ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Israel, Mark

Working Paper

Consumer learning about established firms: Evidence from the automobile insurance

CSIO Working Paper, No. 0020

Provided in Cooperation with:
Department of Economics - Center for the Study of Industrial Organization (CSIO), Northwestern University

Suggested Citation: Israel, Mark (2003) : Consumer learning about established firms: Evidence from the automobile insurance, CSIO Working Paper, No. 0020, Northwestern University, Center for the Study of Industrial Organization (CSIO), Evanston, IL

This Version is available at:
https://hdl.handle.net/10419/38701

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

THE CENTER FOR THE STUDY
OF INDUSTRIAL ORGANIZATION
AT NORTHWESTERN UNIVERSITY

Working Paper #0020

Consumer Learning about Established Firms: Evidence from Automobile Insurance

By

Mark Israel*
Northwestern University

First Draft: April 23, 2001
This Draft: July 2003

*Assistant Professor of Management & Strategy, Kellogg School of Management, Northwestern University. 2001 Sheridan Road, Evanston IL, 60208. Email: m-israel@kellogg.northwestern.edu. This paper is a substantially revised version of the third chapter of my Ph.D. thesis. I thank Tim Bresnahan, Peter Reiss, Roger Noll, Shane Greenstein, Scott Stern, Michael Mazzeo, Ariel Pakes, and Glenn Ellison for numerous suggestions. Any remaining errors are, of course, my own.

Visit the CSIO website at: www.csio.econ.northwestern.edu.
E-mail us at: csio@northwestern.edu.

Abstract

Most research on experience goods embodies the notion that, while direct product experience is required to learn about new goods, information is more complete for established products. This view is supported, at least in part, by three premises – that learning from direct product experiences occurs rapidly; that a consumer's preference for a given firm increases with information (so that firms have strong incentives to disseminate information), and that consumer purchase choices react strongly to that information. However, officials in many industries question these views – arguing that limited consumer information impacts demand even for well-established products, that learning from direct experiences can be quite slow, that consumers are often initially optimistic and then disappointed by experiences, and that by the time consumers learn they may be too "locked-in" to react. Unfortunately the empirical measurement required to settle these issues is nearly impossible in the standard, non-durable product markets generally studied – if consumers learn each time they purchase a product, it is quite difficult to separate learning from other sources of state dependence in demand. Markets for continuously provided services, such as credit cards, telephony, or insurance, are potentially much better venues for such measurement, because consumers learn about service quality at distinct interactions with firms. Unfortunately, the occurrence of these interactions tends to be either endogenous or unobservable. This paper overcomes these problems by considering automobile insurance, where consumers learn about service quality each time they have a claim, and the occurrence of claims is completely distinct from a consumer's satisfaction with her firm and fully observable from company records. Using a panel of 18,595 consumers from one well-established auto insurance company, the paper estimates a structural model of consumers' departure decisions with an imbedded Bayesian learning model. Among the key findings are: patterns of consumer departures by age and claims experience strongly suggest the importance of consumer learning at a long-standing firm; consumers enter the firm optimistic about its quality and are generally disappointed by experiences; and the impact of learning is greatly mitigated by the slow arrival of claims and the accrual of consumer lock-in over tenure with one firm.

I. Introduction

The long-standing literature on experience goods establishes that limited information about product quality can lead to under-provision of high quality goods and a general loss of consumer welfare (Stiglitz, 1989). However, this literature also suggests that informational inefficiencies primarily affect new goods, while information is more complete for well-established products. Theoretical work considers how the development of firm reputations can overcome information failures; empirical work focuses on learning about newly introduced products; and strategic analyses worry about the difficulties faced by late entrants in persuading consumers to try their products. Following this lead, Xiao (2003) uses a firms' age as a direct measure of consumers' knowledge of its quality.

This seemingly common-sense notion that, after some time in the market, a product's quality is well-known is bolstered by three generally accepted premises.[2] First, if they are uninformed prior to purchase, consumers learn quickly from direct product experiences. This not only implies rapid direct learning by those consumers, but also supports equilibria in which firm's can perfectly signal their qualities, and suggests rapid social learning via word of mouth. Second, risk averse consumers prefer products about which they have more information, so that firms have an incentive to disseminate product information as quickly as possible. Finally, consumer purchase choices react strongly to information, suggesting sufficient demand to support "information markets" to fill any remaining gaps.

However, this notion stands in contrast to views expressed by officials in many industries, who see limited consumer information as a prominent feature of demand, even for long-standing firms. For example, an anonymous executive at a well-established credit card firm has argued:

> Honestly, despite all the lip-service it gets, we don't find much benefit to investments in customer service technology or support. I think most of our new customers sort of assume they'll get good service, and by the time they learn otherwise…well, let's just say we're not losing a lot of business at that point.

