 J.L. ECON. & ORG. 356 (2001) (applying option theory to bankruptcy); Lee Anne Fennell, *Revealing Options*, 118 HARV. L. REV. 1399 (2005) (applying option theory to the study of property and liability); Joseph A. Grundfest & Peter H. Huang, *The Unexpected Value of Litigation: A Real Options Perspective*, 58 STAN. L. REV. 1267 (2006) (applying option theory to the study of litigation); Jacob E. Gersen & Eric A. Posner, *Timing Rules and Legal Institutions*, 121 HARV. L. REV. 543, 544-46 (2007); Christopher A. Cotropia, *Describing Patents as Real Options*, 34 J. CORP. L. 1127 (2009) (applying option theory to the study of intellectual property); Matthew Spitzer & Eric Talley, *On Experimentation and Real Options in Financial Regulation*, 43 J. LEGAL STUD. S121 (2014) (applying option theory to the study of financial regulation); Joe Vladeck, *Valuing Regulatory Flexibility: A Real Options Approach to Cost-Benefit Analysis*, 103 GEO. L.J. 797 (2015) (applying option theory to regulatory impact analysis). For an encompassing analysis of how option theory affects the study of the law see IAN AYRES, OPTIONAL LAW: THE STRUCTURE OF LEGAL ENTITLEMENTS (2005).

costs and benefits. The value of the real option to switch at time 1 enhances the value of adoption of the default term at time 0. Therefore, the optimal decision threshold at time 0 will be grater than p^*. Consider again Example 4 and recall that Firm 4 offered the opt out at time 0 when the term was a learning term. Now let us consider the example in the context of an asymmetric, learning-from-default, term and re-evaluate whether Firm 4 should adopt the default.

Firm 4 has a "riskless" strategy. It can opt out at time 0. In this case, the firm does not make any profit from offering this particular clause and does not learn anything. Having learned nothing new, the firm will opt out also at time 1. The firm's profits from opting out are zero in Example 4: consumers value the opt out at zero and the firm faces no costs if it opts out, because, for instance, it has disclaimed liability.

Firm 4 may, however, also consider a more "risky" strategy and choose the default at time 0, that is, choose to be exposed to liability. In that case, the firm makes a loss equal to \$10, which is the difference between the expected costs (.6*\$150 + .4*\$50 = \$110) and the benefits (\$100) of offering the default. This choice, however, allows the firm to learn about costs. Consequently, when the firm chooses the default at time 0, its choice at time 1 will be made under full information. If costs are low, the firm will offer the default and make profits equal to \$100 – \$50 = \$50 at time 1. Instead, if costs are high, the firm will offer the opt out (since offering the default would result in a loss equal to \$150 – \$100 = \$50) and make zero profits.

Which strategy is more profitable? Ex ante, choosing the default generates a loss of \$10 for sure at time 0, but also gives the firm a probability of .6 to learn costs are low and keep offering the default at time 1 if costs are low, while switching to the opt out (and hence avoid a loss) if costs are high. Summing up, the firm pays \$10 at time 0 and earns .4*\$50 + .6*\$0= \$20 in expectation at time 1. In total, the firm expects to earn \$20 – \$10 = \$10 if it chooses the default. Earning an expected payoff of \$10 with the default is better than earning zero for sure by opting out; hence, Firm 4 will offer the default term.

Recall that Firm 4 chose the opt out with (always) learning terms while it chooses the default if the term is a learning-from-default only term. The reason is that now the default has an "option value" of \$20, consisting of the possibility to learn and hence take a better decision in the future, which is greater than its expected costs of \$10 (net of the value v). Yet, learning occurs also in (always) learning clauses. Why is there no option value? The option value comes from the asymmetry in learning associated with the particular modality these terms take. With learning terms, learning occurs irrespective of the firm's choices at time 0 so there is no reason for the firm to alter the choice at time 0. In learning-from-default terms, however, learning is asymmetric and hence the default has a *higher* value than the opt out in terms of learning and the option to switch. Firms are thus induced to choose the default more often for these kinds of terms as compared to terms where learning is symmetric. Summing up, with learning terms Firm 3 chooses the default and Firm 4 chooses the opt out, while with learning-from-default terms both firms choose the default. (It is easy to see that Firm 3, too, would choose the default because its expected costs are even less than those borne by Firm

4.)

Yet, not all firms choose the default. When p is particularly high, the costs of adopting the default at time 0 are large and the probability of learning that costs are low at time 1 is small, leading to a small option value. There will therefore be a value of p above which the firm chooses to opt out even though choosing the default would imply learning. In our examples, this value is $\overline{p} = \frac{2}{3}$, which is greater than p^*. The payoff from choosing the default at time 0 is then $\$100 - \frac{2}{3}\$150 - \frac{1}{3}\$50 = -\frac{\$50}{3}$, while in expectation, the firm gains $\frac{1}{3}(\$100 - \$50) = \frac{\$50}{3}$ at time 1. The gain perfectly balances the loss, making the firm indifferent between the default and the opt out. This also suggests that a firm with $p > \frac{2}{3}$ would incur a net loss if it chose the default and hence will choose to opt out.[40]

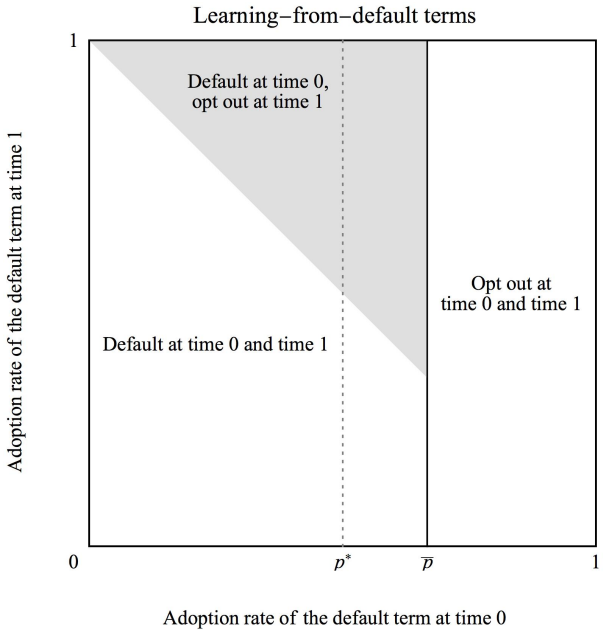
Figure 3 shows the effects of asymmetric learning from the default.[41] The cut-off level of p at time 0 is greater than the one we found with learning and nonlearning terms: more firms choose the default at time 0. Adopters, however, switch to opt-out with relatively high probability, especially in the range $[p^*, \overline{p}]$, that is, in those cases that would have resulted in opt-out at time 0 had the term been of a symmetric type. These are instances in which the default has a negative expected value and it is chosen purely for its learning value, that it, for the option to make a perfectly informed decision at time 1.[42] Note the asymmetry: only those firms that choose the default at time 0 consider switching at time 1.

Figure 3. Learning-from-default terms: adoption decisions at time 0 and time 1

[40] In the general case, a firm considers that if it opts out it will earn 0 from it during both periods. If it adopts the default term, it will earn $v - (1 - p)c_L - pc_H$ at time 0; then it will learn the costs and will keep adopting the default term only if costs are low, which occurs with probability $1 - p$ and yields a payoff equal to $v - c_L$ for sure. Otherwise it will switch to the opt out and earn zero. The ex ante payoff from adoption of the default term at time 0 is hence $v - (1 - p)c_L - pc_H + (1 - p)(v - c_L)$. The ex ante payoff from opting out is 0. The former is greater than the latter if p is greater than $\overline{p} \equiv 2\frac{v - c_L}{c_H - c_L + v - c_L} > p^*$.

[41] We focus again on the special case in which v falls exactly half-way between c_L and c_H. See note 34.

[42] In symmetric learning terms, the alternative (the opt-out) also has the same learning characteristics as the default and hence the choice between the two is not affected by the option to learn and this effect does not arise.



We can use again the different regions in Figure 3 to gain some insights on adoption rates. At time zero, 1/3 (that is, 33.3%) of the firms opt out of the default. Those firms do not learn hence there are no switches at time 1. The remaining 2/3 (66.7%) of the firms offer the default terms—this is substantially more than with symmetric learning terms—but 1/3 of them discover that costs are high and switch to the opt out. In total, 2/9 (that is, 22.2%) of all firms switch at time 1 (this is the area of the gray region in the graph). Compared to the learning clauses, there are substantially more switches from default to opt out: 22.2% instead of 12.5%. (Although in this specific example there are fewer switches in total: 22.2% instead of 25%.)

It is interesting to note that, due to the unidirectional switches for type-*D* terms, at time 1 there are more firms opting out from a learning-from-default term than from a learning term and, namely, 55.6% instead of 50%.[43] That that can happen should not come as a surprise. The option value of the learning modality of a term pushes firms to choose that term far beyond what would be optimal in the short term. As a result, firms choose terms in their learning modality in a set of cases in which their choice is relatively unlikely to be confirmed after learning. This short-term inefficiency is the price that firms pay to learn and choose better terms in the future. The high switching rates away from the learning modality at time 1 are due to firms choosing the learning modality "too often" at time 0 and to the fact that only those "errors" are corrected by learning. Firms that choose the nonlearning modality at time 0 do not learn and hence do not switch. Finally, let us stress that these quantitative examples

[43] With learning terms, in Figure 2 switches are symmetric and hence switches in opposite directions balance each other out. The result is a 50% adoption rate both at time 0 and at time 1 for the default and the opt out. With learning-from-default terms, in Figure 3, switches are unidirectional. 1/3 of the firms opt out at time 0 and keep opting out at time 1. In addition, 2/9 of all firms choose the default at time 0 only to opt out at time 1. Hence in total, 1/3 + 2/9 = 5/9 of all firms opt out at time 1, while only 4/9 choose the default. In percentage terms, these fractions amount to 55.6% for the opt out and only 44.4% for the default.

are limited to the set of cases captured by the figures—those in which v falls halfway between the low and high costs—and hence not all of them hold generally. In particular, the comparison between total switching and adoption rates at time 1 with symmetric learning clauses and with asymmetric learning clauses depends on the value of v compared to the high and low costs. The intuitions and the logic behind them, though, do carry beyond these restrictions.

2. Learning-from-opt-out terms

We now consider other type of asymmetric learning term, the "learning-from-opt-out" type, *O*. Here the firm learns its cost only if it has opted out of the default at time 0. Consider a term offering maintenance and support for the first two months. The default term is that there is no such service. Firms opting out, and hence offering the service, are exposed to customers' requests and learn the costs of offering the term. Firms that adopt the default do not learn.

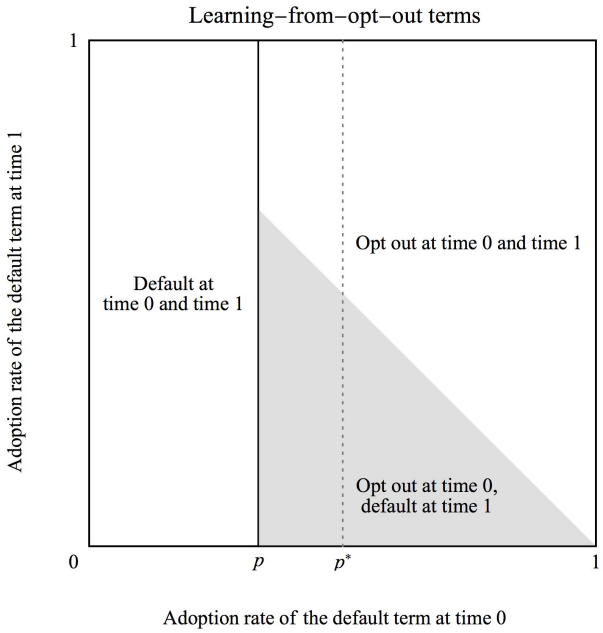
This type of term is the mirror image of the type *D* studied above: the intuition is the same but the results are reversed. Now opting out, rather than adoption of the default, has an added option value, with increased opt-out rates at time 0 and some switches to the default term at time 1. Adoption of the default term at time 0, conversely, implies no learning and hence no switches at time 1.

