Learning in Standard Form Contracts: Theory and Evidence

Dari-Mattiacci, G.; Marotta-Wurgler, F.

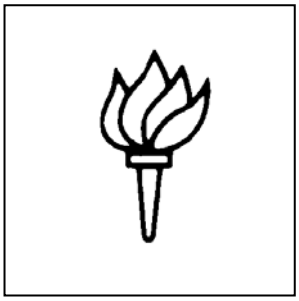
DOI
10.2139/ssrn.3133791

Publication date
2018

Document Version
Final published version

NEW YORK UNIVERSITY SCHOOL OF LAW

NYU Center for Law, Economics and Organization



Learning in Standard Form Contracts: Theory and Evidence

Giuseppe Dari-Mattiacci and Florencia Marotta-Wurgler

March 2018

LAW & ECONOMICS RESEARCH PAPER SERIES
WORKING PAPER NO. 18-11

The Center for Law and Economic Studies
Columbia University School of Law
435 West 116th Street
New York, NY 10027-7201

(212) 854-3739

Learning in Standard Form Contracts: Theory and Evidence

Giuseppe Dari-Mattiacci and Florencia Marotta-Wurgler

Working Paper No. 593

March 2018

LEARNING IN STANDARD FORM CONTRACTS: THEORY AND EVIDENCE

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

March 3, 2018

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 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, Marco Fabbri, Ron Gilson, Scott Hemphill, Gerard Hertig, Keith Hylton, Henrik Lando, Mark Ramseyer, Alan Schwartz, Robert Scott, David Webber, Kathryn Zeiler, Jonathan Zitnyck, and the participants in the Conference on Contractual Black Holes (April 7th and 8th, 2017) at Duke Law School, the annual meeting of the European Association of Law and Economics (September 14th – 16th, 2017) in London, and seminars at NYU Law School, ETH Zurich, the Institute for Information Law at the University of Amsterdam, and the faculty workshop at Boston University School of Law for helpful comments and suggestions.

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 not a general phenomenon. Stickiness is the flipside of change: Some terms seem to be resistant to change, like *pari passu* clauses in sovereign bond agreements,[4] while other get updated very quickly. Consider the rise of class action waivers in privacy policies in the wake of *AT&T Mobility LLC v. Concepcion*.[5] What explains the dynamic evolution of standard form contracts? How do firms decide which terms to change and which to retain? In this paper, we propose a novel account of stickiness and change in standard form contracts based on *experiential* learning by firms. That is, there are certain terms whose relative value is best ascertained through the feedback their use generates. 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.

At the outset, contract drafters may be uncertain about the net value of a particular term. Over time, drafters learn about the relative costs and benefits associated with such terms, leading them to drop some while adding or revising others. Firms can learn in many ways, such as learning

---

[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] Gulati & Scott, *id* (explaining contracting parties' reluctance to revise *pari passu* clauses in sovereign bond agreements after unfavorable interpretations by courts).

[5] Florencia Marotta-Wurgler and Daniel Svirsky, *Do FTC Privacy Enforcement Actions Matter? Compliance Before and After US-EU Safe Harbor Agreement* (NYU Working Paper, 2018). But see Peter B. Rutledge & Christopher R. Drahozal, *Sticky" Arbitration Clauses? The Use of Arbitration Clauses After Concepcion and Amex*, 67 VAND. L. REV. 955, 987-1001 (2014) (failing to find wholesale adoption of arbitration clauses post-Concepcion in franchise agreements).

about the terms offered by competitors, cases litigated in court, technological innovations, and news reports, among others.[6] These learning channels, while important, may depend on the types of terms 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. This is the focus of our paper.

Consider a default implied warranty. A firm may contemplate including a disclaimer of implied warranties in its 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. At the moment the firm makes its choice, the extent to which the costs outweigh the benefits of the warranty to consumers may be uncertain. This might depend on the frequency of product breakdown, the types of consumer losses, as well as the frequency at which consumers claim a remedy. 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. Opting for the disclaimer 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.

Not all terms are created equal. 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. Learning can occur in many ways, but not necessarily from experience in using the term. 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 implied warranty example illustrated above. Adoption of the term modality 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 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

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

costs will allow the firm to revise the term later. If the firm opts out of the default by including a disclaimer of implied warranties in its contract, the firm protects itself against future liability but also forgoes the option to learn and will be less likely to revise the term at a later stage.

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

Stickiness, in our framework of analysis, is the result of the inability to learn. That is, the dynamic of contractual innovation that results from experiential learning produces stickiness for those terms and term modalities where experiential learning doesn't occur. In the case of asymmetric-learning terms, however, stickiness occurs only when the initial balance of costs and benefits induces firms to opt for the non-learning modality. This is especially likely to occur when the non-learning modality of a term is particularly beneficial, such as when it corresponds to a default rule. Default rules, which are typically very attractive for firms and consumers alike, since they are known and easily adoptable, might for this very reason also be more prone to become sticky over time.

Please note that we do not propose that experiential learning is the *only* channel that can lead firms to revise their terms. There are multiple drivers of stickiness and innovation, several of which have been identified in the literature. We posit, however, that these drivers are uncorrelated with terms' inherent potential to allow firms to learn from experience, thus allowing us to track experiential learning empirically.

We classify 32 terms into categories that reflect their opportunity to learn in each of a sample of End User License Agreements (EULAs) used by 246 firms that tracks contractual changes from 2003 to 2010. Specifically, we first determine whether each term is associated with symmetric versus asymmetric learning. For the latter, we identify whether the learning modality of the term occurs at default or when the firm opts out the default. Consistent with previous studies, we find that default rules tend to be sticky. Firms are initially more likely to adopt terms at their default modality, yet terms set at the default are more likely to be revised at a later stage regardless of their learning modality, suggesting that such defaults might have been chosen sub-optimally. Dividing terms into their respective learning categories yields a strikingly different result. Consistent with our hypothesis, we find that symmetric learning terms are equally likely to be revised at a later period irrespective of the modality selected at an initial period. In contrast, we find that asymmetric terms are several times more likely to be revised at a later stage if the firms chose to include the term in its learning modality at an initial period. The results are statistically

significant and robust.

The paper proceeds as follows. In Section I, we review the literature on standard form contracts and contractual innovation. In Section II, we propose a simple theory of experiential learning by firms, which we describe in a stylized model in Section III. 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 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. The Appendix contains all the details of the empirical analysis and a generalization of the theoretical model.

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

The benefits of standardization are well understood and have been explored extensively in the literature. As terms become increasingly common and well-known, they are easier for contracting parties and courts to interpret. In addition, common, well-known terms confer various spillover effects, such as lower reading costs, increased certainty of legal interpretation, and reduced litigation risk.[7] However, the benefits created by standardization, such as learning and network benefits, may stand in the way of change, reducing contracting parties' incentives to revise familiar terms, even when doing so might be efficient.[8] Markets that experience higher network benefits (arising from firms' simultaneous adoption of a term) 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.[9] Weak property rights in contractual innovations are likely to further reduce incentives to innovate.[10]

---

[7] 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 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).

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

[9] *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).

[10] *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

Default rules can also be an important source of contractual stickiness. 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 be reluctant to deviate from them when they perceive that opting out, even if value generating, might signal negative information.[12] Finally, the status quo bias created by default rules can further create inertia that makes switching out difficult.[13] 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 unfavorably by a court, among others.

Despite the obstacles, change and innovation can still happen. Large repeat players, 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.[14] 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.[15] There are some accounts that posit that the opportunity to experiment can result

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 7, at 503 (arguing that because innovations in standard terms are public goods, the absence of intellectual property rights diminishes the incentive to innovate).

[11] Goetz & Scott, *supra* note 10 (discussing how state regulation of contract terms creates barriers to innovation). See also See Alan Schwartz & Robert E. Scott, Contract Theory and the Limits of Contract Law, 113 YALE L.J. 541, 598-601 (2003); Alan Schwartz & Robert E. Scott, The Political Economy of Private Legislatures, 143 U. PA. L. REV. 595, 597 (1995) (noting the inefficiencies likely associated with drafting default rules).

[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).

[13] 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).

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

[15] *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

in learning and change.[16] Change and innovation can also be spurred by "shocks," such as new laws, changes in legal interpretations of terms, or technological advances.

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. Terms related to product attributes tend to fit this category. Consider, for example, a term in a EULA limiting the number of devices where the software can be installed. 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, not one directly driven from experience with a particular version of the term.

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.[17] Firms can also learn from interacting with consumers through employees and customer service channels, where learning isn't necessarily mediated by the contract terms themselves.[18] 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.[19] In addition, law firms can transmit knowledge to their seller-clients, who can then revise their terms accordingly. More recently, legal service firms like Bloomberg Law and Legal Zoom have begun offering standard terms for a

---

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).

[16] 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.