In addition to suggesting limited upfront information, even about a long-standing firm, this view disputes the listed premises on several dimensions. First it implies that new customers have excessively optimistic impressions of products upon initial purchase, and thus tend to be disappointed by actual experiences. This should not be surprising – if there is variation across consumers in expectations about a given firm's quality, such that the *overall* expectation of that distribution is unbiased, the expectation among those who *select* into a given firm will surely be too high. But it contrasts with the typical view that observed preferences increase with experience, and, if correct, it limits firms' incentives to propagate information

[2] This refers to experience goods, for which quality can be learned (perhaps slowly) after purchase. Credence goods, for which quality can not be verified after purchase, raise an entirely separate set of issues.

about their true qualities.[3] Second, it indicates that learning can be quite slow after purchase. Finally, it implies that there may be a "race" between learning and lock-in at any given firm, and that many consumers may be too locked-in to react to information, yielding limited demand for product information.

These contrasting view raise a set of empirical questions that this paper takes up, for the specific case of a well-established automobile insurance firm. Most basically, can we find convincing evidence for consumer learning from direct experiences at this firm? On a related note, is there evidence for learning from other, indirect sources, measured here by increased accuracy of consumer expectations with age? Then, if there is evidence for consumer learning, what is the nature of the learning process? Does a typical consumer's preference for the firm increase or decrease with direct product experiences? After joining a firm, how quickly do consumers learn from direct service experiences? How much impact does learning have on departure decisions? To what extent is this impact limited by consumer lock-in, which may grow over a consumer's tenure with the firm?

Empirical work on this set of questions is quite limited. This reflects several features of the non-durable product markets – say laundry detergent or breakfast cereal -- generally studied using available "scanner data." First, to be sure that all a consumer's experiences with a given product are observed, these papers focus only on new products. Second, for such products, consumer learning essentially adds one more explanation for "tenure dependence" in demand – the tendency of consumers to prefer products they have previously purchased – to a list that already includes switching costs, search costs, unobserved heterogeneity, brand loyalty, etc. To separate learning from these explanations either leans heavily on extremely subtle patterns in the data or (quite often) relies on functional form restrictions, neither of which can provide evidence convincing evidence to challenge standard views about experience goods. Finally, because learning is linked directly to strong patterns of tenure dependence in demand, answers to this paper's questions are "rigged" – we know we will find rapidly increasing preferences with experience, and we can not contrast learning with the development of lock-in over tenure.

Markets for continuously provided services – such as insurance, telephony, or cable television -- provide a more promising venue for this measurement. Datasets tracking consumers' entire relationship with particular firms are readily available. Even more importantly, consumer purchases in these markets are characterized by long-term relationships with particular firms, punctuated by distinct "learning events" at which consumers observe service quality – what marketers call "moments of truth" – such as calls to a service center or visits to a doctor. Empirically, then, patterns of consumer purchase choices

[3] Especially since a firms' current consumers may be the most likely to pay attention to such information.

before and after these events should provide a powerful, and previously untapped, source of information about the importance of learning and the characteristics of the learning process.

Unfortunately, two factors have limited the use of such information. First, data on the occurrence of service events are hard to find. Even when data tracking consumer-firm relationships are available, they generally do not include visits to a bank teller or calls to the phone company. Second, the occurrence of service events is generally endogenous and, most importantly, linked to departure decisions. That is, a consumer is most likely to contact his firm at exactly those moments when he is unhappy, and those consumers who frequently contact their firms are those who tend to be unsatisfied. In either case, learning effects are confounded with unobservable consumer characteristics.

This paper overcomes these difficulties by utilizing claims as service events in automobile insurance. A key feature differentiating insurance companies is the quality of their claims service, including the speed of claims resolution, the value of repair advice, and the courteousness of company officials. Each of these is hard to observe until experiencing a claim. Further, claims data are readily available from company records. And, crucially, claims arise completely separately from departure decisions – each time she has an accident, a consumer receives a draw on the company's service quality and learns accordingly.[4] So this paper relies on the pattern of consumer departures from one automobile insurance firm, in the periods before and after successive claims, to measure the importance of learning and the characteristics of the learning process.

To implement this measurement strategy, this paper considers the claims experience and departure decisions of nearly 19,000 consumers who joined one well-established auto insurance firm in Georgia between 1991 and 1998. The service provided on any given claim is modeled as either satisfactory or unsatisfactory, with the firm's fixed probability of providing good service -- referred to throughout as the firm's quality – denoted q_c.[5] Consumers enter the firm with a prior on this probability – which is based on their knowledge of the firm's reputation and thus allowed to improve with age – and update the prior via direct claims experience. The departure probability is derived as a function of claims

[4] One may argue that claims are endogenous due to moral hazard This has been found to be relatively unimportant in recent work (Chiappori and Salanie, 2001) and is assumed away for this study. Note, that all we really need here is that care in driving is not correlated with the departure decision, which seems quite plausible. Alternatively, one may argue that claims are endogenous due to adverse selection – certain consumers are simply higher risk, and this risk may correlate with unobservables in the departure decision. Such correlation **is** controlled for in all analysis.