Mirroring the previous case, we can also now find a general threshold $\underline{p}$ below which the firm chooses the default, but now this threshold is less than p^*.[44] Figure 4 shows the effects of asymmetric learning in learning-from-opt-out terms.[45] The time-0 adoption rate is less than it was with learning terms, and perfectly symmetrical to the *D*-terms. Those who offer the default at time 0 do so also at time 1. Those who opt out at time 0 switch back to the default with relatively high probability at time 1.

Figure 4. Learning-from-opt-out terms: adoption decisions at time 0 and time 1

[44] The general formula for the threshold is $\underline{p} \equiv \frac{v-c_L}{2(c_H-c_L)-(v-c_L)} < p^*$.

[45] We focus again on the special case in which v falls exactly half-way between c_L and c_H. See note 34.



C. *Other determinants of contractual choice*

In this section, we enrich the model to encompass other determinants of the firm's contractual choices. Table 3 summarizes the results of the discussion that follows. A formal analysis is provided in the Appendix.

Table 3. Other determinants of contractual choice (The symbols +, - and = indicate whether the determinant considered makes the choice of the default respectively more, less or equally likely as compared to the basic model; "?" indicates that the effect is ambiguous.)

	Terms			
	Nonlearning	Learning	Learning-from-default	Learning-from-opt-out
Opting-out costs	+	+	+	+
Switching costs	=	?	–	+
Growth	=	=	+	–

1. Opting-out costs

In the basic model presented above we have assumed that the firm equally weighs the default option and the opt out. In reality, however, there may be several reasons why firms might be, ceteris paribus, more inclined to choose the default option provided by the law. Reasons may include consumers' distrust for non-standard contract terms or legal research costs.[46] In the mass market consumer context, however, some opt outs have become standard. For example, many sellers offer limited warranties and remedies instead of implied warranties of

[46] *See* Goetz & Scott, supra note 9.

merchantability. While common, there are positive research and drafting costs associated with such clauses.

Considering opting-out costs makes the default comparatively more valuable, so in our model, increases the relative value variable *v*. The result is straightforward and easily predictable: the default option will be chosen more often across all typologies of terms.

2. Switching costs

A second and important category of costs that the basic model does not consider are switching costs. Unlike opting-out costs, switching costs are incurred only if the firm switches at time 1 and affect symmetrically both switches from the default term to the opt out and the opposite switches.

Introducing switching costs adds a drag to firm choices, making it more expensive to adapt the contract over time. From a different perspective, switching costs may be perceived as imposing a tax on learning: if the firm learns, it will have to pay the switching-cost tax to switch. Thus, learning becomes less attractive. The effect is intuitive in asymmetric learning terms. Since the "tax" applies only to the learning modality of a term, there will be less learning. In learning-from-default terms the default will be chosen less often and, symmetrically, in learning-from-opt-out terms the default will be chosen more often as switching costs increase.

In nonlearning clauses there is, quite obviously, no effect since there is no learning. In learning terms the effect is ambiguous and depends on the adoption rates of the default. In our examples, we derived an adoption rate equal to 50%, in this case switching costs do not affect choices because they apply equally to switches from the default to the opt out and to the opposite switches. The baseline adoption rate, however, does not need to be 50%. With different values for the variables in our model, any adoption rate between 0 and 100% is feasible. If the adoption rate is below 50%, most of the firms choose the opt out at time 0 and, hence, most of the switches at time 1 will move away from the opt out term. Switching costs will then affect the opt out more than they affect the default term, making the adoption rate increase towards 50%. If the adoption rate is above 50%, the opposite happens. Now the switching costs affect the default more heavily, making the adoption rate decrease towards 50%.

In sum, in asymmetric learning terms, switching costs will tend to reduce adoption of the learning modality of a term. In symmetric learning terms, adding switching costs makes learning more symmetric, working against the learning modality that is most often chosen.

3. Growth prospects

The basic model presented above was based on an assumption of constant profitability. Firms, however, differ substantially in their growth prospects. Young and innovative firms might expect to increase their sales in the future, while older firms or firms in more conventional markets might expect sales to stay constant or to decline.

Introducing consideration for a firm's growth prospects suggests that we should weigh differently sales at time 0 and sales at time 1. If a firm expects to grow, sales at time 1 should receive a larger weight than sales at time 0 and vice versa. In turn, if sales in the future are more important, learning becomes more advantageous. Thus, the expectation of growth works in favor of the learning modality of a term: the default will be chosen more often in learning-from-default terms and less often in learning-from-opt-out terms. Yet, there is no effect on learning terms. The reason is that there learning occurs anyway irrespective of the choice at time 0 and hence putting more or less weight on learning does not affect these choices.

IV. EMPIRICAL ANALYSIS

In this section, we put our theory to bear on the contractual made by real firms. We first derive empirically testable predictions from the theory presented in earlier sections. We then illustrate our dataset and present the empirical results.

A. Empirical implications of the theory

We now offer a number of testable predictions regarding the learning model in Section II about the main determinants of a firm's decision to amend its standard terms over time when learning is experiential. While it is difficult to disentangle empirically the reasons behind firms' adoption of terms in an initial period (given the multitude of factors likely affecting such decisions, many of which are hard to measure), examining firms' decisions to revise such terms at a later period can offer some interesting insights regarding possible drivers of contractual choice. Here we contrast the attractive power of default terms with learning from previous contractual choices.

Prediction 1. The probability that a firm will amend an asymmetric-learning term at time 1 is higher if the firm has chosen the learning modality at time 0.

Asymmetric-learning terms are the most exposed to the effects of learning because only one of the modalities in which the term comes allows the firm to learn, while the other precludes the acquisition of experiential information. Some terms allow the firm to learn only if the default option is chosen (the learning-from-default terms) so that the learning modality is the default. In other cases, it is opting out that generates learning.

Prediction 1 emphasizes these implications: the firm's decision to revise an asymmetric learning term is largely affected by the firm's choice at time 0. Learning puts the firm in the position to re-evaluate past contractual choices and amend them if new information suggests that a different choice is more advantageous. Prediction 1 also identifies a mechanism by which "black holes" could come about. If the firm has chosen a nonlearning modality at time 0, it will not see new information and might fail to revise the term in question at time 1. Inefficient or meaningless terms might survive due to the asymmetric

nature of learning. An interesting result is that firms that chose the nonlearning modality of a term at time 1, inefficient terms might persist at time 1 in the contracts offered by firms that chose the nonlearning modality at time 0, while firms adopting the learning modality of the same term at time 0 stay away from them. Such "black holes" or pockets of inefficiency might affect only a portion of the firms in the market.

Prediction 2. The probability that a firm will amend a symmetric-learning term at time 1 does not depend on the term chosen at time 0.

Prediction 2 focuses on the effect of learning in symmetric learning terms. Contrary to asymmetric-learning terms, here the firm's initial choice does not affect the firm's propensity to revise the term. With nonlearning terms, the result is obvious: the firm does not learn from experience and hence does not revise its terms based on new information. Revisions will likely come from information acquired elsewhere, which is not connected with the firm's contractual choices at time 0. With learning terms, the result is less intuitive. The firm does learn from experience in this case. However, the firm learns symmetrically from both the default and the opt-out option. Consider for example a choice of forum clause. If the firm includes one and gets sued, the firm will experience the relative value of having included one. If the same firm does not include a choice of forum clause in the initial period and gets sued, the firm will also experience the relative value (or cost) of not having one. Learning, while not identical in its form or content, will occur under both conditions. Regardless of the contractual choice, the firm learns about the value of the clause at the later period. As a result, new experiential information informs the firm's decision at time 1 irrespective of the contractual choices made at time 0. We should observe revisions motivated by experience in this case but such revisions should be equally likely for firms that adopted the default and for firms that opted out of it at time 0.

Prediction 3. If default terms are inefficiently often chosen at time 0, default terms will be amended more frequently than non-default terms if they offer an opportunity to learn.

Default contractual terms have long been recognized as important determinants of contractual choice. Implications of this observation come in two guises. On the one hand, if default terms are more frequently chosen, this could apply both at time 0 and at time 1. If, however, the choice of a term is largely determined by the term being a default, default choices at time 0 are more likely to result in inefficient outcomes. We anticipate that such defaults will be more likely to be amended at time 1 if the firm has had an opportunity to learn in the meantime. This effect should be visible both in symmetric and in asymmetric learning terms. In the symmetric ones, the learning terms will be revised at time 1 more often towards the opt-out option if the default was inefficiently chosen at time 0. In asymmetric learning terms, revision should be more frequent when the default is the learning modality (learning-from-default terms) than when it is the nonlearning modality (learning-from-opt-out terms).

Both implications point to an important role of default contractual terms in determining firm choices going forward. If this is the case, switches at time 1 should be largely explained by

the fact that a term is a default. This prediction will allow us to contrast defaults to learning as alternative explanations for change in standard form contracts. We turn to the empirical analysis in the next section.

B. Data and methodology

We test our hypotheses using a sample of software license agreements governing the use of pre-packaged software. End-User License Agreements (EULAs) typically present a rich set of standard terms; while the terms typically vary both across and within markets, EULAs follow a predictable structure.[47] This allows for meaningful comparisons across contracts. We examine the rate of change of terms from 2003 to 2010 in accordance with sellers' opportunity to learn from their presence or absence.

We use the sample of EULAs used in a previous study examining other questions of term change and innovation.[48] The EULAs are from 264 firms with comparable data in 2003 and 2010, ranging from well-known software publishers to smaller companies. For each company and its representative EULA we include information on a representative product as well as various market and company characteristics.

For each EULA in each period, we tabulate the presence of 32 standard terms across seven categories of related terms, such as scope, warranties, limitations of damages, etc. We further classify each term into different categories reflecting the extent to which offering a given term gives sellers an opportunity to learn, either symmetrically or asymmetrically. We also take account of other factors that might affect firms' decisions to revise terms at a later time, such as their size, age, and whether they have in-house counsel.

1. Summary statistics

We used the data set introduced in Marotta-Wurgler and Taylor (2013), which tracks the changes in the terms of EULAs found in typical "prepackaged" (i.e., non-customized) software products and compare their content in 2003 and 2010. That study examined the change in 32 EULA terms from 246 firms that sell their software on their corporate Internet sites, including large, well-known, software publishers, as well as smaller companies. For each of the companies, the dataset includes a representative product along with data on various market, product, company characteristics, and of course, the EULA both in 2003 and in 2010.

Table 5 presents summary statistics. Panel A reports company characteristics for the sample firms. Average revenue in 2003 was $287.5 million and the median was $1.7 million. Average and median revenue in 2010 were $539.1 million and $2.2 million, respectively. The percentage of public companies grew from 11% in 2003 to 14% in 2010.

[47] Marotta-Wurgler, *supra* note 2.
[48] For a full description of the data collection process, see Florencia Marotta-Wurgler and Robert Taylor, *Set in Stone? Change and Innovation in Consumer Standard Form Contracts*, 88(1) N.Y.U. L. REV. 240 (2013).

The sample includes data on legal sophistication in 2010, proxied by firms' choice of legal advice, including whether they have in-house counsel, at least one internal lawyer, or routinely hire outside counsel. All public companies are assumed to receive sophisticated legal advice. In total, 74% of firms for which these data were available received relatively intensive legal advice, which might affect firms' propensity to revise terms at a later date.