[17] See Mitu Gulati & Stephen J. Choi, Innovation in Boilerplate Contracts: An Empirical Examination of Sovereign Bonds, 53 *Emory Law Journal* 930-996 (2004) (positing that firms would revise terms after being interpreted unfavorably by courts); 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) (finding that firms are more likely to include terms in EULAs that become increasingly enforced by courts); Daniel Schwarcz, *The Role of Courts in the Evolution of Form Contracts: An Insurance Case Study* (working paper, 2018) (examining whether insurance contract terms change after changes in court interpretation).

[18] *See* David Hoffman, *Relational Contracts of Adhesion*, 85 U. Chi. L. Rev. (forthcoming 2018).

[19] *See* Goetz & Scott, supra note 10; 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.

number of transactions, allowing firms to innovate at a relatively low cost.[20] Blogs and internet forums are an additional source of free advice regarding terms, in addition to word of mouth and several other mechanisms. In sum, the type of learning that might affect a firm's contractual choices can occur in many ways. Some of these mechanisms have been identified empirically.

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.[21] 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.[22] 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.[23] In the insurance context, Daniel Schwarcz found evidence of innovation in insurance contracts away from the ISO form, which is the standard insurance document.[24] 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 about 20% of such change can be attributed to new terms.[25] Finally, in a study joint with Robert Taylor, one of us found evidence of terms changing in reaction to litigated cases and changes in the enforceability of terms.[26]

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

---

[20] See Triantis, George G., Improving Contract Quality: Modularity, Technology, and Innovation in Contract Design, Stanford Journal of Law, Business, and Finance, Vol. 18, No. 2 (2013) (arguing that the sluggishness to innovation associated with standardized terms can be avoided by exploiting the inherent modularity in standardized agreements

[21] *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).

[22] 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).

[23] GULATI & SCOTT, *supra* note 3, at 10–11; *see also* Hill, *supra* note 9, at 80–81 (arguing that fear of mistakes may discourage attorneys from changing terms).

[24] See Daniel Schwarcz, *The Role of Courts in the Evolution of Form Contracts: An Insurance Case Study* (working paper, 2018) (exploring changes in insurance contracts as a result of litigated cases).

[25] 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).

[26] *See* Marotta-Wurgler & Taylor, *supra* note 17.

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. SYMMETRIC VERSUS ASYMMETRIC LEARNING

We propose a simple theory of contractual choice based on experiential learning. 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 will introduce in Section II.A. After having illustrated our framework of analysis, we will introduce learning in our theory, in Section II.B. In Section II.C, we will illustrate our main results by means of simple numerical examples. (The generality of this results will be demonstrated in Section III, where we develop our theory in a formal model.) In Section II.D, we 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 firm may decide to include a standard default term in the contract or to opt out of it. In the interim period, the firm may have an opportunity to learn the costs associated with offering the term.[27] If it learns new information, the firm may, at time 1, revise the contract.

For example, consider firms selling tax preparation software packages to consumers. Unless, modified or opted out, the standard default rules offered by Article 2 of the Uniform Commercial Code will apply to the transaction, including implied warranties. As they write their contract, firms can choose to keep the defaults or opt-out of them.[28] 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 default term or to opt out depends on two

[27] Note that our model does not require the firm to interact with the *same* consumers at time 0 and time 1. In fact, the model we present in the Appendix is predicated on the assumption that consumers are only active for one period. With repeat consumers learning might be faster because the characteristics of a firm's consumer base could be more stable over time and hence experience acquired at time 0 carries more information about the time-1 contracts. This aspect of the problem is captured in the general model presented in the Appendix where we allow for different degrees of learning. Our results are not affected by it.

[28] For a detailed analysis of how firms adopt or opt-out of UCC Article 2 default rules, see Marotta-Wurgler, supra note 2.

factors: on the one hand, a price increase that the firm is able to capture if it offers the default, as consumer might be willing to pay more for the product if the firm offers the warranty; on the other hand, 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.

Consider a second example, which we will use in the following to contrast the one above. The same firms may consider offering a second term restricting the use of the product by the consumer, a term opting out of the default term with no restriction. A common example is a restriction on commercial use, where the consumer is only able to use the product for personal or household purposes. Consumers value unrestricted use but the firm might benefit from carving out specific uses that, depending on market conditions, could be licensed separately. Therefore, offering the default implies an opportunity cost for the firm—corresponding to value of specific separate licenses—which may be unknown at present. As above, the decision whether to offer the default or opt out of it will depend on the price increase as compared to costs that are uncertain at the outset.

After interacting with consumers at time 0, the firm may learn the actual cost of offering the defaults but will do so in markedly different ways in the two examples above. Consequently, learning will affect both the time-0 and the time-1 contractual choices by firms. The next section will develop this point further.

B. Experiential learning

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. The revised form will apply to all subsequent transactions.

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.

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 two

possible values labeled *S* (for "symmetric") and *A* (for "asymmetric"). The information type relates to how the firm learns about the cost of the default term after time 0.[29] 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.

Terms of type *S* are symmetric-learning terms and are such that the firm receives new information at time 1 irrespective of whether it adopted the default or opted out of it at time 0. In particular, between time 0 and time 1, the firm learns whether costs are high or low. In this case, the firm has an opportunity to revise the contract at time 1 irrespective of the choice made at time 0. In contrast, terms of type *A* are asymmetric-learning terms: between time 0 and time 1, the firm learns the value of *c* only if it has adopted what we call the *learning modality* of the term, which could be either the default or the opt out. If the firm does not offer the learning modality at time 0, it will not learn anything from interacting with consumers. For simplicity, throughout the analysis we focus on cases where the learning modality is the default term but our analysis and results apply unchanged to the opposite case, where the learning modality is the opt out (a detailed analysis of both cases is provided in the Appendix).

Table 1. Information-types and modalities of contract terms

Information-type	Default	Opt-out
Symmetric-learning terms (*S*)	Learning	Learning
Asymmetric-learning terms (*A*)	Learning	Nonlearning

As illustrated in Table 1, symmetric-learning terms are always in the same modality irrespective of whether the firm adopts the default contract term or opts out of it. Examples of such terms are restrictions-on-use terms. In these cases, the firm learns information from its relationship with consumers over time, but choosing the default or opting out of it does not seem to affect learning. Consider a firm that adopted the default of no restriction. The firm might monitor use by consumers and notice that a specific category of its consumers uses the product particularly intensely along a specific dimension. For instance, it may discover that small retailers use the software produced by the firm for marketing purposes. As a consequence, it could in the future restrict the use of the software and sell a license for marketing uses separately at a higher price. Yet, the firm would learn the same information even if it adopted the restrictions-on-use opt-out term. The firm might start out by restricting use to non-marketing uses and then learn about the

[29] Note that also the value of the term for consumers could be uncertain. However, the firm does not need to interact with consumers who purchase the product to learn the value they place on the term. Rather, the firm may learn, for instance, by experimenting with different prices and monitoring the volume of sales. This is an important way to learn but does not exhibit the learning asymmetries on which we ground our theory and hence is not central to our framework. We therefore do not discuss learning about value in the text.

demand for the license for marketing uses directly by offering it for sale separately.[30] In both cases, the firm may acquire valuable information on the market value of specific uses. The opposite situation, where a firm doesn't learn from direct experience with consumers under either modality of a term, is also a symmetric term. The distinctive characteristic of a symmetric term is that a decision of whether to revise a term at a later period will arise irrespective of a term's modality at an initial period.

In contrast, asymmetric-learning terms can be in either learning or nonlearning modality depending on the contractual choice at time 0. Examples of such terms are default warranties, which could potentially expose firms to liability. Firms adopting the default warranty 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 resulting from breach of warranty claims. Firms opting out of the default and disclaiming implied warranties will not face such claims or claims will not result in a judicial assessment of damages resulting from such claims, and hence will not learn the costs of offering a warranty.

In the next Section, we will employ simple numerical examples to address the following question: How does the information-type of a term affect contractual choices at time 0 and at time 1? Analysis of these examples will anticipate results that we will formally demonstrate in Section III.

C. A numerical illustration of the results

To fix intuitions, let us now consider two simple examples involving a symmetric-learning term, that is, the restrictions-on-use term, as explained above.

Example 1. By offering the default term of no restrictions on use, Firm 1 can increase the price of its products by $100. In terms of lost opportunities, the default costs Firm 1 $150 (high costs) with probability 0.4 and $50 (low costs) with probability 0.6.