[5] This service quality is assumed to be the outcome of an unmodeled quality choice game. That is, the firm invests in quality; these investments determine the **probability** that it provides good service on **any given claim,** and then consumers join the firm and learn about this fixed probability from the service they receive. The assumption of a fixed probability across consumers is justified by the fact that claims service employees have essentially no information about the consumer's rating class, history with the firm, etc., when processing claims.

experience, along with consumer characteristics, prices from the study firm, and average prices by consumer type in the broader market, using a structural model of consumer choice that embeds a Bayesian learning model. Estimation proceeds by maximizing the likelihood of observed departures.

Estimating such a learning model is complicated by the fact that, while we see the occurrence of claims, we do not observe a consumer's satisfaction with the service provided. But, with the panel data structure, we do see consumers' decisions to remain with or depart the firm in the periods following a claim. As time passes, consumers who were unsatisfied are more likely to depart, leaving a growing proportion of those who received good service at the study firm, and thus a falling departure probability. And it is well-established that the underlying heterogeneity distribution can be recovered from this "post-claim tenure dependence" in departures (Heckman, 1991). An important methodological contribution of this paper is to demonstrate that -- because this heterogeneity is generated by consumer reactions to observed claims service, and specifically the deviation of this service from prior expectations -- we are able to recover the information and learning parameters of interest from our knowledge of this heterogeneity distribution, even without directly observing consumer satisfaction or firm quality.

A final complication is that firms may learn about consumers during the same interactions in which consumers learn about firms and, in response, change the terms of the relationship in a way that impacts departure decisions. In the current context, claims inform firms about consumers' driving safety, and this learning is explicitly reflected in current and future prices. Fortunately, we have two ways to disentangle firm and consumer learning. First, 55% of all claims are "non-chargeable" – events such as storm damage, roadside breakdowns, or accidents caused by uninsured motorists – which are not the insured's fault, do not reflect his underlying driving safety, and thus do not impact his prices.[6] So, non-chargeable claims are quite powerful -- random events, outside the driver's control, and thus providing no information to the firm, which nevertheless provide the consumer with a draw on the firm's quality. Second, the price effects of chargeable claims are fully observable. So, after controlling for these price effects, any additional impact of chargeable claims can also be used to infer the learning parameters.

The remainder of the paper proceeds as follows. Section II briefly reviews related literature. Section III describes the dataset of insurance consumers used in analysis, along with the relevant institutional features of the Georgia insurance market. Section IV develops the model of consumer departures and the associated learning model. Section V clarifies how the data permit identification of this learning model. Section VI presents the results. Section VII contains extensions designed to test robustness. Section VIII concludes.

[6] These do not include accidents caused by another insured driver, which are covered separately under her insurance.

II. Literature Review

Existing literature has established that limited consumer information on product quality can lead to many market imperfections, including undersupply of quality and consumer welfare loss, For example, Shapiro (1982) demonstrates that if consumers cannot perfectly observe quality before purchase, profit-maximizing firms will undersupply quality, and "the extent to which imperfect information causes quality deterioration depends critically on the speed with which consumer learning occurs." Such findings suggest the need for better empirical work , measuring the quality of consumer information and the speed of learning.

In light of these potential market failures, the early theoretical literature on experience goods focused on mechanisms through which the market could endogenously restore full information, with no need for direct product experience. For example, Nelson (1974) suggested that firms could use advertising as a signal of high quality; Kihlstrom and Riordan (1984) formalized this notion, and Milgrom and Roberts (1986) showed that the signal could also take the form of a low initial price. However, the conditions supporting such separating equilibria are extreme, requiring among other things that quality is revealed immediately after purchase (Horstmann and MacDonald,1994).   If, instead, information arrives slowly -- say only with infrequent insurance claims -- then limited information is more likely to be a persistent feature of market equilibria.

Still, while signals may be imperfect, the extensive literature on firm reputations generally finds that, over time, consumers become more informed, and any inefficiencies from imperfect information fade (Jovanovic, 1982; Bagwell and Riordan, 1986).   This conclusion, however, rests on the fact that firms have an incentive to disseminate information about their qualities, something which may not be true if their consumers tend to be excessively optimistic.