Panel B lists product and market characteristics in 2003 and 2010. The average price of the products in the sample was $812 in 2003 and $841 in 2010. Thirty-six percent of the products are oriented toward consumers or small home businesses, rather than large businesses. One percent of the products in the sample were discontinued, but the company used the same EULA for all their products in 2003 and 2010. Firms are classified firms into 114 distinct software markets, as classified by Amazon.com, the largest Internet software retailer.[49]

Panel C reports contract characteristics. We first record whether at least one of the thirty-two terms we track was revised in any way during the sample period. Of the entire sample, 40% of contracts changed at least one substantive term. Of the 103 contracts that had at least one change (39% of 264), change was limited to one or two terms, but a few firms changed their contracts significantly, including some that changed more than ten terms. Contract length increase, from 1,517 words in 2003 to 1,938 in 2010, or an average of 27 percent. The median word increase in contracts with no material changes was one word, whereas the median word increase in the EULAs with material changes was 435 words.

2. Classifying symmetric and asymmetric learning terms

We classify the 32 terms into four categories that reflect drafters' opportunity to learn. Each term is described in detail in Marotta-Wurgler and Taylor (2013) and its presence is measured against the benchmark of the default rules of Article 2 of the Uniform Commercial Code. We note if a term matches the default rule provided in Article 2 (given that such rules would fill any gaps to the extent a contract is silent on a given issue) and if a term deviates or opts-out of such default rule. A contract can adopt the default rule either by including a term that matches such rule or by remaining silent. These classifications are outlined in Table 7.

Not all terms give sellers the same opportunities to learn. Table 7 also reports how we classify each term depending on whether some terms allow for symmetric learning (or failure to learn) or whether learning is asymmetrically tied to the seller adopting the default rule or opting out of it. A further explanation of the reasons behind the learning classification of each term is in the Appendix. Consider a term that allows the seller to collect and/or share the consumer's personal information. Whether that term is offered or not, the seller is likely to receive feedback regarding the value of such activity. The act of collecting information will

[49] These markets are very finely defined and can be grouped into larger, more general, markets. For example, Amazon defines one market as "Office Suites," which is included in a larger market labelled "Business and Office." For a detailed account of these variables and the methodology used, see Florencia Marotta-Wurgler, *Competition and the Quality of Standard Form Contracts*, 5 J. EMPIRICAL LEGAL STUD. 447, 457–67 (2008).

inform seller about the value of the activity. Failure to collect may also inform the seller over time whether the product or service is hurting the seller's competitive advantage or whether it makes the product more appealing to consumers. Learning is symmetric for all modalities of the term. The table labels such terms as "S (L)"—i.e., symmetric learning. We identify three additional terms as symmetric learning terms. These include two terms related to dispute resolution: the seller gets to experience whether the chosen law (or the failure to offer one) or whether who pays for attorneys' fees is optimal. The last symmetric learning term allows the seller to disable the software remotely in case the buyer breaches. Again, regardless of its modality, a seller learns whether it is desirable to have such a clause (assuming it is feasible for the seller to offer it) whenever the seller experiences a buyer breach. All terms and the rationale for coding decisions are explained in the Appendix.

Terms that never allow learning regardless of their modality are labelled "S (N)"—symmetric nonlearning. We identify eleven such terms. These include: a term notifying the consumer that the product can be returned if she declines terms; one change of terms clause that allows the seller to unilaterally amend the contract; one term noting whether the licensed product includes updates or upgrades; one term delineating the scope of the use rights granted by limiting the buyer's ability to modify or alter the program; three terms explaining whether there are transfer limitations or other license grant restrictions; one noting whether the disclaimer is in caps or otherwise conspicuously presented (this is not a term *per se*, but one that tracks a requirement under Magnuson Moss Warranty Act); two terms related to the rights of third parties; and one term informing consumers of their statutory rights outside the contract. A common characteristic of S (N) terms is that they either supply information about the product or define the features of the product as opposed to allocating rights and risks between sellers and buyers. While important, given their nature, these types of terms do no lend themselves to allow seller to learn about their value from experiencing their use with consumers. This doesn't mean that such terms will not be revised. Indeed, demand for more flexible products, or products that can be installed in multiple devices, might lead sellers to revise these terms. But the mechanism through which sellers learn will be less direct.

The coding for most of these clauses is straightforward. Of course, one could disagree with our classification and argue that a nonlearning term would actually allow the seller to learn, very much like a symmetric learning term. Consider a change of terms clause, which allows the seller to modify the agreement. We currently code such clause as nonlearning (because it establishes a mechanism by which modifications will occur, without imparting or withholding any rights from consumers—the modification could also happen absent such clause), but one could imagine that a seller that uses that clause and fails to adequately inform consumers of the modification or does not provide them with an opportunity to reject the modification, might find itself without an enforceable modification or, worse, without any term to enforce if the court decides such an expansive term renders the contract illusory. In this circumstance, the clause exposes the seller to learning. Failure to include the clause also allows a seller who wishes to modify the agreement in a simple, streamlined way, and thus also allows the seller to learn. For this reason, we group symmetric learning clauses together

in our empirical analysis.

We now turn to asymmetric learning clauses. In contrast to the pure information or product feature terms, a term like an express warranty results in asymmetric learning, as the seller learns its relative value only by offering one. We separate these terms into those where sellers learn by experience only when they opt out of the default and when then they learn only when they adopt the default. There are no default express warranties, so the seller learns only by opting out of the default (or, A (O)). We identify five such clauses. These include: one term allowing the drafter to install software to monitor users' activities, three tracking whether the seller offers limited or full warranties, and one tracking whether the software includes maintenance and support services (here, the sellers is not obligated to do so unless this is promised in the contract; such promise exposes the seller to consumer demands).

In contrast, if the seller offers default implied warranties, it might learn the value of such offering. In this case, adopting the default allows the seller to learn. We label these clauses A (D)—i.e., asymmetric default. We find twelve such terms. These include two clauses allowing the buyer to create derivative works and reverse engineering (which are allowed under intellectual property laws), a choice of forum clause (where the learns seller the costs of not providing one if sued in inconvenient forum), as well as nine clauses disclaiming implied warranties, various risks, or damages.

For each term and category of term, Table 7 reports the mean opt-out from the relevant default rules in both 2003 and 2010, as well as the mean change during the sample period. For example, in 2003, 55.3 percent of firms included a term capping damages at less or equal the purchase price, a term we classify as A(D)—which our hypothesis predicts sellers would be more likely to revise in the later period if they offer the learning modality of the term. This number decreased slightly in 2010, to 51.9 percent of firms choosing to opt out of the default rules. The difference of 3.4 percent, while small, is statistically significant at the 10% level.

C. Analysis

We now explore the extent to which the changes reported in Table 7 are more likely depending on the initial choice of terms as well as when sellers have an opportunity to learn. Panel A in Table 7 begins by exploring the stickiness of default rules in the data by reporting the extent to which sellers chose to match the default rules of the UCC at the initial period as well as the probability of revising a term given their initial modality in the previous period. The top right figure shows that among 32 terms in total, and 8,448 EULA-term observations, 30.8% of all terms in 2003 were at the opt-out value, whereas the remainder, or 69.2%, matched the default rules, indicating a strong gravitational pull towards the default previously identified in the literature.

Yet default terms are not set in stone. In 2010, the fraction of terms that match the default decreased to 66.7%. Indeed, 65.3% of all terms were at default values in both 2003 and 2010, but 3.9% were at default values in 2003 and opted out in 2010. In terms of

probabilities, the right panel shows that the probability of changing a term in 2010 given that a term was in an opt-out and default value in 2003 was 0.045 and 0.056, respectively. The 0.011 difference is statistically significant at the 5% level. While terms are more likely to begin at the default, the probability that they will be revised at a later period is larger if the term starts at the default, offering support to the known view that sellers might be inefficiently choosing default due to stickiness. Marotta-Wurgler and Taylor (2013) posit that this may be caused by sellers' incentives to opt-out of consumer-friendly UCC defaults, despite stickiness or inertia.

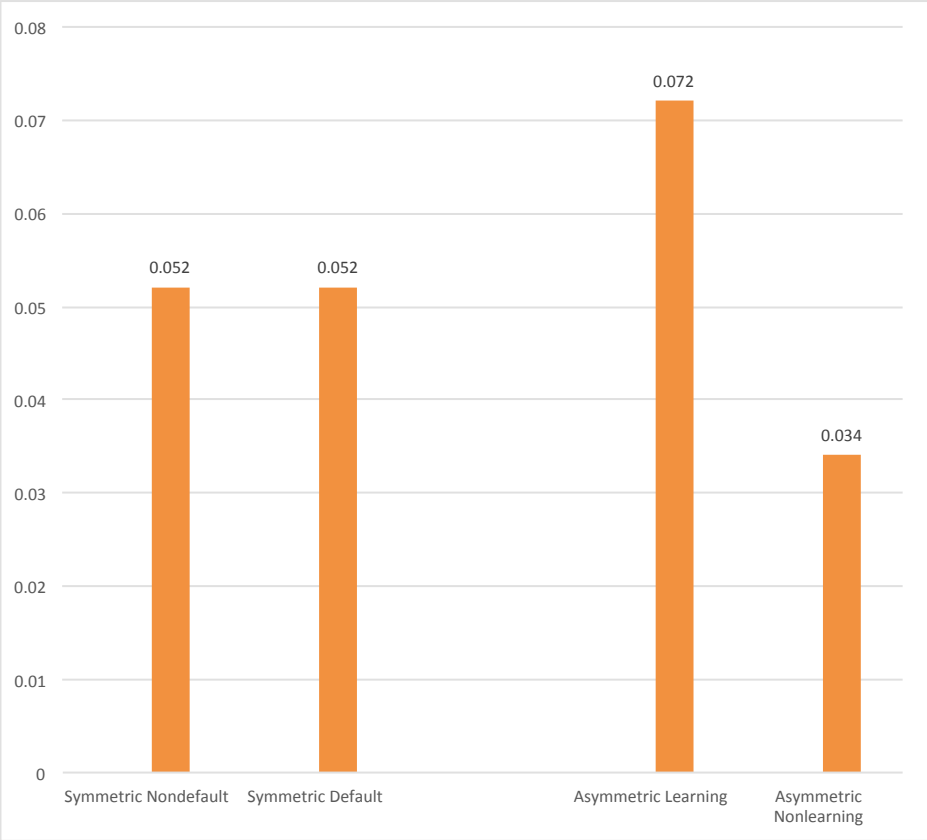
With this baseline in mind, we test predictions 1 and 2 by dividing the data into whether the term generates symmetric or asymmetric learning opportunities. Panel B presents data on symmetric learning by grouping both learning and nonlearning terms alike. As noted earlier, sellers might be learning about these terms through other means, independent from experience and irrespective of whether the term matches the default rule or not. We have no a priori hypotheses as to how these additional sources may inform sellers. We thus combine all symmetric terms. For our purposes, all we care is to know whether change is more likely to be associated with one modality of the term or the other.

The results show that, again, defaults are powerful determinants of contract terms in the initial period. In this case 75% of symmetric terms match the default rule in 2003, only to change to 72.4% in 2010, indicating some change away from defaults. More interesting for our purposes, however, is the probability of change conditional on the starting point. Recall that we predicted that the starting point for these types of clauses would be a poor predictor of change. In fact, the probability of changing a term is precisely the same, or 5.2% depending on where the term is in 2003.

Contrast this with Panel C, the results for asymmetric terms. In 2003, 64.2% of all such terms matched the default rules of the UCC, a number that shrank to 61.8% in 2010. The right panel shows that the probability of change for terms that matched the default in 2003 is 6.1%, in contrast to 4.2% for non-defaults. The difference is significant at the 5% level. Even for the asymmetric learning clauses—and consistent with the findings in Panel A examining all terms—terms are more likely to be revised when they start at the default rule, regardless of the learning modality.