Should the firm offer the default term at time 0? Doing so costs Firm 1 an expected amount equal to 0.4*$150 + 0.6*$50 = $90. Consumers value the term at $100; hence, the firm can make a positive profit of $10 at time 0 by offering the default term. A restrictions-on-use term is a symmetric learning term. After time 0, Firm 1 will learn whether costs are in fact high or low. If the firm learns that costs are low, it will keep offering the default at time 1. If, instead, the firm learns that costs are high, it will revise its choice and opt out at time 1. It is important to stress that

[30] It is not crucial to our analysis that the symmetry in learning be perfect. The firm may learn in slightly different ways by offering the default or opting out. In the general model presented in the Appendix, we allow for imperfectly symmetric forms of learning. In the text, we focus on the sharp scenario in which learning is either perfectly symmetric or perfectly asymmetric.

learning will occur anyway, irrespective of the choice at time 0 and hence this choice has no effect on it. Consequently, Firm 1 should offer the default term and capture the $10 profit at time 0.

Consider a similar case, but now a different firm faces high costs with a higher probability.

Example 2. *By offering the default term of no restrictions on use, Firm 2 can increase the price of its products by $100. In terms of lost opportunities, the default costs Firm 1 $150 (high costs) with probability 0.6 and $50 (low costs) with probability 0.4.*

The parameters of the example are all the same but for the probability of facing high costs, which is now equal to 0.6. Should Firm 2 offer the default term? Doing so costs Firm 2 an expected amount equal to 0.6*$150 + 0.4*$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 chose the default term and should thus opt out. As Firm 1, also Firm 2 will learn after time 0 whether costs are high or low and will re-evaluate its choices accordingly.

To summarize, at time 0, Firm 1 will choose the default term of no restrictions on use and Firm 2 will opt out of it and restrict use by consumers. At time 1, both firms will offer the same version of the term because both of them will have learned whether costs are high or low. The rate of keeping the default at time 1 is equal to the probability that Firm 2—which is the only firm opting out of the default at time 0—faces high costs, that is 0.4.

Let us now consider two additional examples concerning the default warranty for consequential damages illustrated above, which is an asymmetric-learning term.

Example 3. *By offering consumers the default warranty, Firm 1 can increase the price of its products by $100. Offering the default costs Firm 1 $150 (high costs) with probability 0.4 and $50 (low costs) with probability 0.6.*

Here, we ask whether the firm will offer the default warranty. As before, doing so costs Firm 1 an expected amount equal to 0.4*$150 + 0.6*$50 = $90. Consumers value the term at $100; hence, the firm can make a positive time-0 short-term profit equal $10 by offering the default term. In addition, by offering the default the firm will make sure it learns the costs of the warranty and makes better choices in the future. This learning value, which we will quantify momentarily after having examined the next example, casts yet another vote in favor of offering the default as compared to the opt out. If the firm opted out, it would not learn. Firm 1 should thus offer the default at time 0, as it did in Example 1. At time 1, it will learn and keep offering the default only if costs are low, and revise to the opt out otherwise. The next example will show more clearly why learning plays an important role in firm choices when terms are asymmetric-learning terms.

Example 4. *By offering consumers the default warranty, Firm 2 can increase the price of its products by $100. Offering the default costs Firm 2 $150 (high costs) with probability 0.6 and $50 (low costs) with probability 0.4.*

Note that only the probability of high costs has changed with respect to Example 3. Should Firm 2 offer the default warranty? Doing so costs Firm 2 an expected amount equal to 0.6*$150 + 0.4*$50 = $110. Now the expected costs from breach of warranty claims exceed the value of the term. Consequently, Firm 2 would make a loss of $10 at time 0 if it offered the default term. Yet, this is not enough to conclude that Firm 2 should opt out.

In contrast to Example 2, which concerned a symmetric-learning term, in Example 4 Firm 2 learns only if it offers the default. Learning is lost if the firm chooses to opt out. Assume that Firm 2 chooses to opt out at time 0. The firm will not learn and hence will not have new information at time 1. Consequently, the firm's best choice at time 1 is to confirm the choice already made at time 0. The firm will keep the opt out provision in the contract it offers to consumers. Choosing the opt out at both times results in no gains and no costs for Firm 2. Overall, the firm earns zero from opting out at time 0.

Consider now the situation in which Firm 2 offers the default at time 0. As explained above, the firm will make a short-term loss equal to $10. However, the firm will learn after time 0 whether the warranty imposes high or low costs. If Firm 2 learns that costs are high it will revise the contract and opt out at time 1, avoiding further losses. If, instead, Firm 2 learns that costs are low it will confirm the choice made at time 0 and keep the default warranty in the contract. In that case, the firm will earn $100 and pay costs equal to $50, making a profit of $50.

At time 0, a firm offering the default will expect to learn that costs are high and hence to have to revise the contract with probability 0.6 (time-1 profits are then zero); it will learn that costs are low and will keep offering the default with probability 0.4 (in this case, time-1 profits are equal to $50). Summing up, offering the default at time 0 will allows the firm to capture an expected profit equal to 0.4*$50 + 0.6*$0 = $20 at time 1. The expected time-1 profits from offering the default, $20, more than offsets the time-0 loss of $10 and induce also Firm 2 to offer the default warranty at time 0.

Contrast this result with what we found when considering a symmetric-learning term. In Example 1 and Example 2, firms made different choices depending on whether the expected costs where low or high: Firm 1 chose the default while Firm 2 opted out of. In contrast, the firms in Example 3 and Example 4 make the same choices (they both offer the default). In particular, both firms choose the learning modality of the term, even when doing so results in short-term losses. Note that the numbers in the two examples are identical, the only change is the information-type of the term. When learning is asymmetric, choosing a term in its learning modality (in this case, the default) has an additional value that adds what is known as a "real option"[31] to the firm's short-

[31] 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*

term gains and costs: the option to revise the contract following learning. This option allows the firm to capture additional profits at time 1. While Firm 2 can do so irrespective of its time-0 choices in Example 2—so that the default and the opt out term give the firm the same options with respect to future profits—in Example 4, to learn the firm has to choose the default at time 0—and hence the default (which in this case allows the firm to learn the value of the term) gives the firm an option that the opt out denies. This asymmetry tilts the scale in favor of choosing the default even in the face of short-term losses at time 0.

In addition, there is an effect at time 1. Since both firms choose the default in our examples when learning is asymmetric, switches away from the default occur more frequently that with symmetric learning. Firm 1 will switch to the opt out with a probability equal to 0.4; in addition, Firm 2 will also switch to the opt out with a probability equal to 0.6, resulting in more switches away from the default when compared to symmetric learning. Recall that with symmetric learning only Firm 1 may switch away from the default with a probability equal to 0.4. This effect will be central in the empirical analysis that follows. We will fully exploit the implications of this observation momentarily, after having expanded on the scope on our theory.[32]

D. 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 to consumers could 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

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).

[32] Asymmetric learning may also come in a different guise: the firm may learn from opting out rather than from offering the default. This case is the mirror image of the case discussed in the examples and the results are the same: the learning modality of the term (the opt out in this case) is chosen more often at time 0 and hence there are more switches away from it at time 1. Our formal model and the empirical analysis take this case into account but we do not discuss it in the text because the analysis and the results are the same as with the asymmetric-learning terms where learning follows from offering the default.

supplier. We do not elaborate on this alternative learning motive but we stress that this is also a form of experiential learning. Conversely, 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. Learning about these costs may induce the firm to amend the clause at a later time. Our analysis applies to these cases.

More generally, our framework applies to all cases in which firms offer standard form contracts to consumers, learn from interacting with them and may amend their contracts as a result. Yet, firms routinely attempt to tailor their contracts to the specific characteristics of individual or groups of consumers. Firms in our model offer standard form contracts to consumers, but costs associated with offering a specific term may vary with consumer characteristics. In these cases, firms might find it advantageous to tailor contracts to specific consumers or consumer groups rather than offering all of their consumers the same contract. One might also question how the method of learning we propose interacts with other learning mechanisms. We discuss tailoring first.

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.[33] However, firms in our model cannot do so 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.[34] If offered different contracts, consumers in our model would all choose the same contract, defeating the firms' attempt to screen among them.[35]

Is this a realistic assumption? We think so. First, 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.[36] Second,

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

[34] 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.

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

[36] 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

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. Empirical evidence confirms this. For example, Della Vigna and Gentzkow find that retail chains in the United States do not adapt their prices to local conditions. The authors found that the retailers studied fail to tailor their contracts to easily identifiable groups of consumers living in different states 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; the fixed costs associated with tailoring other, more difficult to individualize, terms might be cost prohibitive as well.[37] Finally, even if firms were able to discriminate among consumers depending on their use patterns, there is no guarantee that they will have incentives to do so perfectly. In a recent study, Xinyu Hua and Kathryn Spier show that a monopolist may fail to supply the optimal level of product warranties even when it can price discriminate among consumers. The intuition is that the monopolist will tend to cater to the interests of the marginal consumer, which is typically not representative of the population of consumers.[38] Therefore, the possibility to learn may also alter the behavior of price-discriminating firms. By abstracting from price discrimination, we zero in on the important details of the analysis and offer insights that might be more generally applicable.