In an empirical context, Xiao (2003) begins from the premise that reputations spread as firms age. So, upon finding that formal quality certification has a greater impact on the demand for new child care firms than established ones, she concludes that the information failures are most important at new firms. The current paper, in contrast, uses micro data to demonstrate that consumers have very limited information about a well-established firm. And it suggests a reconciliation of this finding with Xiao's. If, for any reason, consumers become locked in as they remain with one firm, and if much of the business at well-established firms comes from long-term consumers, then the aggregate demand at such firms will show limited reaction to *any source of information*. And this limited reaction from the bulk of consumers also suggests why demand for quality certification services may be limited, even given imperfect information about those quality attributes that consumers value.

Aside from Xiao's paper, most empirical studies of consumers learning consider non-durable product markets, relying on scanner data. Leading papers in this vein are Erdem and Keane (1996) and Ackerberg (2002), which consider laundry detergent and yogurt, respectively. As noted above, separating learning from other sources of state dependence in demand is quite challenging in this context, relying on extremely subtle patterns in the data plus the functional form restrictions in a Bayesian learning model. This is not a major concern for these authors, whose focus is on the informational content in advertising. But to meet this paper's objective – determining whether learning is an important phenomenon at a well-established firm -- requires more compelling sources of identification.

Finally, Tom Hubbard's 2002 paper is quite close to the current study in spirit, as it uses micro data to study learning in a service market --vehicle inspections in California -- modeling consumers as preferring shops which are more likely to give them a "passing grade." However, the key to identification in this paper is that Hubbard observes both the overall percentage of passing grades given by different firms and the outcomes on particular inspections for a set of consumers. So, he relies on the relative impact of these two measures of "quality" on subsequent purchase decisions to measure the impact of a firm's reputation vs. learning from direct experience. This method, while very powerful in his context, will not work in those experience good markets where we can not observe a consumer's satisfaction with the service at any particular event nor the firm's overall probability of providing good service. The current paper considers the measurement of learning in such an environment.

III. Auto Insurance Data

To better understand the modeling choices made in Section IV, it useful to have a clear picture of the available data. Analysis relies on a panel of 18,595 consumers joining one Georgia auto insurance firm between October 1991 and December 1998. Each consumer is observed beginning with her initial purchase from the firm. 31.5% of these consumers voluntarily depart the firm during the sample period, presumably by switching to another firm.[7] All other consumers either randomly fall off the sample or survive through 1998.[8] Table 3.1 presents details on entry and exit from the sample.

The record for each consumer is divided into 6-month periods, the length of an auto insurance policy. The dependent variable for all analysis is an indicator for whether the consumer completes the period with the study firm, which she does in 92.7% of all observed periods.

[7] Auto insurance is mandatory in Georgia. For simplicity, we ignore both violations of this law and decisions to stop driving altogether . This means that a decision to depart the study firm is equivalent to a decision to switch insurers.
[8] Consumers randomly fall off the sample if they move to another state, or for a variety of administrative reasons which lead them to be assigned a new policy number.

Explanatory variables include the *rating class* variables used to compute price – age, gender, marital status, vehicle usage, vehicle characteristics, zip code, etc. -- which are assumed to be exogenous. Data also include all details of the consumer's policy – in addition to the core liability policy required by law, consumers can choose to purchase extra liability coverage and/or "first party" coverage for damage to their own person or vehicle. For tractability these policy choices are assumed to be based on the consumer's underlying insurance needs, and thus to be exogenously specified when considering the departure decision.[9] Of particular interest is the amount of liability coverage purchased, as the study firm indicates that these "liability limits" are highly correlated with consumer income. So, in analysis, we use a categorical measure of liability limits, $y_t \in \{0,1,2,3\}$, as a proxy for income. Collectively, a consumer's rating class variables and policy choices in period *t* are referred to as x_t.

Most importantly, the panel also includes data on the complete claims experience for each consumer with the study firm and a summary of claims experience prior to joining this firm.[10] Table 3.2 contains a frequency tabulation of total, chargeable, and non-chargeable claims for the 85,931 consumer-periods on the panel. Because periods with more than 1 claim are very rare, analysis relies on categorical variables for 0 or 1^+ chargeable and non-chargeable claims in the period, denoted as l_t, and n_t, respectively.

In addition, we use three summaries of claims history. First, since learning accrues with each claim at the firm, a key variable is the total number of claims experienced at the study firm, C_t. The last panel of Table 3.2 contains a frequency distribution for this cumulative claims experience. Second, to compute a consumer's price, firms rely on the number of chargeable claims she has experienced over the last 3 years, L_t.[11] Finally, prices also rely on τ_t^{nc}, the number of consecutive periods with the current firm with no chargeable claims, as firms give discounts after 6 and 12 such periods.