Once we divide asymmetric terms up into their learning modalities, a new picture emerges, as seen in the bottom panel of Panel C. In 2003, asymmetric terms are included in their learning and nonlearning modalities about equally. However, and in contrast to the symmetric terms, where the probability of changing a term was independent of the original allocation of the term between default and nondefault, in the asymmetric scenario, the original learning modality matters. The probability of changing a term given that the 2003 contract included such term in its learning modality is 0.072, in sharp contrast to the 0.034 that occurs when the term is not in its learning mode. The findings support the basic prediction that opportunity to learn helps to explain contractual change and innovation.

Figure 1. Probability of Term Change


0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0.052
0.052
0.072
0.034
Symmetric Nondefault
Symmetric Default
Asymmetric Learning
Asymmetric Nonlearning

These findings are illustrated in Figure 1. The left bars show the probability of change conditional on their 2003 starting point (default versus opt-out). The bars are the same height, consistent with the modality of the term conferring no consistent learning advantage. Contrast this to the bars on the right. Change is more likely to happen if the terms are switched on their learning modes in 2003, as opposed to their nonlearning mode.

Table 9 reports regressions including company, product, and market control variables. The first column just repeats the results from the bottom of Panel C of Table 8. The second column adds firm (contract) fixed effects, controlling for the overall propensity of a given contract to change. The fact that the coefficient on learning does not budge indicates that there is not a tendency for some firms to make wholesale changes to their policies, including their learning terms; a given learning term is equally likely to change "within" a contract whether the same firm is changing many or few other terms. The third and fourth columns show that the probability of changing away from a term at the default in 2003 is robust to the overall propensity to change the contract, but the effect is only half that of the probability of changing the term as a function of the term's learning status, and is a distinct effect.

The last two columns add a variety of potentially interesting control variables, but with no effect on the learning coefficient of interest. Note that fixed effects cannot be included here because the variables do not vary within a given contract. We see that multi-user licenses are less likely to change. One hypothesis, which we cannot test, is that such licenses were, in general, given more thought in the first place. It also appears that when the firm is selling increasingly expensive products, its contract terms are more likely to change. Finally, the presence of lawyers is associated with change.

Finally, Table 10 presents some refinements by dividing asymmetric terms into whether the learning modality is at the default or at opt out. It repeats the exercise in Table 8 and reveals that, when learning occurs by keeping the default, firms are more likely to include the term at the initial period (59.9%, as compared to 40.1%, as seen in the left portion of Panel A). This is not the case for when learning occurs at opt out (where only 25.5% of such terms are operationalized in their learning modality), as noted in Panel B. The latter might be the result of the stickiness of defaults. Change in the later period, however, is more likely when terms are set in their learning modality in their initial period, regardless of whether learning occurs at the default or at opt-out, consistent with our prediction. The right hand of Panel A shows that when learning occurs at the default, terms that were offered in their learning mode in 2003 had a 7.3% probability to change, compared to 3.2% of terms that were in their nonlearning mode. The difference is significant at the 1% level. The same is true for terms where learning occurs from opt-out. These are 7.1% likely to change when offered in their learning mode, compared to 3.5% when they are not. Again, the results are highly statistically significant. Note that the results in Panel A support Prediction 3, which states that terms are more likely to be revised from inefficiently chosen defaults when such defaults carry and opportunity to learn.

Discussion

While the stickiness of default rules is apparent from the findings, the results support the hypothesis that learning plays a role in how standard form contract terms change over time. Our study focuses on a particular setting—consumer EULAs—, but the learning mechanism we present could be present and examined in other markets and settings, involving other types of contracting parties.

Of course, there are competing hypotheses that could explain the desire to revise terms, such as opting out of defaults. After all, defaults may be chosen because firms are unaware about a term and might revise them because they become aware of it later. Also, default terms tend to benefit consumers, so sellers might revise as way to allocate part of the surplus to themselves. Assuming all (or most) default terms benefit consumers relative to the opt-out, then we would expect a shift away from *all* defaults with the same frequency. Yet we don't see this. Rather, it's those defaults that carry an opportunity to learn those that get revised more frequently. Defaults that possess different learning modalities also tend to benefit consumers (e.g., a contract without a choice of law clause gives consumers more options in where to bring suit; similarly, a contract that doesn't include restrictions on the consumer's ability to modify the software is also more beneficial to consumers, all else equal). Yet only those that carry an opportunity to learn are more likely to be revised. Of course, both motivations could co-exist. Sellers could be opting out of consumer-friendly defaults with the motive of drafting more self-serving contracts, but also be more likely to revise those terms and term modalities that are associated with learning.

Our theory and findings suggest that, to the extent that having an opportunity to learn encourages contracting parties to revise their terms in ways the increase the benefits from the contract. Normatively, this insight could play a role in law makers' decision-making when deciding whether to create or modify default rules (e.g., in the context of revising Articles or provisions of the U.C.C). All else equal, it might be more desirable to adopt defaults that carry an opportunity to learn. When learning occurs by opting out of the default, the stickiness inherent (and oftentimes built up) in default rules hampers learning. This offers an additional reason for not making default rules sticky in these particular circumstances.[50]

V. CONCLUSIONS

Standard form contracts include terms that may benefit consumers and generate costs for the firm in ways that are not perfectly predictable at the outset. Adopting a contract term is often akin to experimentation: the firm may accept the risk of short-term losses in order to learn the net value of the term and take a better-informed decision in the future. Yet, only some terms offer an opportunity to learn and may do so in different ways.

[50] See Schwartz, Alan, and Robert E. Scott, *The Common Law of Contract and the Default Rule Project,* 102 Va. L. Rev. 1523, 1566-68 (2016) (offering a critique of projects of law reform that seek to establish contract default rules and discussing the problems with the creation of such default rules).

We have introduced a distinction between two main categories of terms: symmetric-learning terms are terms that offer symmetric opportunities to learn to firms that adopt them and to firms that do not adopt them; asymmetric-learning terms are those that offer an opportunity to learn either to adopting firms or to non-adopting firms, but not to both. Exploiting differences in the way firms learn from their contractual choices, we have built a theory of experiential learning in standard form contracts. The theory predicts that firms will be more likely to revise terms that offer an opportunity to learn and might fail to revise terms that do not offer such an opportunity. Through this lens, we have examined and classified the terms included in the End User Software License Agreements (EULAs) by a sample of 264 firms across 114 different software markets in 2003 and in 2010. We found that learning opportunities are a determinant of change, overcoming the stickiness of defaults. When such opportunities are absent, terms may survive long enough to appear obsolete and out of touch with the rest of the contract.

A. Does experiential learning produce biases pro or against consumers?

Our analysis shows that a firm's contractual choices are affected by experiential learning if the term is of the asymmetric learning type. Here we stress a possible empirical regularity: the modality of a term that allows learning is most likely the one exposing the firm to costs rather than the safe option, from the firm perspective, irrespective of whether that modality is the default or the opt out. In our analysis, we have seen that with respect to liability for consequential damages, the learning modality is the default: the firm learns about the costs of consequential damages liability if it is exposed to it. In contrast, with respect to maintenance and support, the learning modality is the opt out, because the default would not additional maintenance and support and would hence not expose the firm to learning.

In both of these case and, plausibly, in the vast majority of the cases, the modality that exposes the firm to experiencing costs is also the modality that expands the rights of consumers, offers more generous warranties, protects against a broader range of damages, and grants more effective right of redress. These terms are generally seen as being pro-consumer. Does experiential learning push firms to adopt pro-consumer terms more often than they would otherwise do?

To be sure, the question is not whether firms are biased pro or against consumers, but rather whether experiential learning induces them to behave as if they were. Our analysis reveals that, at time 0, when costs are still unknown, firms are more likely to offer the learning modality of a term rather than its nonlearning modality. At time 1, however, precisely because of the time-0 choices, firms switch in large numbers from the learning to the non-learning modality, while switches in the other direction are rare because no learning occurs if the firm chose the nonlearning modality at time 0.

Observationally, these trends produce the empirical equivalent of a pro-consumer *bias* at time 0—because there is a push towards choosing the modality that most favors consumers at the outset—and of the opposite *trend* against consumers at time 1, when most of the pro-consumer provisions are taken out of contracts. Our analysis cautions against interpreting the

resulting empirical regularities as biases, as they may be the effect of the firms' strategic learning choices.

B. The normative implications of experiential learning

Our analysis focuses on learning from experience and we have stressed the firm's behavior in response to information about the costs of offering certain clauses.[51] In general, such learning is beneficial because it allows the firm to offer terms that maximize the value of consumer contracts. This observation speaks against the stickiness of default terms: defaults should not be sticky because stickiness distorts the process of learning and prevents firms from opting out of a default terms in cases in which this choice would otherwise be optimal. From a normative viewpoint, the law should make contractual choices as neutral as possible as leveraging on the attractiveness of default provisions comes with a possibly high cost.

C. Open questions for further research

The analysis we present in this article opens, we hope, interesting avenues for further theoretical and empirical inquiry. To our knowledge, we are the first to identify the learning modalities of different terms and to draw conclusions for contractual choices. Yet, we use a rather rigid, binary classification that does not allow us to distinguish modalities that imply more or less learning. Further research could provide interesting insights into the learning potential of different terms: which terms allow firms to learn the most? Learning also occurs through different channels, as we have emphasized. How do these channels interact and what channels do firms mostly use? Finally, how does new technology affect the way in which firms learn—does it make experiential learning faster, more effective or redundant?

VI. APPENDIX

A. Theory

1. Model setup

We introduce here a formal model of contractual choice, which generalizes the setup presented in the text. There are two sets of players: (i) a population of firms described by the density function $f(p)$, where $p \in [0,1]$ is a firm-specific characteristic as specified below; and (ii) a population of homogeneous consumers. Consumers and firms interact at two points in time, denoted time 0 and time 1; however, each individual consumer is on the market only once, that is, either at time 0 or at time 1. A representative firm offers its consumers a standard form contract that can come in two guises. The firm can either adopt a default term

[51] We recognize that firms may also experiment ways in which they could exploit consumers. There is a large literature about this and similar problems and we do not examine it here.

prescribed by the law or opt out of it.

The default term has a known value v^D for consumers, while the alternative opt-out term has known value v^O. The costs of the default, c^D, and the opt out, c^O, however, are uncertain. With probability p, the firm faces a high-cost (H) default term, which costs $c^D = c_H^D$, or a high-cost opt-out term, which costs $c^O = c_H^O$. With the complementary probability $1-p$, the firm faces a low-cost term (L), which costs $c^D = c_L^D$ or $c^O = c_L^O$, respectively. [52] The upper part of Table 4 illustrates this setup.

Table 4. Values and costs of default and opt-out

	Value with default	Value with opt-out	Cost with default	Cost with opt out	Probability
H	v^D	v^O	c_H^D	c_H^O	p
L	v^D	v^O	c_L^D	c_L^O	$1-p$

	Relative value of default	Relative cost of default	Probability
H	$v = v^D - v^O$	$c_H = c_H^D - c_H^O$	p
L	$v = v^D - v^O$	$c_L = c_L^D - c_L^O$	$1-p$

We assume that the firm can set the price at the consumers' willingness to pay.[53] Hence, the firm maximizes its net payoff—that is, the difference between the value and the costs of the term—while choosing between the default and the opt out. If it had full information, the firm should adopt the default if $v^D - c^D > v^O - c^O$ (the net payoff from offering the default is larger than the net payoff from offering the opt out) and opt out otherwise.