A second concern is whether the contract changes that we attribute to learning from experience might result from different learning mechanisms. As noted earlier, firms invest relevant resources in learning through other channels. The most intuitive account of contract change should consider in *all* the mechanisms by which learning and innovation can come about, given that it is unlikely that change is driven by only one form of stimulus. Our model does not exclude these contemporaneous drivers of change. Importantly, these alternative forms of learning are orthogonal to the experiential learning capability of each term so they do not affect contractual choice and modification in the same way. This allows us to measure experiential learning empirically, as we do in Section IV.

III. A FORMAL THEORY OF EXPERIENTIAL LEARNING IN CONTRACTS

We introduce a formal implementation of the theory presented in the previous section.[39] We will

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).

[37] 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).

[38] Xinyu Hua & Kathryn E. Spier, *Product Safety, Contracts, and Liability*, Working Paper (documenting market failures in the provision of product safety).

[39] 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.

employ the same framework of analysis introduced above, augmented with a simple formal apparatus that will allow us to generalize the results more systematically beyond the examples provided earlier. Formalization will also allow us to better appreciate the applicability of the theory. We will consider the following setup. At time 0, a firm offers a standard form contract to consumers and may offer a standard default term—say, a warranty or unrestricted use of the product as in the examples above—or opt out of it—no warranty or a restrictions-on-use provision, respectively.[40] In the interim period, the firm may learn and revise the contract.[41] For now, we assume that there are no switching or opting-out costs. We also assume that the volume of sales 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 choices in Section III.C.)

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 product if the firm offers the default warranty or unrestricted-use provision—but exposes the firm to either high costs, $c_H > v$ (with probability p), or low cost, $c_L < v$ (with the complementary probability, $1 - p$).[42] In the examples, these are the costs associated with consumer claims following the warranty and the foregone income associated with the unrestricted-use provision (where the firms fails to charge extra for additional uses of the product). Opting out does not produce any value and does not impose any cost.[43] In theory, firms would want to 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

[40] 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).

[41] 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 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.

[42] 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 be negative, which would indicate that the opt out is more valuable or costlier to implement, respectively, than the default. Our model covers also these case, although we do not discuss them explicitly in the text.

[43] 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.

contract to the actual costs they will face.[44] Firms can, however, tailor the contract to expected costs. This setup is summarized in Table 2.[45]

Table 2. 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	$1-p$

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. For simplicity, we assume that in the population of firms the firm-specific characteristic p is uniformly distributed from 0 to 1.[46]

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. 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 low-p firms facing low expected costs when offering the default—and firms with a p above the cutoff opt out—these

[44] 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 seen in Section II.D, 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.

[45] Table 2 does not account for a second type of asymmetric-learning terms: those in which the firm learns only if it opts out. We deal with this mirror-image case in the Appendix. Results are the same as in the baseline case. Additionally, 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.

[46] This assumption greatly simplifies the discussion and, in particular, allows us to easily characterize adoption and switching rates. The assumption, however, is innocuous for our results and it is relaxed in the Appendix.

are high-p firms. This cut-off level of p will depend on the information-type of the term and hence we will find two different cutoff levels when analyzing symmetric-learning terms, such as a restrictions-on-use term, and asymmetric-learning terms, such as a warranty.

A. Symmetric-learning terms

First, we focus on symmetric-learning terms, such as restrictions-on-use terms. While the real cost of the term is unknown to the firm at time 0, it will be known at time 1 due to the firm's experience with consumers in the interim period irrespective of whether the firm offered the default term or opted out at time 0. Since learning occurs anyway (either from experience or from another channel), or not at all, the firm's choice at time 0 is only affected by the expected costs and benefits of adopting a particular term in the short-term, that is, at time 0.

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 $pc_H + (1 - p)c_L < v$.[47] 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:

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

Firms characterized by $p < p^S$ 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 contrast, firms with $p > p^S$ 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.[48]

To fix intuitions, consider again the numerical examples presented above. In the examples, we have v = \$100, c_L = \$50, c_H = \$150. By plugging in these parameters in Expression (1), we can

[47] 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. Fourth, in a competitive market prices track costs, not consumer surplus. Hence firms might adjust the price they charge to consumers after learning the costs or different clauses. These adjustments may erode firms' profits but should not affect the key mechanism behind our model. A formal model of standard form contracts in competitive markets might unveil additional implications.

[48] If $p = p^S$, expected costs are exactly equal to the value of the term; for ease of notation we disregard this possibility.

easily see that, under the examples specifications, $p^S = 0.5$. This confirms that Firm 1 in Example 1, which has a $p = 0.4 < p^S$, will adopt the default term, while Firm 2 in Example 2, which is characterized by a $p = 0.6 > p^S$, will opt out.

After interacting with consumers at time 0, the firm learns whether it faces low or high costs. Hence, 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 1. Learning terms: adoption decisions at time 0 and time 1

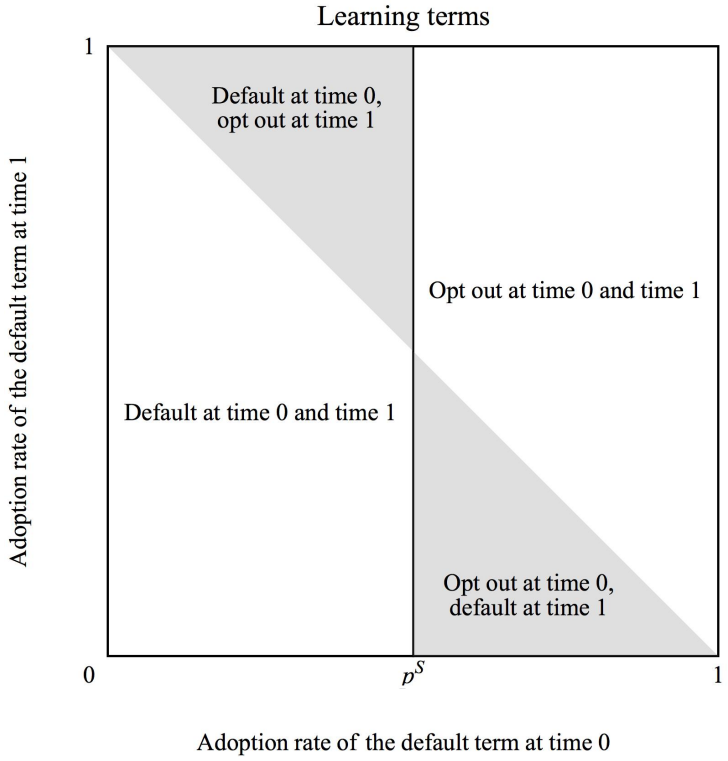


Figure 1 shows the adoption rates of the default term at time 0 and time 1 for symmetric-learning terms.[49] At time 0 firms that adopt the default are those with $p < p^S$, like Firm 1. At time 1, however, there is new information available. Adopters may discover that costs are high—this happens with probability p—and decide to switch to the opt out; with the complementary probability $1 - p$, adopters discover that costs are in fact low and keep adopting the default term. The area of the grey triangle in the upper-left corner of the graph 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 Example 1, Firm 1 will switch to the opt out at time 1 with probability 0.4.

[49] In the figures, we focus on the special case in which v falls exactly half-way between c_L and c_H and hence $p^S = 0.5$. The examples in the text are a subspecies of this case, with v = \$100 being exactly half way between c_L = \$50 and c_H = \$150.

Conversely, firms with $p > p^S$, like Firm 2, decide not to adopt the default term at time 0 when learning is symmetric. 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 in the lower-right corner of the graph depicts the ex ante probability of switches from opt out to default. Firm 2 in Example 2 will switch away from the opt out to the default at time 1 with probability 0.4. Considering all firms in the explanatory case described in Figure 1, 50% of the firms choose the default at time 0 and 25% of them (that is, 12.5% of the total number of firms)[50] switch to the opt out at time 1. The remaining 50% of the firms 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.

Since the probability of time-1 switches away from the default is central to our analysis, it is worth giving it a more general characterization. This probability is equal to the area of the gray region in the upper-left corner of Figure 1. Using simple arithmetic, this area is easily calculated and is equal to $\frac{(p^S)^2}{2}$. When $p^S = 0.5$, this is in turn equal to $\frac{1}{8}$ or 12.5%.

B. Asymmetric-learning terms

Here we consider asymmetric-learning terms A. Now the firm learns its costs only if it has adopted the learning modality of the term—which we assume to be the default—at time 0. Adoption of the learning modality of the term 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.

The asymmetric learning value of the alternatives available to the firm at time 0 gives rise to a real option. By choosing the learning modality—that is, the default term—the firm acquires 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 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 greater than p^S.