Prices for all firms in Georgia are subject to "prior approval" regulation. Each firm's menu of prices – as a function of rating class and coverage, chargeable claims in the last 3 years, and tenure at one firm with no chargeable claims -- is negotiated with the state insurance commission, which imposes

[9] More formally, we assume that consumers choose an optimal policy based on their insurance needs and the prices at their *current* firm. Then, when considering departures, they compare this pre-set policy across firms. In this way, the two decisions are tractably separated into two problems and we consider only the departure portion.

[10] Note that data are only available for the consumer's automobile insurance experience. So if consumers learn about the firm from homeowners or other lines of insurance, this is not captured. However, the firm maintains completely separate divisions for these different lines – with strikingly little contact across the divisions – so it seems reasonable to assume that the learning processes are separate.

[11] The dataset includes each consumer's chargeable claims history for 3 years prior to joining the firm, so we are able to compute L_t for all observations. Note that in Georgia, as in most other states, insurance companies share claims data in a central database, so it's safe to assume that they know the claims history for the years prior to joining the firm, and that consumers can't switch firms to flee a bad record.

substantial restrictions on permissible prices.[12] The prices arising from this process are taken as exogenous. The study firm's entire menu of prices, $p_c(x_t, L_t, \tau_t^{nc})$, is included in the data. Prices from "outside" firms in the broader market are summarized by a measure of the average price by rating class and claims history, $p_m(x_t, L_t)$.[13] Additional details on computation of this average price are contained in Appendix I. Table 3.3 summarizes these price data. While the study firm's prices are somewhat lower than average, it is certainly not seen as a "discount" company.

It is important to be clear that, because all consumers in the data have selected the study firm, this is not a random sample of insurance consumers. However, three factors lessen this concern here. First, the study firm's "internal market research" indicates that their consumer base is "quite similar" to the overall set of insurance consumers. Second, this firm's market share and entry rates were very stable over the study period, so changes in the entry process should not be driving the results. And, most importantly, while sample selection affects results for variables that impact the unmodeled entry decision, the primary variables of interest here are post-entry claims, which are only relevant for the estimated exit decision.

IV. Model

The core logic of the model is that consumers update preferences for their current firm using information from claims experiences, and then compare these updated preferences to "offers" drawn randomly from the broader market to decide whether to depart. This offer structure is used as a simple way to incorporate the commonly held view that consumers have limited information about alternate insurance firms (Joskow, 1973). It also provides a simple source of econometric randomness – observationally equivalent consumers may make different choices due to different, random offers.

A. Claims Arrival

As noted above, consumers are assumed to have 0 or 1 chargeable claims per period, denoted l_t, and 0 or 1 non-chargeable claims, n_t. The associated probabilities are given by:

$$\begin{aligned} \Pr(l_t = 1) &= \lambda^l(x_t;\eta) \equiv \lambda_t^l \\ \Pr(n_t = 1) &= \lambda^n(x_t) \equiv \lambda_t^n \\ \lambda^l(x_t;\eta) + \lambda^n(x_t) &\equiv \lambda(x_t;\eta) \equiv \lambda_t \end{aligned} \tag{4.1}$$

[12] Regulators' main objective in this process is to achieve prices which are not "inadequate, excessive, or unfairly discriminatory." In practice, this results in a "cost-plus' type restriction. However, because each firm uses its own actuarial data and methods, and for a variety of other institutional reasons, the process leads to prices which vary greatly across firms, with the standard deviation of prices for a given rating class roughly 25% of the mean.

[13] Note that $\tau_t^{nc} = 0$ for any potential alternate firm by definition. Also note the key assumption here that all firms set prices based on the same classifications, which is a reasonable approximations to the Georgia insurance market. .

where η is a fixed, individual specific claims risk error, unobservable to both the firm and the econometrician, assumed to impact only the risk of chargeable claims.[14] The exact functional forms for $\lambda^l(x_t;\eta)$ and $\lambda^n(x_t)$ are specified in Appendix II and estimated jointly with the departure probability.

B. Preferences

To define preferences for the current firm, begin by specifying a consumer's ex-post utility, conditional on choosing her current firm, f_c, and whether she experiences a chargeable and/or a non-chargeable claim in period t:

$$U(f_c) = h(y_t) - \alpha_1(y_t)p_{c,t} - \alpha_2(x_t)[l_t(1-g_{l,t}) + n_t(1-g_{n,t})] + f_c(x_t, \tau_t, \tau_t^{nc}, \eta) \tag{4.2}$$

The first term is a general function for utility given current income. This is reduced by the price paid for insurance, with the impact of price allowed to vary with income.[15] Utility is further reduced by each chargeable or non-chargeable claim experienced, *unless the firm provides satisfactory service on the claim,* with satisfactory service on a chargeable or non-chargeable claim denoted by $g_{l,t} = 1$ or $g_{n,t} = 1$, respectively. For simplicity, assume that all claims are fully paid, so that the "cost" of the claim does not depend on the amount of the loss. Instead, this is a pure service quality term -- the firm can provide either satisfactory or unsatisfactory service on any given claim, with satisfactory service eliminating the "inconvenience" cost of the claim. We allow this cost to vary by consumer type as $\alpha_2(x_t)$, with the exact form of this term and its estimates presented in Section VI.