For the purpose of the analysis, it is therefore useful to specify explicitly the difference between the value and costs of the default and those of the opt out. Let $v \equiv v^D - v^O$ be the value of the default relative to the opt out and $c \equiv c^D - c^O$ be the cost of the default relative to the opt out. Using this more compact notation, we can write the condition for adopting the default with full information as a comparison between the relative value and the relative cost of the default:

$$v > c \tag{2}$$

Note that c can take either of two values, depending on whether the term is a high- or low-cost term. With probability p, we have $c = c_H$, where $c_H \equiv c_H^D - c_H^O$ is the cost of the default relative to the opt out for high-cost terms. With the complementary probability $1-p$, we have $c = c_L$, where $c_L \equiv c_L^D - c_L^O$ is the cost of the default relative to the opt out for low-

[52] Note that we do not make any assumption as to the relationship between c_L^D and c_H^D and, in particular, we allow for $c_L^D < c_H^D$, $c_L^D = c_H^D$ and even $c_L^D > c_H^D$. Similarly, for c_L^O and c_H^O. The meaning of *high* and *low* will be clarified below in the text in relative terms as a comparison between the costs of the default and the costs of the opt out.

[53] See the main text for a discussion of this assumption.

cost terms, as summarized in the bottom part of Table 4.

Consistently with our notation and without loss of generality, we assume that $c_H > c_L$, which allows us to finally give real meaning to the labels H and L and suggests the following (relativistic) interpretation of the costs imposed by different terms. Low-cost (L) terms are those that impose *lower* costs on the firm when the firm offers the default version of the term than when the firm opts out; similarly, high-cost (H) terms are those that impose *higher* costs on the firm when the firm offers the default version of the term than when the firm opts out.[54]

To capture the simple idea that the firm chooses between the default and the opt out depending on which option guarantees the larger net payoff, we investigate the relationship between the relative costs and benefits of different contractual choices. If $v > c_H$ or $v < c_L$, the firm invariably offers the default or the opt out, respectively, and hence there is no scope for learning. These are uninteresting cases for our purposes. If, instead, $c_L < v < c_H$, the firm should ideally offer the default if the default term imposes relatively low costs (L) and the opt out if the default term imposes relatively high costs (H).

However, since neither the firm nor its consumers observe costs at time 0, it is useful to define a measure of relative expected costs. At time 0, the expected cost of the default relative to the opt out is:

$$\mathrm{E}[c] = \mathrm{E}[c^D - c^O] = pc_H + (1-p)c_L \tag{3}$$

Before examining the firm's contractual choices, let us add two additional ingredients to the model. First, switching at time 1 entails a cost for the firm, denoted $s \geq 0$, which captures the costs of rewriting the contract, doing additional legal research, informing the consumers and so on.

Second, time-0 and time-1 profits might weigh differently on the firm's time-0 choices because the expected volume of sales at time 0 might be greater or less than that at time 0. Let w be the weight of time-1 sales in the time-0 decision. Firms that expect to grow are characterized by $w > 1$: time-1 sales are expected to be greater than time-0 sales (which have weight equal to 1). Conversely, firms that expect to lose consumers are characterized by $w < 1$.

2. Information and learning

At time 0 the firm knows the value of a term—that is, it knows v^D and v^O—and the distribution of costs—that is, it knows the probability p and the values of c_L^D, c_H^D, c_L^O and c_H^O—but it does not know whether the costs associated with the default term and the opt out are high or low—that is, at time 0 the firm does not observe c^D and c^O and hence does not know c. Therefore, contractual choices at time 0 are made in a condition of incomplete

[54] Note that this is an innocuous labeling choice. Assuming $c_H < c_L$ would simply require us to give a somewhat unintuitive interpretation of the cost variables.

information. Consumers only know their valuations v^D and v^O.[55]

Time-0 purchases by consumers make the firm experience the costs associated with the term chosen. In particular, the firm will observe c^D or c^O depending on whether it offers the default or the opt out. Whether observation of the costs leads to learning useful information depends on whether costs are correlated across consumers or not. We consider two alternative—and somewhat extreme—scenarios.[56]

Uncorrelated costs. In this scenario, each contract may result in either high or low costs with probability p and $1-p$, respectively. Although the firm observes the costs following by individual contracts, it does not learn anything new about the costs it will face in the future. The firm faces and will continue to face average costs when offering a standard form contract to its consumers.[57] Terms characterized by uncorrelated costs are *nonlearning terms*.

In the opposite scenario, when costs are perfectly correlated, either all purchases result in high costs or all result in low costs. The ex ante probability is still the same as before, but now the firm has a possibility to learn exactly whether costs are high or low by observing realized costs ex post. If the firm observes high costs, it may learn that in the future the opt out will yield the larger payoff. Conversely, if the firm observes low costs, it may conclude that the default option is preferred. In this case, the variable p is the ex ante probability that each contract—and, given perfect correlation, all contracts—results in high costs. However, since the firm observes costs, learning crucially depends on the firm's ability to tell high and low costs apart, which it will not be able to do if high and low costs are of the same magnitude. With this observation in mind, we can consider three subcases.[58]

Perfectly correlated costs with $c_L^D < c_H^D$ *and* $c_L^O < c_H^O$. In this case the firm, by observing realized costs, can infer it is facing high- or low-cost terms irrespective of the term it offered at time 0. Terms characterized by symmetric learning, irrespective of the option chosen, are labelled *learning terms*.

Perfectly correlated costs with $c_L^D < c_H^D$ *and* $c_L^O = c_H^O$. In this case the firm, by observing realized costs, can infer whether it is facing high- or low-cost terms only if it offered the default at time 0. The opt out is characterized by the same cost for both high- and

[55] We discuss these assumptions in the main text. In addition, note that the firm would not be able to screen consumers and offer different contracts at different prices to different set of consumers even if consumers were aware of the costs. This is because they are identical with respect to the amount of money, v, that they are willing to pay for the default over the opt out. Moreover, even if the firm could learn the costs imposed by a specific consumer at time 0, consumers buy only once and hence the firm would not be able to condition the contract offered to individual consumers at time 1 on their type. Thus, as we will see below, information on costs (even if related to a specific consumer's characteristics, his or her "type") is only valuable if it can be used to make inference about the population of consumers, rather than on individual consumers.

[56] Considering intermediate cases would yield less sharp results.

[57] In fact, the actual costs faced by the firm in this case vary stochastically. Yet, if the population is large, the expected costs are tightly distributed around the population average.

[58] In what follows, we use an intuitive ordering of costs with $c_L^D \leq c_H^D$ and $c_L^O \leq c_H^O$ for ease of interpretation. However, none of our results depends on this assumption and we could allow for $c_L^D > c_H^D$ and / or $c_L^O > c_H^O$. What is important in our classification of terms that allow the firm to learn is only whether we have $c_L^D = c_H^D$ or $c_L^D \neq c_H^D$, and similarly for the costs of the opt out.

low-cost terms and hence observing the realized cost does not yield information on the type of term: even after observing the cost, the firm cannot distinguish between c_L^O and c_H^O. Terms characterized by asymmetric learning that occurs only if the option chosen is the default are labelled *learning-from-default terms*.

Perfectly correlated costs with $c_L^D = c_H^D$ *and* $c_L^O < c_H^O$. In this case the firm, by observing realized costs, can infer whether it is facing high- or low-cost terms only if it opted out at time 0. This is the mirror image of the previous case. Now the default is characterized by the same cost for both types of terms and hence observing the cost does not yield to learning if the firm chooses the default term at time 0. Terms characterized by asymmetric learning that occurs only if the option chosen is the opt out are labelled *learning-from-opt-out terms*.[59]

Note that, although the different types of clauses arise with different values of the absolute costs c_L^D, c_H^D, c_L^O and c_H^O, our results will depend exclusively on the values of the relative cost c_H and c_L. Therefore, the comparisons that we will make below have to be understood as being ceteris paribus with respect to the relative costs, while the underlying absolute costs will typically be different, as should be clear from the classification above.

3. Determinants of contractual choice

As usual, we proceed backwards. In the last period, time 1, the firm is better off adopting the term that maximizes its payoff given the information available. Backing up to the earlier period, time 0, the firm will have to weigh two possibly opposing interests: maximize the expected time-0 payoff and improving the time-1 payoff by learning. Denote the payoff at time $t \in \{0,1\}$ when the firm adopts term $d \in \{D, O\}$ at time 0 as $\Pi_t^d = v^d - c^d$. The firm maximizes the following payoff

$$\max_{d\in\{D,O\}}\{\mathrm{E}[\Pi_0^d] + w\mathrm{E}[\Pi_1^d] - q^d s\} \tag{4}$$

That is, the firm maximizes the expected payoff it receives at time 0 (given that it chooses without knowing the types of consumers) plus the expected payoff it receives at time 1 (which accounts for the possibility to learn in the meantime and is discounted by a factor that captures changes in the market share of the firm), minus the expected cost of switching at time 1, where the probability of switching, q^d, will depend on the choice made at time 0. In the next sections, we will analyze the firm's decision problem for each of the four types of terms identified above.

4. Nonlearning terms

If the firm cannot learn from observing costs, the choice at time 1 must necessarily be identical to the choice made at time 0, because both choices are based on similarly

[59] The residual case in which $c_L^D = c_H^D$ and $c_L^O = c_H^O$ yields no learning and is uninteresting as it is a special case of the nonlearning terms discussed above.

incomplete information. Thus, we have $\mathrm{E}[\Pi_0] = \mathrm{E}[\Pi_1]$. Absent new information there is no reason to switch at time 1, thus $q^d = 0$. The firm's problem in (4) reduces to

$$\max_{d\in\{D,O\}}\{(1+w)\mathrm{E}[\Pi_1^d]\} \tag{5}$$

The firm's payoff is maximized by the default if $\mathrm{E}[\Pi_0^D] > \mathrm{E}[\Pi_0^O]$ and, vice versa by the opt out. Note that, from (3), $\mathrm{E}[\Pi_0^D] - \mathrm{E}[\Pi_0^O] = v - \mathrm{E}[c] = v - pc_H - (1-p)c_L$, which is greater than zero (and hence the default yields the larger payoff) if $p < p^*$ and lower than zero (and hence the opt out yields the larger payoff) if $p > p^*$, where:

$$p^* \equiv \frac{v - c_L}{c_H - c_L} \tag{6}$$

If $p = p^*$, the firm is indifferent between adopting the default and opting out. For ease of notation, here and in what follows we disregard this case. (Note that we found the same threshold in the simplified model presented in the text.) Lemma 1 summarizes these results.

Lemma 1. With nonlearning terms (N) the default term is adopted both at time 0 and at time 1 if and only if $p < p^*$, irrespective of w and s; p^* increases in v. The probability of switching at time 1 is $q^D = q^O = 0$.

5. Learning terms

If the firm can learn from observing costs irrespective of the term chosen at time 0, the choice at time 1 is made under complete information. If the firm learns that costs are low, it will adopt the default, while if it learns that they are high, it will opt out. The firm will switch away from the term chosen at time 0 with a positive probability. We make the simplifying assumption that s is small enough so that the presence of switching costs does not prevent switching at time 1 but only affects choices at time 0.[60]

The firm's time-0 expected payoff is $\mathrm{E}[\Pi_0^d] = v^d - pc_H^d - (1-p)c_L^d$ and depends on the term chosen in an intuitive way. The firm's time-1 expected payoff is $\mathrm{E}[\Pi_1^d] = p(v^O - c_H^O) + (1-p)(v^D - c_L^D)$ and does not depend on the term chosen at time 0. The reason is that the firm learns and adopts the optimal term irrespective of its time-0 choice: if costs are low it adopts the default while if costs are high it opts out. The probability of switching at time 1 is $q^D = p$ if the firm chose the default at time 0 and hence switches only if costs are discovered to be high, that is, with probability p. Likewise, if the firm chose the opt out, we have $q^O = 1 - p$.