More precisely, 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 $v - (1 - p)c_L - pc_H + (1 - p)(v - c_L)$, where the latter term captures the option value of adopting the default, which adds to the time-0 short-term profits captured by the first three terms in the expression. The ex ante payoff

[50] 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.

from opting out is 0. The firm's payoff is greater with the default if p is greater than a value p^A defined as follows:

$$p^A \equiv 2\frac{v - c_L}{c_H - c_L + v - c_L} \quad (2)$$

Firms with $p < p^A$ choose the default, while firms with $p > p^A$ opt out. Simple calculations show that $p^A > p^S$: more firms choose the default with asymmetric-learning terms than they did with symmetric-learning terms. These additional adopters sustain short-term losses, which can be interpreted as the price the pay to learn. In this way, they purchase an option to make better choices in the future.

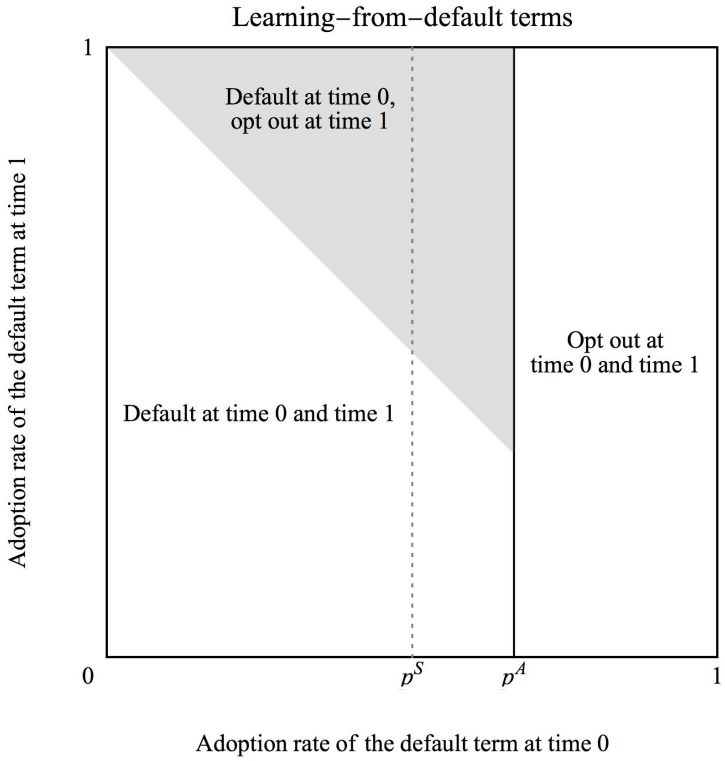
To fix intuition, consider again our running examples. By plugging in the parameters v = \$100, c_L = \$50, c_H = \$150 in Expression (2) we have that $p^A = \frac{2}{3}$. This value is greater than the value of p^S = 0.5 that we found with symmetric learning terms. Firms with a p falling between these two thresholds—that is, with a p greater than p^S but less than p^A—would not choose the default if the term is a symmetric-learning term but do so if learning is asymmetric. Consider Firm 2 facing p = 0.6: we have 0.6 > 0.5 in Example 2, and hence the firm opts out; but we have 0.6 < 2/3 in Example 4, and hence the firm offers the default.

Figure 2 shows the effects of asymmetric learning.[51] The cut-off level of p at time 0 is greater than the one we found with symmetric learning: more firms choose the default at time 0. Adopters, however, switch to opt-out with relatively high probability, especially in the range $[p^S, p^A]$, 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.[52] Note the asymmetry: only those firms that choose the default at time 0 may switch at time 1. Firms that choose to opt out irreversibly confirm that decision at time 1 because no new information is acquired in the meantime.

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

[51] We focus again on the special case in which v falls exactly half-way between c_L and c_H. See note 49.

[52] 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 2 to gain some insights on adoption rates. At time 0, 1/3 (that is, 33.3%) of the firms opt out of the default. These firms do not learn hence there are no switches at time 1. The remaining 2/3 (66.7%) of the firms offer the default term—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, 1/3 of 1/3 or 22.2%) of all firms switch at time 1 (this is the area of the gray region in Figure 2). Compared to the symmetric-learning terms, there are substantially more switches from default to opt out: 22.2% instead of 12.5%.[53]

Let us give again a general characterization of the ex post probability of switches away from the default at time 1. This probability is again to area of the gray region in Figure 2. Using again simple arithmetic, this area is easily calculated and is equal to $\frac{(p^A)^2}{2}$. Since we have established that $p^A > p^S$, the probability of switches away from the default is now larger than the one we found in symmetric learning terms, which was equal to $\frac{(p^S)^2}{2}$. In the ecosystem of the above examples, where $p^A = \frac{2}{3}$, the ex post probability of switches away from the default is to $\frac{2}{9}$ or 22.2%, as we found in the example. Summing up, when learning is asymmetric the learning modality is chosen more often at time 0 but is also abandoned at higher rates at time 1, if compared to symmetric learning terms.

[53] Note that in this specific example there are fewer switches in total when learning is asymmetric: 22.2% instead of 25%, which we had with symmetric learning.

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	
	Symmetric-learning terms	Asymmetric-learning terms
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.[54] 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 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 of the 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,

[54] *See* Goetz & Scott, supra note 10.

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 and the learning modality will be chosen less often. If the learning modality is the default, as we assume here, greater switching costs will result in a lower adoption rate of the default at time 0 and hence also in a lower switching rate away from the default at time 1.[55]

In symmetric-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 if it is the learning modality.[56] Yet, growth has no effect on symmetric-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.

[55] If the learning modality is the opt out, higher switching costs will result in a smaller adoption rate for the opt out at time 0 and a lower switching rate away from the opt out at time 1 (and hence greater adoption and switching rates for the default). See the Appendix.

[56] Vice versa, the default will be chosen less often if the opt out is the learning modality.

IV. EMPIRICAL ANALYSIS

We now put our theory to bear on the contractual choices made by real firms. We first derive empirically testable predictions from our theory. Next, we present our dataset and empirical results.

A. Empirical implications of the theory

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. We explore learning from previous contractual choices against the attractiveness and stickiness of default terms that have been explored and identified in previous literature.

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. In contrast, 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. For these terms, experiential learning, or learning from other channels, occurs (or not) irrespective of the contractual choice at time 0. If the firm learns from experience, learning

will occur 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. The same is true for when the firms learns from other means. Revisions of the term at a later date will be uncorrelated with the contractual choices made during the earlier period.

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 than when it is the nonlearning modality.

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.[57] This allows for meaningful comparisons across contracts. We examine the rate of

[57] Marotta-Wurgler, *supra* note 2.

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.[58] 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 by Marotta-Wurgler and Taylor,[59] 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 4 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.

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

[58] For a full description of the data collection process, see Marotta-Wurgler & Taylor (2013), *supra* note 17.
[59] *See* Marotta-Wurgler & Taylor, *supra* note 17.

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.[60]

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 by Marotta-Wurgler and Taylor[61] 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 5.

Not all terms give sellers the same opportunities to learn. The first column of Table 5, labelled "Learning Category," 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, though not necessarily from direct experience with consumers. The act of collecting information will 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. As noted earlier, symmetric learning terms might allow firms to learn about the cost of offering such terms from experience under all modalities of the term or under none of them. Sometimes learning from

[60] 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).

[61] *See* Marotta-Wurgler & Taylor, *supra* note 17.

experience is not the most direct form of learning for a particular term. Consider, for example, a term limiting the transfer of the software. Sellers might consider revising the scope of this right by evaluating consumer demand or by studying what competitors offer. Consumer might also request for broader transfer rights. The mechanism by which sellers learn, whichever it may be, will not depend on the initial choice on a particular modality of the term.

The table labels such terms as "S"—i.e., symmetric learning. We identify fourteen such terms, including the two described above. Two terms in this group are 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. Another 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. Additional terms 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.