Finally, utility is increased by the term $f_c(x_t, \tau_t, \tau_t^{nc}, \eta)$. This includes consumer characteristics and the state of the relationship with the current firm, both of which may influence a consumer's preference to remain with her current firm rather than switch.[16] x_t is included to allow for variation in switching costs by consumer type. The number of periods spent with the study firm, τ_t, is included to

[14] This follows from the definition of a non-chargeable claim. If η impacted the occurrence of non-chargeable claims they would be informative about its value and thus used by the firm in setting prices.
[15] To be completely formal, the first two terms should actually be combined as a function of income less price, $h(y_t - p_{c,t})$. However, with only a categorical measure of income, estimation of such a function is not feasible. So, this linear approximation is used instead.
[16] Recall that we are assuming that all consumers purchase auto insurance in all periods. So, there is no outside good, and the relevant comparison in estimation will be between the current firm and potential alternate firms. Hence, all utility calculations are really comparisons between firms, and $f_c(x_t, \tau_t, \tau_t^{nc}, \eta)$ captures both positive things about the current firm and negative things – such as switching costs – about moving to any other firm.

allow lock-in to grow with tenure.[17] The last two terms control for alternate impacts of claims, to ensure that they are not confounded with learning. As consumers accrue periods with one firm with no chargeable claims, they move closer to the claims-free discounts described above, which may increase their preference for the current firm. So the value of τ_t^{nc} is included. And preferences are allowed to vary directly with the claims risk error, η. The exact form of $f_c(x_t, \tau_t, \tau_t^{nc}, \eta)$ is presented, along with its estimates, in Section VI.

In choosing a firm, consumers do not know if they will experience a claim, nor what service they will receive if they do. So, in practice, their choice depends on the expected utility function – defined as the mathematical expectation of $U(f_c)$ with regard to the occurrence of a claim and the service received -- which after some rearranging can be written as:

$$\begin{aligned} &E_t(U(f_c)) = h(y_t) - \lambda_t \alpha_2(x_t) \\ &+\lambda_t \alpha_2(x_t) E(q_c \mid x_t, C_t, G_t) - (\alpha_1(y_t) p_{c,t} - f_c(x_t, \tau_t, \tau_t^{nc}, \eta)) \end{aligned} \tag{4.3}$$

The key assumption used here is that the firm's probability of providing satisfactory claims service is fixed at q_c for all consumers and all claims.[18] A consumer's expectation of this probability is based on her characteristics, x_t – which impact her pre-entry learning -- and her claims experience, summarized by the total number of claims experienced, C_t, which is observable, and the number of claims with satisfactory service, G_t, which is not.[19] The exact functional form of this expectation is derived next.

C. Learning

Consumer learning is assumed to be Bayesian. We begin with a description of the learning process for consumers with minimal information prior to joining the firm, then add the possibility of pre-entry learning that accrues with age. The minimally informed consumers enter with a Beta prior on the firm's quality:

$$\begin{aligned} &q_c \sim Beta(a,b) \Rightarrow \\ &f_0(q_c) \propto (q_c)^{a-1}(1-q_c)^{b-1} \end{aligned} \tag{4.4}$$

In practice, we rely on a one-to-one transformation of the Beta parameters:

---

[17] Actually, preferences may change with τ_t for two reasons – either because an individual consumer's preferences grows with tenure (true tenure dependence) or because consumers have heterogeneous preferences and those that like the firm more are more apt to remain (unobserved heterogeneity). Israel (2002) considers the separate identification of these two effects. That is not the focus here, so we simply include a general function of tenure.
[18] Section VII tests the robustness of the results by relaxing this assumption.
[19] Any claims in period t are assumed to happen at the **start** of the period, so that all decisions during the period are conditional on the learning from those claims.

$$E_0(q_c) \equiv \mu_0 = \frac{a}{a+b} \tag{4.5}$$

$$\psi_0 = a+b$$

Consumers update this prior with the Bernoulli draws on claims service at each claims experience. So, it is easy to show that for any level of claims experience:

$$E(q_c \mid C_t, G_t) = \frac{\psi_0 \mu_0 + G_t}{\psi_0 + C_t} \tag{4.6}$$

This provides a clean interpretation. Before joining the firm, consumers are endowed with some "knowledge" of its quality, with that knowledge taking the form of ψ_0 "draws" of an expected quality $\mu_0 \; \varepsilon \; [0,1]$. Then after joining, each claim provides them with another draw on quality, $g \; \varepsilon \; \{0, 1\}$. Their expectation given any level of experience is just the average of all these pre and post entry draws.