Note that since $\mathrm{E}[\Pi_1^D] = \mathrm{E}[\Pi_1^O]$, the firm's maximization problem can be simplified to

$$\max_{d\in\{D,O\}}\{\mathrm{E}[\Pi_0^d] - q^d s\} \tag{7}$$

The firm's payoff is maximized by the default if $\mathrm{E}[\Pi_0^D] - q^D s > \mathrm{E}[\Pi_0^O] - q^O s$ and vice versa by the opt out. Note that, using (3), $\mathrm{E}[\Pi_0^D] - \mathrm{E}[\Pi_0^O] - (q^D - q^O)s = v - pc_H -$

[60] If switching costs were particularly high, some switches would not take place and we would partially revert back to the nonlearning case. In addition, depending on the costs, switching might be advantageous only in one direction, making the dynamics of learning asymmetric.

$(1-p)c_L-(2p-1)s$, which is greater than zero (and hence the default yields the larger payoff) if $p<p^{**}$ and lower than zero (and hence the opt out yields the larger payoff) if $p>p^{**}$, where:

$$p^{**}\equiv\frac{v-c_L+s}{c_H-c_L+2s} \tag{8}$$

and

$$\frac{\partial p^{**}}{\partial s}=\frac{c_H+c_L-2v}{(c_H-c_L+2s)^2} \tag{9}$$

The latter expression is positive if $v<\frac{c_H+c_L}{2}$ (which implies $p^*<\frac{1}{2}$), zero if $v=\frac{c_H+c_L}{2}$ (which implies $p^*=\frac{1}{2}$) and negative if $v>\frac{c_H+c_L}{2}$ (which implies $p^*>\frac{1}{2}$). To rationalize this result, consider that if $s=0$ then $p^{**}=p^*$: the firm's choice concerning a learning term is the same as the choice the firm would make with a nonlearning term. Absent switching costs, the firm chooses at time 0 the alternative giving the higher expected payoff and switches to the optimal term inevitably at time 1. (This is the case considered in the simplified model presented in the main text.)

If $s>0$, switching costs weigh favourably on the choice of the alternative that is most likely the optimal choice. If $p<\frac{1}{2}$, switching costs make the default option more attractive because there is a high chance of low costs and hence switching to the opt out is not very likely. If instead $p>\frac{1}{2}$, switching costs tend to make the opt out more attractive.

An interesting benchmark is $v=\frac{c_H+c_L}{2}$. In this case, we have $p^{**}=p^*=\frac{1}{2}$. Since the two alternative cost scenarios are equally likely, the prospect of facing switching costs does not make either choice more attractive. If $v<\frac{c_H+c_L}{2}$, we have that $p^{**}<p^*$:[61] the need to avoid switching costs makes the firm more inclined to choose the opt out, which gives a lower probability of switching compared to the nonlearning case. If instead If $v>\frac{c_H+c_L}{2}$, we have that $p^{**}>p^*$ for the opposite reason.[62]

Lemma 2. *With learning terms (L) the default term is adopted at time 0 if and only if $p<p^{**}$ and is adopted at time 1 if and only if $c=c_L$, independently of w; p^{**} increases in v. Ceteris paribus, we have that: (i) if $v<\frac{c_H+c_L}{2}$, then $p^{**}<p^*$ and p^{**} increases in s; (ii) if $v=\frac{c_H+c_L}{2}$, then $p^{**}=p^*$ and p^{**} does not change in s; and (iii) if $v>\frac{c_H+c_L}{2}$, then $p^{**}>p^*$ and p^{**} decreases in s.*

6. Learning-from-default terms

If the firm can learn from observing costs only if it chose the default at time 0, the choice at

[61] This is because p^{**} increases in v toward p^* in that region until it reaches p^* for $v=\frac{c_H+c_L}{2}$.

[62] This is because p^{**} increases in v away from p^* in that region.

time 1 is made under complete information only in the latter case. If instead the firm chose the opt out at time 0, it does not have any additional information at time 1 and hence it does not switch. We focus again on cases in which the switching costs are small enough so that switching always takes place when advantageous.

The firm's time-0 expected payoff is $\mathrm{E}[\Pi_0^d] = v^d - pc_H^d - (1-p)c_L^d$ as before. The firm's time-1 expected payoff is the same as above, $\mathrm{E}[\Pi_1^O] = v^O - pc_H^O - (1-p)c_L^O$, if it chose the opt out but becomes $\mathrm{E}[\Pi_1^D] = p(v^O - c_H^O) + (1-p)(v^D - c_L^D)$ if the firm chose the default and hence learned costs. The probability of switching at time 1 is $q^D = p$ if the firm choose the default at time 0 and $q^O = 0$ if it chose the opt out. We now need to consider the maximization problem in (4) in its complete form.

The firm's payoff is maximized by the default if $\mathrm{E}[\Pi_0^D] + w\mathrm{E}[\Pi_1^D] - ps > (1+w)\mathrm{E}[\Pi_0^O]$ and, vice versa by the opt out. Note that, using again (3), $\mathrm{E}[\Pi_0^D] + w\mathrm{E}[\Pi_1^D] - (1+w)\mathrm{E}[\Pi_0^O] - ps = v - pc_H - (1-p)c_L + w(1-p)(v-c_L) - ps$, which is greater than zero (and hence the default yields a larger payoff) if $p < \overline{p}$ and lower than zero (and hence the opt out yields a larger payoff) if $p > \overline{p}$, where:

$$\overline{p} \equiv \frac{(1+w)(v-c_L)}{c_H - c_L + w(v-c_L) + s} \tag{10}$$

with

$$\frac{\partial \overline{p}}{\partial s} = -\frac{(1+w)(v-c_L)}{(c_H - c_L + w(v-c_L) + s)^2} < 0 \tag{11}$$

and

$$\frac{\partial \overline{p}}{\partial w} = \frac{(v-c_L)(c_H - v + s)}{(c_H - c_L + w(v-c_L) + s)^2} > 0 \tag{12}$$

and finally

$$\frac{\partial \overline{p}}{\partial v} = \frac{(1+w)(c_H - c_L + s)}{(c_H - c_L + w(v-c_L) + s)^2} > 0 \tag{13}$$

The threshold below which the firm chooses the default term at time 0 increases—thereby making the choice of the default more likely—if switching at time 1 is less expensive, if learning is more important because profits at time 1 weigh more and if the relative value of the default increases.

Compared with the learning terms, the learning-from-default terms make firms choose the default more often, $\overline{p} > p^{**}$,[63] because, in addition to the usual considerations, the firm also learns by choosing the default. The same result obtains with respect to the nonlearning case, $\overline{p} > p^{*}$.[64]

[63] This is very easy to see by noticing that $\overline{p} > p^*$ if $w = 0$ and that $\overline{p}$ increases in w.

[64] Compared with the nonlearning terms, there are two countervailing factors at play. On the one hand the firms choose the default more often because they can learn (this is the positive effect of w); on the other hand, the choice of the default is discouraged by the presence of switching costs (this is the negative effect of s). On balance, however, the first effect prevails because if w is low than the firm does not switch (a case that we have not explicitly accounted for in the analysis) and hence there are no switching costs.

Lemma 3. With learning-from-default terms (D) the default term is adopted at time 0 if and only if $p < \overline{p}$ and is adopted at time 1 if and only if the firm choose the default at time 0 and $c = c_L$; $\overline{p}$ increases in w and in v but decreases in s. Ceteris paribus, we have that: (i) $\overline{p} > p^{**}$ and (ii) $\overline{p} > p^{*}$.

7. Learning-from-opt-out terms

This is the mirror image of the case discussed above: now the firm only learns if it chose the opt out at time 0. The firm's time-0 expected payoff is $\mathrm{E}[\Pi_0^d] = v^d - pc_H^d - (1-p)c_L^d$ as before. The firm's time-1 expected payoff is the same as above, $\mathrm{E}[\Pi_1^D] = v^D - pc_H^D - (1-p)c_L^D$, if it chose the default but becomes $\mathrm{E}[\Pi_1^O] = p(v^O - c_H^O) + (1-p)(v^D - c_L^D)$ if the firm chose the opt out. The probability of switching at time 1 is $q^D = 0$ if the firm choose the default at time 0 and $q^O = 1 - p$ if it chose the opt out.

The firm's payoff is maximized by the default if $(1+w)\mathrm{E}[\Pi_0^D] > \mathrm{E}[\Pi_0^O] + w\mathrm{E}[\Pi_1^O] - (1-p)s$ and, vice versa by the opt out. Note that, using again (3), $(1+w)\mathrm{E}[\Pi_0^D] - \mathrm{E}[\Pi_0^O] - w\mathrm{E}[\Pi_1^O] + (1-p)s = v - pc_H - (1-p)c_L - wp(c_H - v) + (1-p)s$, which is greater than zero (and hence the default yields a larger payoff) if $p < \underline{p}$ and lower than zero (and hence the opt out yields a larger payoff) if $p > \underline{p}$, where:

$$\underline{p} \equiv \frac{v - c_L + s}{c_H - c_L + w(c_H - v) + s} \tag{14}$$

with

$$\frac{\partial \underline{p}}{\partial s} = \frac{(w+1)(c_H - v)}{(c_H - c_L + w(c_H - v) - s)^2} > 0 \tag{15}$$

and

$$\frac{\partial \underline{p}}{\partial w} = -\frac{(v - c_L - s)(c_H - v)}{(c_H - c_L + w(c_H - v) + s)^2} < 0 \tag{16}$$

and finally

$$\frac{\partial \underline{p}}{\partial v} = \frac{c_H - c_L + w(c_H - c_L - s) + s}{(c_H - c_L + w(c_H - v) + s)^2} > 0 \tag{17}$$

The threshold above which the firm chooses the opt out term at time 0 decreases—thereby making the choice of the opt out more likely—if switching at time 1 is less expensive, if learning is more important because profits at time 1 weigh more and if the relative value of the default decreases.

Compared with the learning terms, the learning-from-default terms make firm choose the default more often, $\underline{p} < p^{**}$,[65] because, in addition to the usual considerations, the firm also learns by choosing the opt out. The same result obtains with respect to the nonlearning case, $\underline{p} < p^{*}$.[66]

[65] This is very easy to see by noticing that $\overline{p} > p^{*}$ if $w = 0$ and that $\overline{p}$ increases in w.
[66] The reason is similar to what observed for learning-from-default terms.

Lemma 4. *With learning-from-opt-out terms (O) the default term is adopted at time 0 if and only if $p < \underline{p}$ and is adopted at time 1 if the firm choose the default at time 0 or if the firm choose the opt out at time 0 and $c = c_L$; $\underline{p}$ increases in s and in v but decreases in w. Ceteris paribus, we have that: (i) $\underline{p} < p^{**}$ and (ii) $\underline{p} < p^{*}$.*

8. Summary table for the comparative statics results

Table 5 summarizes the comparative statics results from the lemmata above.

Table 5. Comparative statics

	Terms			
	Nonlearning	Learning	Learning-from-default	Learning-from-opt-out
Changes in v	$\frac{\partial p^*}{\partial v} > 0$	$\frac{\partial p^{**}}{\partial v} > 0$	$\frac{\partial \overline{p}}{\partial v} > 0$	$\frac{\partial \underline{p}}{\partial v} > 0$
Changes in s	$\frac{\partial p^*}{\partial s} = 0$	$\frac{\partial p^{**}}{\partial s} =?$	$\frac{\partial \overline{p}}{\partial s} < 0$	$\frac{\partial \underline{p}}{\partial s} > 0$
Changes in w	$\frac{\partial p^*}{\partial w} = 0$	$\frac{\partial p^{**}}{\partial w} = 0$	$\frac{\partial \overline{p}}{\partial w} > 0$	$\frac{\partial \underline{p}}{\partial w} < 0$

9. Switching rates

Switching rates depend on two factors: the probability of switching conditional on having made a particular contractual choice at time 0 and the base rate at which choices at time 0 are made, which in turn exerts a selection effect on switching at time 1. Consider the general density function $f(p)$ on $[0,1]$ for the probability of having high costs p, where p can be interpreted as capturing all factors affecting costs for a particular firm, so that different firms might be characterized by different levels of p and $f(p)$ is the density function characterizing the population of firms.