We now turn to asymmetric learning clauses, labeled as "A". In contrast to the terms described above, 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 (or "A(O)") and when then they learn only when they adopt the default (or "A(D)"). Narrowing the classifications in this manner allows us to examine the relative attractiveness of default rules. There are no default express warranties, so the seller learns only by opting out of the default. 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 find twelve such terms. These include two clauses allowing the buyer to create derivative works and reverse engineering (which are allowed under intellectual property law), 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 5 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 5 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[62] 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. 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. 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,

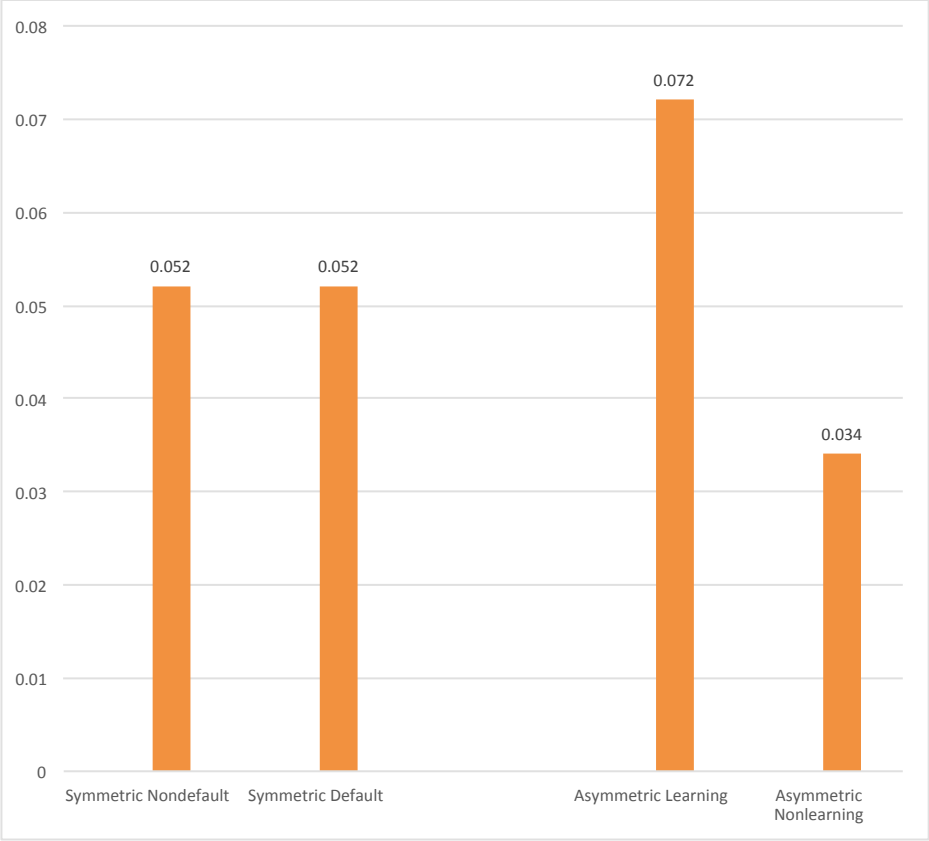
[62] *See* Marotta-Wurgler & Taylor, *supra* note 17.

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 explain contractual change and innovation. That being said, while the findings are consistent with our theoretical predictions, observational data can never prove that learning is one of the causes of changes in terms. Future work could complement the large-sample evidence with case studies and interviews with general counsel to give a nuanced, descriptive view of why standard terms change over time.

Figure 3. 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 3. 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 7 reports regressions including company, product, and market control variables. The first column just repeats the results from the bottom of Panel C of Table 6. 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.[63]

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, suggesting that lawyers might be part of the mechanism by which experiential knowledge generates change in standard terms.

Finally, Table 8 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 6 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

[63] The terms more likely to change when set in their learning modes are those that address the buyer's ability to reverse engineer the product, and those that determine the sellers' implied warranty obligations and liability for consequential damages of the buyer (unreported).

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.

D. 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. For example, we would expect to observe even more experiential learning in insurance markets, given the nature of the terms and the stakes at issue.

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 a 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.

There are alternative hypotheses that could explain change as well. The changes we observe could, in theory, be explained by a general trend to revise terms in a particular direction. Yet we established earlier that experiential learning had its own effect despite general tendencies to revise a contract. Or sellers offering terms to learn about a supplier's reliability would respond by switching suppliers and not necessarily terms. This could be the case, and it would work against finding any changes in the contract. Finally, a note on learning. It would be natural to expect in many cases for experiential learning to be sequential, where sellers learn by trial and error with different iterations. Our findings are consistent with this version of learning as well.

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. Some normative implications arise. First, law makers' deciding whether to create or modify default rules (e.g., in the context of revising Articles or provisions of the U.C.C) might

want to factor the extent to which certain default rules allow for experiential learning, especially when default rules are sticky. All else equal, it might be more desirable to adopt defaults that carry an opportunity to learn. Second, 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.[64] Finally, experiential learning weighs against the implementation of mandatory rules, since, all else equal, these never allow learning.

V. CONCLUSION

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.

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.

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. Understanding how these interact, as well as how new technology affects

[64] 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).

the way in which firms learn, are important questions that are the subject of future work.

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.[65] 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.

[65] 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.

A. Tables

Table 4. 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 5. 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, when learning occurs or not 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	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	Are license's terms subject to change?	0 = no 1 = yes	0.106 (0.309)	0.076 (0.265)	0.030** (0.012)
S	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	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	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	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	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	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	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	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	0.553 (0.498)	0.519 (0.501)	0.034* (0.019)

		1 = cap on damages less than or equal to purchase 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	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	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	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	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	Does license inform licensee of statutory rights?	0= no or no mention			
	Total Mean Change				**0.583*** **(0.128)**

Table 6. 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; 8,448 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	2010 term: 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	2010 term: 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 7. *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***	0.0332***	0.0412***	0.0323***	-0.0757	-0.246**
	(0.00533)	(0.00399)	(0.00525)	(0.00588)	(0.0507)	(0.0996)

Fixed Effects	None	Firm	Firm	Firm	None	None
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 8. 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***

Table 9. 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***

B. Term Classification

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, when learning occurs or not 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). For each term, the Table reports the rationale support a particular experiential learning classification, as noted in the last column, "Classification Rationale."

Term #	Learning Category	Term (t)	Classification Rationale	Learning (0=no; 1=yes)
		Acceptance		
x_1	S	Does license alert consumer that product can be returned if she declines terms? 1=yes; 0=no	Notice term giving pure information to the consumer. Feedback about the value of such term is unlikely to arise from experience.	0
		Modification and Termination		
x_2	S	Are license's terms subject to change? 0=no; 1=yes	Term altering the process of contractual modification. Feedback about the value of such term is unlikely to arise from experience.	0
x_3	S	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	Does definition of "licensed software" include updates, enhancements, versions, releases, patches, etc.? 1=yes;0=no mention/no	Product attribute. Feedback occurs in either case, though not necessarily from experience.	0
x_5	S	Can licensee alter/modify the program? 0=yes or no mention; -=no	Product attribute. Feedback occurs in either modality, though not necessarily from experience.	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	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 attribute. Feedback occurs in either modality, though not necessarily from experience.	0
		Information Collection		
x_9	S	Does license allow licensor to collect and /or distribute licensee's information? 0=no/no mention 1=yes	Some feedback in either case. Seller will learn in the future whether collecting information gives him a competitive advantage or not-collecting information makes the 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	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 attribute. Feedback occurs in either modality, though not necessarily from experience.	0

x_{12}	S	Can Licensee transfer the software if end user accepts license terms? 0=yes or no mention; 1=no	Product attribute. Feedback occurs in either case, though not necessarily from experience.	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	Is the disclaimer in caps? 0=yes or no disclaimers appear; 1=no	Product attribute. Feedback occurs in either case, though not necessarily from experience	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	Law specified? 0=same as forum or no mention; 1=yes and different from forum	Sellers receives feedback in either modality.	1

x_{28}	S	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	Does license require licensee agree to third party licenses or terms? 0=no; 1=yes	Product attribute. Feedback occurs in either case, though not necessarily from experience.	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	Does license allow licensor or third parties to install additional software? 0=no; 1=yes	Product attribute. Feedback occurs in either case, though not necessarily from experience.	0
		Consumer Protection		
x_{32}	S	Does license inform licensee of statutory rights? 0=no; 1=yes	Product attribute. Feedback occurs in either case, though not necessarily from experience.	0

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 cumulative distribution function $F(p)$,[66] 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 prescribed by the law or opt out of it.

The default term has a value v^D for consumers, while the alternative opt-out term has known value v^O. The costs of the default, $c^D \in \{c_L^D, c_H^D\}$, and the opt out, $c^O \in \{c_L^O, c_H^O\}$, however, are uncertain and can be either high (H), with probability p, or law (L), with the complementary probability $1-p$.[67] The upper part of Table 10 illustrates this setup.

Table 10. 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.[68] Hence, the firm maximizes 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 would adopt the default if

[66] We assume that $F(p)$ is continuously differentiable over the interval $(0,1)$.

[67] 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.

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

$v^D - c^D > v^O - c^O$ (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:

$$v > c \tag{3}$$

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-cost terms, as summarized in the bottom part of Table 10.

Consistently with our notation and without loss of generality, we assume that $c_H > c_L$, which allows us to give concrete 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.[69]

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).

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. Firms that expect to grow are characterized by $w > 1$: time-1 sales

[69] 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.

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 $w < 1$.

Finally, if switching costs are prohibitively high compared to the efficiency gains switching brings, then switching does not take place and learning becomes an uninteresting option. Therefore, we focus on the case in which switching at time 1 is a realistic option for the firm, that is, we restrict the analysis to cases in which $s < w(c_H - v)$ and $s < w(v - c_L)$: the firm's switching cost is less than the efficiency gain from switching, that is, the difference between the relative value and the relative cost at time 1 both when the opt out is the efficient choice and when the default is the efficient choice, respectively.