This form provides a natural way to allow pre-entry learning to accrue with age, a_t. Before entering consumers are assumed to have received $\psi(a_t)$ draws on the firm's true service quality – through word-of mouth, explicit signals, or any other sources of information on the firm's reputation. Since, the youngest drivers on the dataset are 16, the number of draws is defined as:

$$\psi(a_t) \equiv \psi_1 * (a_t - 16) + \psi_2 * (a_t - 16)^2 \tag{4.7}$$

So, all consumers enter the firm with a Beta prior, but now it is updated with these pre-entry draws, with its parameters adjusted in the natural way as:

$$\begin{aligned} \psi_p &\equiv \psi_0 + \psi(a_t) \\ \mu_p &\equiv \frac{\psi_0 \mu_0 + \psi(a_t) q_c}{\psi_p} \end{aligned} \tag{4.8}$$

This implies that the expected claims service for any level of claims experience is:

$$E(q_c \mid x_t, C_t, G_t) = \frac{\mu_p \psi_p + G_t}{\psi_p + C_t} \tag{4.9}$$

This is the equation which enters the expected utility function (4.3).

D. External Offers and Departure Probability

Decisions about whether to depart one's current firm are triggered by offers from the external market. Since the focus here is not on the consumer search process, the arrival of offers is assumed to be exogenous. However, we define a "search event" as any claim or any change in a consumer's rating class, to capture the idea that consumers are more likely to search for alternate firms following such events. Periods following such a change (20.2% of all periods) are indicated by setting the variable $\Delta_t = 1$. Consumers receive Ω_t offers per period -- where an offer is simply a price quote drawn randomly from the distribution of market prices associated with the consumer's type, $p_\Omega(x_t, L_t)$ – such that:[20]

$$\Omega_t = \Omega_0 + \Omega_1 \Delta_t \tag{4.10}$$

An offer is a consumer's only specific information about any alternate firm.[21] Consumers evaluate offers according to an appropriately modified version of 4.3:

$$\begin{aligned} E_t(U(f_\Omega)) &= h(y_t) - \lambda_t \alpha_2(x_t) \\ &+ \lambda_t \alpha_2(x_t) \mu_\Omega - \alpha_1(y_t) p_{\Omega,t} \end{aligned} \tag{4.11}$$

μ_Ω is the consumer's expected claims service from alternate firms – it is assumed to be uncorrelated with price and thus fixed across offers. Similar to the expected quality at the study firm, we let μ_Ω evolve with age as:

$$\mu_\Omega = \mu_0 + \mu_1 * (a_t - 16) + \mu_2 * (a_t - 16)^2 \tag{4.12}$$

meaning that consumers start with the same uninformed expectation that they have for the study firm, μ_0, and then learn more about average quality in the market as they age.[22]

The distribution of price quotes is assumed to be Normal for all rating classes:

$$p_\Omega \sim N(p_m(x_t, L_t), \sigma_p^2) \tag{4.13}$$

where $p_m(x_t, L_t)$ is in the data, as discussed above. Since previous work has suggested that the standard error of prices is proportional to the mean, we model σ_p as $\overline{\sigma}_p p_m(x_t, L_t)$ and estimate the proportionality factor, $\overline{\sigma}_p$. This implies that the probability of departure in period *t*, denoted as $d_t = 1$, is equal to 1 minus the probability that all Ω_t offers are less attractive than the current firm:

$$D_t \equiv \Pr(d_t = 1) = 1 - \left[\Phi\left(\frac{p_m(x_t, L_t) - p_{c,t}}{\overline{\sigma}_p p_m(x_t, L_t)} + \frac{\delta_{c,t}}{\alpha_1(y_t)\overline{\sigma}_p p_m(x_t, L_t)}\right)\right]^\Omega \tag{4.14}$$

---

[20] In estimation, alternate specifications were considered. In particular, lags of Δ_t were tested, to allow search effort to continue for several periods after a search event, but none were even remotely significant. Alternate definitions of search events were also considered – requiring the event to be something more substantial, like a move, rather than any class change – but results were effectively the same. Also note that while the Ω parameters were left unrestricted in estimation, they remained positive throughout all iterations, avoiding the strange implications of a negative number of offers. This specification does allow a non-integer number of offers, which may seem strange, but simply generalizes the idea of a higher value as more information from the external market.

[21] In particular, this means that consumers have no recall of their previous firms.