We are interested in calculating the ex ante probability to observe switching at time 1 for a given population of firms. The following lemma presents self-evident results.

Lemma 5. *The ex ante probability of switches is equal to: 0 for nonlearning terms; $\int_0^{p^*} pf(p)dp + \int_{p^*}^1 (1-p)f(p)dp$ for learning terms, where the first addendum is the probability of switching from default to opt-out and the second addendum is the reverse probability of switching from opt-out to default; $\int_0^{\overline{p}} pf(p)dp$ for learning-from-default terms where switches are unidirectional from default to opt out; and, finally, $\int_{\underline{p}}^1 (1-p)f(p)dp$ for learning-from-opt-out terms and switches where unidirectional from opt out to default.*

Assuming a uniform distribution for p provides a quantitative example.

Corollary 1. With a uniform distribution, the ex ante probability of switches is equal to: 0 for nonlearning terms; $\frac{(p^*)^2}{2} + \frac{(1-p^*)^2}{2}$ *for learning terms;* $\frac{\overline{p}^2}{2}$ *for learning-from-default terms; and, finally,* $\frac{\left(1-\underline{p}\right)^2}{2}$ *for learning-from-opt-out terms.*

B. Tables

Table 6. Company, Product, Market, and Contract Characteristics

	Obs	Mean	SD	Min	Median	Max
Panel A. Company Characteristics						
Revenue 2003 ($000)	259	287,499	2,490,751	30	1700	36,800,000
Revenue 2010 ($000)	259	539,091	4,225,384	90	2200	60,400,000
Change Revenue ($)	254	256,679	1,917,968	-723,200	111.5	23,600,000
Change Revenue (%)	254	226	627	-90	24.08	5000
Public 2003	264	0.11	0.32	0	0	1
Public 2010	264	0.14	0.35	0	0	1
Age 2003 (Yrs)	264	13.62	8.01	0	13	68
Age 2010 (Yrs)	264	20.62	8.01	7	20	75
Lawyers	118	0.74	0.44	0	1	1
Pro-Consumer State	264	0.32	0.61	-1	0	1
Panel B. Product and Market Characteristics						
Trial 2003	264	0.73	0.45	0	1	1
Trial 2010	264	0.77	0.42	0	1	1
Median Price 2003 ($)	264	812	1,310	14.99	360	12,000
Median Price 2010 ($)	256	841	1,686	8.99	350	20,995
Consumer Product	264	0.36	0.48	0	0	1
Multi-User License	264	0.08	0.28	0	0	1
Developer License	264	0.08	0.27	0	0	1
H-H Index	236	0.37	0.24	.065	.30	1
Panel C. Contract Characteristics						
Any Terms Changed	264	0.39	0.49	0	0	1
Number of Words 2003	264	1,517	1,365	33	1,152	8,406
Number of Words 2010	262	1,938	2,077	106	1,354	13,416

Table 7. EULA Terms and Bias: 2003 vs. 2010

EULA terms are classified into 32 common terms that allocate rights and risks between buyers and sellers across seven categories of related terms, according to the degree the terms either match the default rules of UCC Article 2 (Adoption of Default = 0) or deviate from them (Opt-out= 1). "Learning Category" refers to the type and modality that allows sellers to learn from a term. Terms allow for symmetric learning, denoted S (L), when learning either happens regardless of the modality of the term, and S (N) when learning never happens regardless of the modality of the term. Some terms allow for asymmetric learning, allowing sellers to learn as long as the modality adopted enables learning. Terms that enable learning when the seller adopts the default rule but not otherwise are denoted A (D) (i.e., asymmetric learning by adopting the default). Terms that enable learning when the seller opts out of the default are denoted A (O) (i.e., asymmetric learning by opting out of the default). The table reports the mean opt-out of UCC Article 2 default in 2003 and 2010, as well as the mean change and statistical significance. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Learning Category	Category and Term	Adoption of Default=0 Opt-out=1	Mean 2010 (SD)	Mean 2003 (SD)	Mean Change (SE)
	Acceptance	1 = yes 0 = no	**0.458** **(0.499)**	**0.470** **(0.500)**	**0.011** **(0.022)**
S (N)	Does license alert consumer that product can be returned if she declines terms?				
	Modification and Termination		**0.227** **(0.539)**	**0.167** **(0.439)**	**0.061***** **(0.021)**
S (N)	Are license's terms subject to change?	0 = no 1 = yes	0.106 (0.309)	0.076 (0.265)	0.030** (0.012)
S (L)	Does license allow licensor to disable the software remotely if licensee breaches any EULA terms, according to licensor?	0 = no 1 = yes	0.121 (0.327)	0.091 (0.288)	0.030** (0.013)
	Scope		**1.792** **(1.169)**	**1.659** **(1.162)**	**0.133***** **(0.046)**
S (N)	Does definition of "licensed software" include regular updates such as enhancements, versions, releases, etc.?	1 = yes 0 = no or no mention	0.170 (0.377)	0.136 (0.344)	0.034** (0.015)
S (N)	Can licensee alter/modify the program?	0 = yes or no mention 1 = no	0.640 (0.481)	0.598 (0.491)	0.042*** (0.015)
A (D)	Can licensee create derivative works?	0 = largely unrestricted or no mention	0.379	0.352	0.027*

		1 = strict prohibition, derivative works owned by licensor, or need permission of licensor	(0.486)	(0.479)	(0.015)
A (D)	Does license prohibit reverse engineering of the software?	0 = no or no mention 1 = yes	0.716 (0.452)	0.663 (0.474)	0.053*** (0.017)
S (N)	Are there license grant restrictions?	0 = no or no mention 1 = yes (e.g., for business tgbnhoriented products, "for business purposes" or "internal purposes only" language; for consumer-oriented products, restrictions on commercial use)	0.227 (0.420)	0.182 (0.386)	0.045*** (0.018)
	Information Collection		**0.117** **(0.367)**	**0.061** **(0.269)**	**0.057*** **(0.017)**
S (L)	Does license allow licensor to collect and /or distribute licensee's personally identifiable information?	0 = no or no mention 1 = yes	0.102 (0.304)	0.053 (0.225)	0.049*** (0.014)
A (O)	Does license allow licensor to install software that will track licensee's activity?	0 = no or no mention 1= yes	0.015 (0.122)	0.008 (0.087)	0.008 (0.005)
	Transfer		**1.466** **(0.584)**	**1.394** **(0.595)**	**0.072*** **(0.021)**
S (N)	Are there limitations on transfer?	0 = no or no mention 1 = some or full restrictions (licensee cannot assign, transfer, lease, sublicense, distribute, etc.; or, needs written consent of licensor)	0.955 (0.209)	0.943 (0.232)	0.011* (0.007)
S (N)	Can licensee transfer the software to an end user who accepts the license terms without licensor's prior permission?	0 = yes or no mention 1 = no	0.511 (0.501)	0.451 (0.499)	0.061*** (0.017)
	Warranties and Disclaimers		**0.871** **(0.994)**	**0.875** **(0.973)**	**0.004** **(0.028)**
A (O)	Are there express warranties?	1 = yes 0 = no	0.042 (0.200)	0.042 (0.200)	0.000 (0.005)

A (O)	Is there a limited warranty stating that software is free from defects in materials and workmanship or that the software will work according manual specifications in force for a limited period?	1 = yes 0 = no	0.311 (0.464)	0.295 (0.457)	0.015 (0.017)
A (O)	Is there a limited warranty stating that the media of software distribution and documentation are free from defects in force for a limited period?	1 = yes 0 = no	0.280 (0.450)	0.269 (0.444)	0.011 (0.017)
S (N)	Is the disclaimer in caps, bold, or otherwise conspicuously presented?	0 = yes or no disclaimers appear 1 = no	0.231 (0.422)	0.261 (0.440)	0.030** (0.013)
A (D)	Disclaims IWM and IWFPP or contains "AS IS" language?	0 = no 1 = yes	0.913 (0.283)	0.890 (0.313)	0.023** (0.009)
A (D)	Disclaims warranty that software will not infringe on third parties' intellectual property rights?	0 = no 1 = yes	0.360 (0.481)	0.330 (0.471)	0.030** (0.014)
	Limitations on Liability		**2.413** **(1.221)**	**2.273** **(1.187)**	**0.140*** **(0.047)**
A (D)	Who bears the risk of loss?	0 = licensor, for losses caused by factors under licensor's control, or no mention 1 = licensee	0.167 (0.373)	0.152 (0.359)	0.015 (0.012)
A (D)	Who bears the performance risk?	0 = licensor (for causes under licensor's control), or no mention, or licensee (for uses expressly forbidden by licensor) 1 = licensee (language "licensee assumes responsibility of choice of product and functions," etc)	0.299 (0.459)	0.277 (0.448)	0.023 (0.015)
A (D)	Disclaims consequential, incidental, special, or foreseeable damages?	0 = no or no mention 1 = yes	0.924 (0.265)	0.902 (0.299)	0.023** (0.009)
A (D)	Are damages disclaimed under all theories of liability (contract, tort, strict liability)?	0 = no or no mention 1 = yes	0.299 (0.459)	0.273 (0.446)	0.027* (0.015)
A (D)	What is the limitation on damages?	0 = no mention or cap on damages greater than purchase price 1 = cap on damages less than or equal to purchase	0.553 (0.498)	0.519 (0.501)	0.034* (0.019)

		price			
A (D)	Is there an indemnification term?	0 = no, no mention, or twoway indemnification 1 = indemnification by licensee	0.170 (0.377)	0.152 (0.359)	0.019 (0.015)
	Maintenance and Support	1 = yes 0 = no or no mention	**0.667** **(0.472)**	**0.663** **(0.474)**	**0.004** **(0.014)**
A (O)	Does base price include M&S for 31 days or more?				
	Conflict Resolution		**0.341** **(0.513)**	**0.284** **(0.476)**	**0.057***** **(0.019)**
A (D)	Forum specified?	0 = court, choice of licensee, or no mention 1 = specific court or mandatory arbitration	0.322 (0.468)	0.273 (0.446)	0.049*** (0.017)
S (L)	Law specified?	0 = same as forum or no mention 1 = yes and different from forum	0.011 (0.106)	0.008 (0.087)	0.004 (0.004)
S (L)	Who pays licensor's attorney fees?	0 = paid by losing party or no mention 1 = paid by licensee	0.008 (0.087)	0.004 (0.062)	0.004 (0.004)
	Third Parties		**0.216** **(0.574)**	**0.098** **(0.346)**	**0.117***** **(0.028)**
S (N)	Does license require licensee agree to third party licenses or terms?	0 = no or no mention 1 = yes	0.121 (0.327)	0.064 (0.246)	0.057*** (0.015)
A (D)	Does license disclaim licensor's liability for any included third party software?	0 = no or no mention 1 = yes	0.080 (0.271)	0.034 (0.182)	0.045*** (0.015)
S (N)	Does license allow licensor or third parties to install additional software?	0 = no or no mention 1 = yes	0.015 (0.122)	0.000 (0.000)	0.015** (0.008)
	Consumer Protection	1= yes, contract informs consumer about state law rights they may have	**0.473** **(0.500)**	**0.417** **(0.494)**	**0.057***** **(0.017)**
S (N)	Does license inform licensee of statutory rights?	0= no or no mention			
	Total Mean Change				**0.583***** **(0.128)**