2. Information and learning

At time 0 the firm observes 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 the type *H* or *L* of the term. More simply put, the firm observes v but does not observe c. Therefore, contractual choices at time 0 are made in a condition of incomplete information. Consumers only know their valuations v^D and v^O.[70]

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 the information-type of the term. To generalize the framework presented in the text, we postulate that the firms will learn c with probability $\lambda^D \in [0,1]$ if it adopts the default and with probability $\lambda^O \in [0,1]$ if it opts out. The information-type of a term is then the pair (λ^D, λ^O).[71]

Therefore, when $\lambda^D = \lambda^O = 1$ we have the symmetric-learning term described in the analysis. (More generally, whenever $\lambda^D = \lambda^O$ learning is symmetric).[72] Instead, when $\lambda^D = 1$ and

[70] 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) is only valuable if it can be used to make inference about the population of consumers, rather than on individual consumers.

[71] In the model, we take (λ^D, λ^O) to be exogenous. However, the firm might be in control of these probability by, for instance, deciding to experiment with a term on a subset of its consumers. By doing so the firm reduce short-term losses in exchange for a reduced probability of learning. While this is an interesting extension of our model it is unlikely to affect our results: after the firm has set (λ^D, λ^O) our model would carry on virtually unchanged. In addition, the extent to which firms can treat similar consumers differently is limited both by law and by reputational concerns.

[72] Note that in the degenerate case of $\lambda^D = \lambda^O = 0$ the firm does not learn anything from interacting with consumers. Our analysis also covers this case.

$\lambda^O = 0$ we have the case of asymmetric learning from the default option described in the text. When $\lambda^D = 0$ and $\lambda^O = 1$, we have the mirror-image case of asymmetric learning from the opt-out option. The model presented here is more general than the cases discussed in the main text and allows for different degrees of learning from the default and the opt-out option and hence different degrees of asymmetry in learning. The remaining of this section provides a motivational example for the learning process implied by this reduced-form model. (Readers who are only interested in the analysis may skip it.)

In the model, we postulate that the firm may experience different costs for terms of types *L* and *H* and yet may not learn the term's type. To see how this may be possible, consider that costs may be high or low for various reasons, which imply different degrees of learning. For instance, costs could be either high or low because they depend on the personal characteristics of consumers, such as use patterns, which in turn vary stochastically in the population. In this case, a firm facing, say, a probability of high costs $p = 0.4$ simply faces a population that is composed for 40% of individuals whose characteristics imply high costs. The firm experience with its consumer base does not bring new knowledge. Consider instead a different scenario where the firm faces a homogenous population of consumers whose use patterns imply *either* high *or* low cost, again with $p = 0.4$. In this case, the firm will experience either high or low costs, but not both, and hence has a possibility to learn. Correlation in consumer use patterns related to the costs of offering a term is thus one of the reasons why firms may or may not learn from experience.

To motivate asymmetric learning, assume that observing costs is revealing, as illustrated above, and consider the following three scenarios.[73] (i) If costs are such that $c_L^D < c_H^D$ and $c_L^O < c_H^O$, 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. This is a case of symmetric learning. (ii) If costs are such that $c_L^D < c_H^D$ and $c_L^O = c_H^O$, 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 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. This is a case of asymmetric learning that occurs only if the option chosen is the default. Finally, (iii) if costs are such that $c_L^D = c_H^D$ and $c_L^O < c_H^O$, 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. This is a case of asymmetric learning that occurs only if the option chosen is the opt out. The

[73] 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.

learning probabilities λ^D and λ^O capture the fact that observing actual costs may or may not yield useful information about future costs.

3. Determinants of contractual choice

As usual, in the analysis 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 $\tau \in \{D, O\}$ at time 0 as $\Pi_t^\tau = v^\tau - c^\tau$. The firm maximizes the following payoff

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

That is, the firm maximizes the expected payoff it receives at time 0 (given that it chooses without knowing the type of the term) 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^τ, will depend on the choice made at time 0.

The expected payoff from choosing a term τ at time $t = 0$ is simply the difference between the value of the term and its expected costs:

$$\mathrm{E}[\Pi_0^\tau] = v^\tau - pc_H^\tau - (1-p)c_L^\tau \tag{5}$$

The expected profit at time $t = 1$ is more involved. If the firm learns, with probability λ^τ, the firm will adopt the efficient term, which is the opt out with probability p and the default with probability $1 - p$. If, with the complementary probability $1 - \lambda^\tau$, the firm does not learn, then the firm will not switch and will earn again its time-0 profits.

$$\mathrm{E}[\Pi_1^\tau] = \lambda^\tau[p(v^O - c_H^O) + (1-p)(v^D - c_L^D)] + (1-\lambda^\tau)\mathrm{E}[\Pi_0^\tau] \tag{6}$$

Filling is the appropriate values for the default and the opt out and rearranging, we have:

$$\mathrm{E}[\Pi_1^D] = \lambda^D p(c_H - v) + \mathrm{E}[\Pi_0^D] \tag{7}$$

and

$$\mathrm{E}[\Pi_1^O] = \lambda^O(1-p)(v - c_L) + \mathrm{E}[\Pi_0^O] \tag{8}$$

These expressions indicate that at time 1 the firm is guaranteed its time-0 profits and, in addition, can capture the efficiency gains due to learning. These gains are equal to the probability of learning times the probability of switching efficiently times the profits of the (efficient) alternative not chosen relative to the one chosen at time 0.

Finally, the probability of switching is $q^D = \lambda^D p$ if the firm chooses the default at time 0

and $q^O = \lambda^O(1-p)$ if the firm chooses the opt out. This probability depends on learning and on the probability that the opposite option to the one chosen initially is the efficient one. A firm characterized by p and maximizing the net payoff in (4) is indifferent between the default and the opt out if:

$$\mathrm{E}[\Pi_0^D] + w\mathrm{E}[\Pi_1^D] - q^D s = \mathrm{E}[\Pi_0^O] + w\mathrm{E}[\Pi_1^O] - q^O s \tag{9}$$

After substituting (7) and (8) into (9) and rearranging, we can derive a cutoff level of p below which the firm chooses the default term and above which the firm chooses the opt out. This cutoff defines the take-up rate of the default option at time 0 as being $F(p^*)$, since firms with a p below the threshold adopt the default and firms with a p above the threshold opt out of it. We have:

$$p^* \equiv \frac{(1+w)(v-c_L) - \lambda^O[w(v-c_L)-s]}{(1+w)(c_H-c_L) - \lambda^D[w(c_H-v)-s] - \lambda^O[w(v-c_L)-s]} \tag{10}$$

The second measure in which we are interested is the mass of switches at time 1. The probability that a firm p switches away from the default to the opt out at time 1 is equal to the probability of learning λ^D, given that the firm has chosen the default at time 0, times the probability that the opt out is the superior choice, which is equal to p. Considering all firms, we have:

$$\Delta^*(p^*) \equiv \int_0^{p^*} \lambda^D p dF(p) \tag{11}$$

Similarly, the mass of switches away from the opt out to the default at time 1 is equal to

$$\Omega^*(p^*) \equiv \int_{p^*}^1 \lambda^O(1-p) dF(p) \tag{12}$$

We will now analyze how the firm's decisions depend on the information-type of a term, (λ^D, λ^O). The information-type of a term is perfectly symmetric if $\lambda^D = \lambda^O$. If $\lambda^D > \lambda^O$ learning is asymmetric and the default provides for more effective learning. We say in this case that the default is the learning modality of the term; the case presented in the main text is an extremely sharp one, with $\lambda^D = 1 > 0 = \lambda^O$. Analogously, the opt out is the learning modality if $\lambda^D < \lambda^O$. We demonstrate here two general results.

Proposition A1. The learning modality of a term is chosen increasingly often at time 0 as learning becomes more asymmetric.

Proof. The rate of adoption of the default at time 0 is given by the mass of firms with $p < p^*$, which is equal to $F(p^*)$ and increases monotonically in p^*. Similarly, the rate of adoption of the opt out at time 0 is given by the mass of firms with $p > p^*$, which is equal to $1 - F(p^*)$ and decreases monotonically in p^*. Therefore, we can focus on how learning asymmetry affects p^*.The proposition implies that, if $\lambda^D > \lambda^O$, the default is chosen more often at time 0 as the difference $\lambda^D - \lambda^O$ increases; this is the case if $F(p^*)$ increases in λ^D and decreases in λ^O, which in turn is

the case if p^* increases in λ^D and decreases in λ^O. Analogously, if $\lambda^D < \lambda^O$, the opt out is chosen more often at time 0 as the difference $\lambda^O - \lambda^D$ increases; this is the case if $1 - F(p^*)$ decreases in λ^D and increases in λ^O, which occurs, again, if p^* increases in λ^D and decreases in λ^O. Since the condition that applies when the default is the learning modality is the same as the condition that applies when the opt out is the learning modality, to prove Proposition A1 it is enough to show that $\frac{\partial p^*}{\partial \lambda^D} > 0$ and $\frac{\partial p^*}{\partial \lambda^O} < 0$. We have:

$$\frac{\partial p^*}{\partial \lambda^D} = \frac{[w(c_H - v) - s][(1 + w - w\lambda^O)(v - c_L) + \lambda^O s]}{\left[(1 + w)(c_H - c_L) - \lambda^D[w(c_H - v) - s] - \lambda^O[w(v - c_L) - s]\right]^2} \qquad (13)$$

The latter expression is positive if $s < w(c_H - v)$, that is, if switching costs are not so high as to prevent switching at time 1, which we have assumed at the outset. Similarly, we have:

$$\frac{\partial p^*}{\partial \lambda^O} = -\frac{[w(v - c_L) - s][(1 + w - w\lambda^D)(c_H - v) + \lambda^D s]}{\left[(1 + w)(c_H - c_L) - \lambda^D[w(c_H - v) - s] - \lambda^O[w(v - c_L) - s]\right]^2} \qquad (14)$$

which is negative if $s < w(v - c_L)$, as we assume. *Q.E.D.*

Proposition A2. Switches away from the learning modality of a term occur increasingly often at time 1 as learning becomes more asymmetric.

Proof. The proposition implies that if $\lambda^D > \lambda^O$, there are more switches away from the default at time 1 as the difference $\lambda^D - \lambda^O$ increases; that is, it implies that $\Delta^*(p^*)$ increases as the difference $\lambda^D - \lambda^O$ increases. Note that $\Delta^*(p^*)$ increases in p^*. Proposition A1 shows that as the difference $\lambda^D - \lambda^O$ increases, then p^* increases, which implies that $\Delta^*(p^*)$ also increases. Conversely, the proposition also implies that if $\lambda^D < \lambda^O$, there are more switches away from the opt out at time 1 as the difference $\lambda^O - \lambda^D$ increases; that is, it implies that $\Omega^*(p^*)$ increases as the difference $\lambda^O - \lambda^D$ increases. Note that $\Omega^*(p^*)$ decreases in p^*. Proposition A1 shows that as the difference $\lambda^O - \lambda^D$ increases, then p^* decreases, which implies that $\Omega^*(p^*)$ increases. *Q.E.D.*

These two general results apply directly also to the sharp cases discussed in the main text, where we assumed for simplicity that the population of firms is uniformly distributed, that is, that $F(p) = p$. In what follows, we characterize these cases more explicitly and derive the comparative statics presented in the main text.

4. Symmetric-learning terms

Let us first consider symmetric-learning terms with $\lambda^D = \lambda^O = 1$. Plugging these values into (10), we have the adoption rate for the default term at time 0:

$$p^S = \frac{v - c_L + s}{c_H - c_L + 2s} \tag{15}$$

The expression we found in the main text did not consider switching costs and growth opportunities. Thus, that expression is a special case of the more general formulation in (15) and can be obtained from it by setting $s = 0$. Similarly, the switching rates are:

$$\Delta^S = \int_0^{p^S} p dp = \frac{(p^S)^2}{2} \tag{16}$$

and

$$\Omega^S = \int_{p^S}^1 (1 - p) dp = \frac{(1 - p^S)^2}{2} \tag{17}$$

Note that the adoption rate p^S is independent of w and increases in v. Moreover, we have:

$$\frac{\partial p^S}{\partial s} = \frac{c_H + c_L - 2v}{(c_H - c_L + 2s)^2} \tag{18}$$

The latter expression is positive if $v < \frac{c_H + c_L}{2}$ and negative if $v > \frac{c_H + c_L}{2}$. To rationalize this result, consider that switching costs put more weight on the choice that is most likely to be sticky later on. If v is small, switching at time 1 to the default is relatively expensive for the firm compared to the value it brings and hence the firm, to compensate, chooses the default more often at time 0. The opposite occurs if v is large.

Finally, note that the switching rates at time 1 depend directly only on p^S. Therefore, Δ^S increases in v as p^S does, while s has an ambiguous effect. In contrast, Ω^S moves in the opposite direction as compared to p^S and hence decreases in v, while s has again an ambiguous effect. These observations prove the results of the comparative statics presented in the main text.

5. Asymmetric-learning terms

Here we consider asymmetric-learning terms with $|\lambda^D - \lambda^O| = 1$, which encompasses the case discussed in the main text of asymmetric-learning from the default, with $\lambda^D = 1$ and $\lambda^O = 0$, and its mirror image of asymmetric learning from the opt out, with $\lambda^D = 0$ and $\lambda^O = 1$. As we will see, the results derived in the text for the former case apply to the latter.

If the learning modality is the default, its adoption rate at time 0 can be obtained by plugging $\lambda^D = 1$ and $\lambda^O = 0$ into (10) to obtain:

$$\overline{p}^A \equiv \frac{(1 + w)(v - c_L)}{c_H - c_L + w(v - c_L) + s} \tag{19}$$

Similarly, if the learning modality is the opt out, its adoption rate at time 0 is $1 - \underline{p}^A$ can be obtained by plugging $\lambda^D = 0$ and $\lambda^O = 1$ into (10) to obtain:

$$\underline{p}^A \equiv \frac{v - c_L + s}{c_H - c_L + w(c_H - v) + s} \tag{20}$$

If $v \neq \frac{c_H - c_L}{2}$, then the two adoption rates diverge because the default and the opt out option bring different net profits when used efficiently. To illustrate, if $v > \frac{c_H - c_L}{2}$, then the net profits from adopting the default efficiently at time 1—that is, $v - c_L$—are greater than the net profits from adopting the opt out efficiently at time 1—that is $c_H - v$—and, consequently, the learning modality is chosen more often when it coincides with the default (and vice versa if $v < \frac{c_H - c_L}{2}$).

Instead, if $v = \frac{c_H - c_L}{2}$, that is, when the relative value of the default falls exactly in between the high and low cost, the net value of the learning modality is not affected by it being the default or the opt out, as we have assumed in the examples in the text. In this special case, the adoption rate of the learning modality at time 0 is equal to $\overline{p}^A = 1 - \underline{p}^A$. The value we found in the main text can be obtained again from (19) by ignoring the switching costs and the growth options ($s = 0$ and $w = 0$). Note that, following Proposition A1 the learning modality is chosen more often when learning is asymmetric if compared to the same modality of the term with symmetric learning, that is, $\overline{p}^A > p^S$ and $1 - \underline{p}^A > 1 - p^S$.

The switching rates are

$$\Delta^A = \int_0^{\overline{p}^A} p dp = \frac{\left(\overline{p}^A\right)^2}{2} \tag{21}$$

if the learning modality is the default and is easy to see that, following Proposition A2, switches away from the learning modality are more frequent if learning is asymmetric, that is, $\Delta^A > \Delta^S$. If the learning modality is the opt out, we have:

$$\Omega^A = \int_{\underline{p}^A}^1 (1 - p) dp = \frac{\left(1 - \underline{p}^A\right)^2}{2} \tag{22}$$

Again, following A2 switches away from the learning modality are more frequent if learning is asymmetric, that is, $\Omega^A > \Omega^S$. Now let us investigate how adoption decisions respond to changes in the parameters. If the learning modality is the default, we have

$$\frac{\partial \overline{p}}{\partial s} = -\frac{(1 + w)(v - c_L)}{(c_H - c_L + w(v - c_L) + s)^2} < 0 \tag{23}$$

and

$$\frac{\partial \overline{p}}{\partial w} = \frac{(v - c_L)(c_H - v + s)}{(c_H - c_L + w(v - c_L) + s)^2} > 0 \tag{24}$$

and finally

$$\frac{\partial \overline{p}}{\partial v} = \frac{(1+w)(c_H - c_L + s)}{(c_H - c_L + w(v - c_L) + s)^2} > 0 \tag{25}$$

If the learning modality is the opt out, we have:

$$\frac{\partial \left(1 - \underline{p}\right)}{\partial s} = -\frac{(w+1)(c_H - v)}{(c_H - c_L + w(c_H - v) - s)^2} < 0 \tag{26}$$

and

$$\frac{\partial \left(1 - \underline{p}\right)}{\partial w} = \frac{(v - c_L - s)(c_H - v)}{(c_H - c_L + w(c_H - v) + s)^2} > 0 \tag{27}$$

and finally

$$\frac{\partial \left(1 - \underline{p}\right)}{\partial v} = -\frac{c_H - c_L + w(c_H - c_L - s) + s}{(c_H - c_L + w(c_H - v) + s)^2} < 0 \tag{28}$$

These results suggest that the learning modality is adopted more frequently at time 0 if the switching costs go down (smaller s), if the growth prospects improve (larger w) and if the relative value of the option improves (that is, larger v if the learning modality is the default and smaller v if the learning modality is the opt out). Switching rates at time 1 react in the same way.