Table 8. Learning and Changing Terms

Fraction of terms that change between 2003 and 2010 depending on whether their 2003 values are at the default or, for asymmetric terms, at the learning value. In Panel A, for example, 29.4% of terms were at opt-out values in both 2003 and 2010 and 1.4% were at a opt-out value in 2003 and changed to a default value by 2010. The probability of a change for a term that was at a opt-out value in 2003 is 0.045 (0.014/0.308), while the probability of a change for a term that was at the default in 2003 is 0.056 (0.039/0.692), which is a statistically significant difference of -0.011. Asymmetric terms can also be at a learning or nonlearning value. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel A. All Terms (32 terms; 8448 EULA-term observations)

		2010 term				
	(Fractions)	opt-out	default	total		
2003 term	opt-out	0.294	0.014	0.308	Prob(change | 2003 at opt-out)	0.045
	default	0.039	0.653	0.692	Prob(change | 2003 at default)	0.056
	total	0.333	0.667	1	difference	-0.011**

Panel B. Symmetric Learning Terms (15 terms; 3,696 EULA-term observations)

		2010 term				
		opt-out	default	total		
2003 term	opt-out	0.238	0.013	0.251	Prob(change | 2003 at opt-out)	0.052
	default	0.039	0.711	0.750	Prob(change | 2003 at default)	0.052

total	0.277	0.724	1	difference	0

Panel C. Asymmetric Learning Terms (17 terms; 4,752 policy-term observations)

		2010 term				
		opt-out	default	total		
2003 term	opt-out	0.344	0.015	0.359	Prob(change | 2003 at opt-out)	0.042
	default	0.039	0.603	0.642	Prob(change | 2003 at default)	0.061
	total	0.383	0.618	1	difference	-0.019**

		2010 term				
		learning	nonlearning	total		
2003 term	learning	0.461	0.036	0.497	Prob(change | 2003 at learning)	0.072
	nonlearning	0.017	0.485	0.502	Prob(change | 2003 at nonlearning)	0.034
	total	0.478	0.521	1	difference	0.038***

Table 9. Learning and Changing Terms: Robustness

The sample is asymmetric terms only in 264 contracts. Least squares regressions where the dependent variable is a 0-1 indicator that the term changed between 2003 and 2010. Learning means that the term was set at a learning value in 2003. Default means that the term was set at the default in 2003. Standard errors in parentheses are clustered by firm. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) Change	(2) Change	(3) Change	(4) Change	(5) Change	(6) Change
Learning	0.0392***	0.0402***		0.0394***	0.0401***	0.0420**
	(0.00920)	(0.00801)		(0.00815)	(0.00984)	(0.0145)
Default			0.0187**	0.00204	0.0003	0.0138
			(0.00818)	(0.00831)	(0.00958)	(0.0152)
Multi-User License					-0.0417***	-0.0778***
					(0.0147)	(0.0173)
Developer License					-0.0104	-0.00121
					(0.0280)	(0.0328)
Ln Price					0.0103	0.0338**
					(0.00627)	(0.0128)
Change Ln Price					0.0497**	0.0647
					(0.0223)	(0.0404)
Consumer Product					0.00400	0.0376
					(0.0159)	(0.0265)
Ln Revenue					0.00393	-0.000247
					(0.00348)	(0.00564)
Change Ln Revenue					0.0219***	0.0290***
					(0.00662)	(0.0100)
Ln Age					0.00122	0.0142
					(0.0117)	(0.0214)
Lawyers						0.0611*
						(0.0329)
Pro-Consumer State					-0.00448	-0.0298
					(0.0110)	(0.0198)
H-H Index					0.0279	0.0217
					(0.0247)	(0.0377)

Constant	0.0337*** (0.00533)	0.0332*** (0.00399)	0.0412*** (0.00525)	0.0323*** (0.00588)	-0.0757 (0.0507)	-0.246** (0.0996)
Firm Fixed Effects	No	Yes	Yes	Yes	No	No
Observations	4,488	4,488	4,488	4,488	3,791	1,751
Adjusted R^2	0.007	0.160	0.154	0.160	0.026	0.050

Table 10. Asymmetric Learning by Default vs. Opt-out

Rate of learning values chosen for asymmetric terms, where asymmetric terms are broken down into those where learning is by adoption of the default rules of UCC and those where learning is by opting-out of such default rules.

Panel A. Asymmetric Learning Terms -- Learning from Defaults (12 terms; 3,168 EULA-term observations)

		2010 term		
		learning	nonlearning	total
2003 term	learning	0.555	0.044	0.599
	nonlearning	0.013	0.388	0.401
	total	0.568	0.432	1

Prob(change | 2003 at learning)	0.073
Prob(change | 2003 at nonlearning)	0.032
difference	0.041***

Panel B. Asymmetric Learning Terms -- Learning from Opt-out (5 terms; 1,320 EULA-term observations)

		2010 term		
		learning	nonlearning	total
2003 term	learning	0.237	0.018	0.255
	nonlearning	0.026	0.719	0.745
	total	0.263	0.737	1

Prob(change | 2003 at learning)	0.071
Prob(change | 2003 at nonlearning)	0.035
difference	0.036***

C. Empirics

Term #	Learning Category	Term (t)	Classification Rationale	Learning (0=no; 1=yes)
		Acceptance		
x_1	S (N)	Does license alert consumer that product can be returned if she declines terms? 1=yes; 0=no	Pure information given to consumer; no feedback.	0
		Modification and Termination		
x_2	S (N)	Are license's terms subject to change? 0=no; 1=yes	Pure information given to consumer; no feedback.	0
x_3	S (L)	Does license allow licensor to disable the software if licensee breaches any EULA terms, according to licensor? 0=no; -1=yes	Clause makes enforcement easier. Feedback occurs in either case.	1
		Scope		
x_4	S (N)	Does definition of "licensed software" include updates, enhancements, versions, releases, patches, etc.? 1=yes;0=no mention/no	Pure information given to consumer; no feedback.	0
x_5	S (N)	Can licensee alter/modify the program? 0=yes or no mention; -=no	Product feature; no feedback in either case.	0
x_6	A (D)	Can licensee create derivative works? 0=largely unrestricted or no mention; 1= strict prohibition, derivative works owned by licensor, or need permission of licensor	Seller does not know value of derivative work for consumers. Prohibiting it hinders learning, while allowing it possibly also allows the seller to learn.	1 if t = 0
x_7	A (D)	Does license allow reverse engineering of the software? 0=yes 1=no	Seller might not know whether reverse engineering is possible, cost-effective and damaging for seller. Prohibiting it impairs learning.	1 if t = 0

x_8	S (N)	Are there restrictions on use? 0=no or no mention; 1=yes (e.g., for business-oriented products, "for business purposes" or "internal purposes only", or "within the same building" language; for consumer-oriented products, restrictions on commercial use)	Product feature, no feedback in either case.	0
		Information Collection		
x_9	S (L)	Does license allow licensor to collect and /or distribute licensee's information? 0=no/no mention 1=yes	Product feature. Some feedback in either case. Seller will learn in the future whether collecting information gives him a competitive advantage or not-collecting information makes his product more appealing to consumers.	1
x_{10}	A (O)	Does license allow licensor to install software that will track licensee's activity? 0=no or no mention 1=yes	Seller learns the value of the clause of if allows to track activity (for enforcement purposes).	1 if t = 1
		Transfer		
x_{11}	S (N)	Are there limitations on transfer? 0=no or no mention; 1=some or full restrictions (licensee cannot assign, transfer, lease, sublicense, distribute, etc.; or, needs written consent of licensor)	Product feature; no feedback in either case.	0
x_{12}	S (N)	Can Licensee transfer the software if end user accepts license terms? 0=yes or no mention; 1=no	Product feature; no feedback in either case.	0
		Warranties and Disclaimers		
x_{13}	A (O)	Are Express Warranties made? 1=yes; 0=no	Seller learns the value of the warranty only if warranty is included.	1 if t = 1

x_{14}	A (O)	Is there a limited warranty (e.g. stating that software is free from defects in materials and workmanship or that it will perform substantially in accordance to material documentation) in force for 31 days or more? 1=yes; 0=no	Seller learns the value of the warranty only if warranty is included.	1 if t = 1
x_{15}	A (O)	Is there a limited warranty stating that the media of software distribution and documentation are free from defects in force for 31 days or more? 1=yes; 0=no (RECORD AS #)	Seller learns the value of the warranty only if warranty is included.	1 if t = 1
x_{16}	S (N)	Is the disclaimer in caps? 0=yes or no disclaimers appear; 1=no	Pure information given to consumer; no feedback.	0
x_{17}	A (D)	Disclaims IWM, EW, and IWFPP or contains "AS IS" language? 0=no; 1=yes	Seller learns the value of the warranty only if warranty is included.	1 if t = 0
x_{18}	A (D)	Disclaims warranty that software will not infringe on third parties' intellectual property rights? 0=no ;1=yes	Seller learns the value of the warranty only if warranty is included.	1 if t = 0
		Limitations on Liability		
x_{19}	A (D)	Who bears the risk of loss? 0=licensor, for losses caused by factors under licensor's control, or no mention; 1=licensee	Seller learns exposure to liability only if bears the loss.	1 if t = 0
x_{20}	A (D)	Who bears the performance risk? 0=licensor, for causes under licensor's control, or no mention, or licensee, for uses expressly forbidden by licensor; 1=licensee (language "licensee assumes responsibility of choice of product and functions, etc.)	Seller learns exposure to liability only if bears the loss.	1 if t = 0
x_{21}	A (D)	Disclaims incidental, consequential and special damages? 0=no or no mention; 1=yes	Seller learns exposure to liability only if there is no disclaimer.	1 if t = 0

x_{22}	A (D)	Are damages waived under all theories of liability (contract, tort, strict liability)? 0=no; 1=yes	Seller learns exposure to liability only if there is no waiver.	1 if t = 0
x_{23}	A (D)	What is the limitation on damages? 0=no mention or cap on damages greater than purchase price; 1=cap on damages less than or equal to purchase price	Seller learns exposure to liability only if there is no limitation.	1 if t = 0
x_{24}	A (D)	Is there an indemnification clause? 0=no, no mention, or two-way indemnification; 1=indemnification by licensee	Sellers from exposure by being liable for any infringement.	1 if t = 0
		Maintenance and Support		
x_{25}	A (O)	Does base price include M&S for 31 days or more?1=yes; 0=no or no mention	Seller learns only if M&S included.	1 if t = 1
		Conflict Resolution		
x_{26}	A (D)	Forum specified? 0=choice of licensee or no mention; 1=specific court or mandatory arbitration	Seller learns risks of non-specified forum only if no choice of forum is made.	1 if t = 0
x_{27}	S (L)	Law specified? 0=same as forum or no mention; 1=yes and different from forum	Seller learns risks of non-specified law only if no choice of law is made.	1
x_{28}	S (L)	Who pays licensor's attorney fees? 0= paid by losing party or no mention; 1=paid by licensee	If there is litigation, seller learns anyway the costs.	1
		Third Parties		
x_{29}	S (N)	Does license require licensee agree to third party licenses or terms? 0=no; 1=yes	Pure information given to consumer; no feedback.	0
x_{30}	A (D)	Does license disclaim licensor's liability for any included third party software? 0=no - 1=yes	Seller learns exposure to liability only if there is no disclaimer.	1 if t = 0

x_{31}	S (N)	Does license allow licensor or third parties to install additional software? 0=no; 1=yes	Product feature; no feedback in either case.	0
		Consumer Protection		
x_{32}	S (N)	Does license inform licensee of statutory rights? 0=no; 1=yes	pure information given to consumer; no feedback.	0