[22] Consumers could be learning about the reputation of each firm in the market separately. As long as this reputation is uncorrelated with individual consumer's price quotes, all analysis goes through.

where Φ is the standard normal CDF, and:

$$\delta_{c,t} = \lambda_t \alpha_2(x_t)[E(q_c \mid x_t, C_t, G_t) - \mu_\Omega] + f_c(x_t, \tau_t, \tau_t^{nc}, \eta) \tag{4.15}$$

In estimation, $\alpha_1(1)$ is normalized to 1, while $\alpha_1(0)$, $\alpha_1(2)$ and $\alpha_1(3)$ are left free.

E. Likelihood Function

For ease of estimation, the claims risk error, η, is assumed to take on three possible values, η_0, η_1, or η_2, with probabilities ρ_0, ρ_1, ρ_2, respectively. This distribution is subject to the standard restrictions that $\rho_2 = 1 - \rho_0 - \rho_1$ so that the probabilities sum to 1, and that $\eta_2 = -(\eta_0\rho_0 + \eta_1\rho_1)/\rho_2$ so that the expected value is 0.

An individual's likelihood contribution is the joint probability of her observed claims and departure decisions. Consider two possible cases. First, suppose consumer *i* joins the firm in period *t* and *is observed* to depart during period t+s. Then we exclude claims experience in *t+s* from the likelihood function because the full period is not observed. So, the likelihood of this sequence is based on claims from *t* through *t+s-1,* along with *s-1* decisions to remain with the firm and 1 departure. So, using 4.1 and 4.14, we have:

$$l_i = \sum_{G_t^{t+s}} \sum_{j=0}^{2} \Pr(G_t^{t+s})\rho_j \left[(D_{t+s} \mid \eta_j, G_t^{t+s}) \prod_{r=0}^{s-1} (1-(D_{t+r} \mid \eta_j, G_t^{t+s}))(\lambda_{t+r}^l \mid \eta_j)^{l_{t+r}} (1-\lambda_{t+r}^l \mid \eta_j)^{1-l_{t+r}} (\lambda_{t+r}^n)^{n_{t+r}} (1-\lambda_{t+r}^n)^{1-n_{t+r}} \right] \tag{4.16}$$

where G_t^{t+s} is a particular sequence of unobservable service qualities on the claims experienced by the consumer, $\Pr(G_t^{t+s})$ is the probability of this particular sequence given q_c, and we sum over all possible sequences and all possible values for η.

Second, suppose consumer *k* joins the firm in period t, remains through period *t+s-1*, but is right-censored in period *t+s*. In that case, period *t+s* contributes nothing to the likelihood, and we have:

$$l_k = \sum_{G_t^{t+s}} \sum_{j=0}^{2} \Pr(G_t^{t+s})\rho_j \left[\prod_{r=0}^{s-1} (1-(D_{t+r} \mid n_j, G_t^{t+s}))(\lambda_{t+r}^l \mid \eta_j)^{l_{t+r}} (1-\lambda_{t+r}^l \mid \eta_j)^{1-l_{t+r}} (\lambda_{t+r}^n)^{n_{t+r}} (1-\lambda_{t+r}^n)^{1-n_{t+r}} \right] \tag{4.17}$$

Estimation proceeds by maximizing the sum of the log of the likelihood contributions for each of the 18,595 individuals on the panel.

Finally, note that while the model has many pieces, it produces simple functional forms. All probabilities are based on Normal CDFs, so at core this is a multivariate probit. The dependent variables are independent save for the explicit cross equation restrictions, capturing the fact that consumers' valuation on claims service depends on claims risk, and allowing the departure probability to vary directly with the claims risk error. While it may seem like a complication, the offer structure actually generalizes the estimating equations, by allowing the departure probability to take the form of a Normal CDF *raised to any power*.[23] And while the Bayesian learning model is somewhat restrictive, the expected claims service term which it produces is quite intuitive.

V. Identification

The paper's primary goal is to identify the impact of changes in a consumer's expected service quality, resulting from claims experience. This impact is given by:[24]

$$\alpha_2(x_t)E(q_c \mid x_t, C_t, G_t) = \alpha_2(x_t)(\frac{\mu_p \psi_p + G_t}{\psi_p + C_t} - \mu_p) = \alpha_2(x_t)(\frac{G_t - C_t \mu_p}{\psi_p + C_t}) \tag{5.1}$$

That is, the change in expectation is a function of the difference between the observed number of claims with good service and the number which would have been expected under the prior, with the impact of this difference a function of the faith in the prior and the value on good claims service.

The dataset has two features which are especially helpful here. First, the panel structure means that we observe departure decisions for several periods following each claim. Second, we observe multiple claims for many consumers. A convenient way to think of the data, then, is as a set of departure decisions before any claims, plus a set of departure decisions as tenure accrues after a first claim, plus a set of departure decisions as tenure accrues after a second claim, etc.

Identification of q_c, $\alpha_2(x_t)$, and μ_p relies on the first feature. Suppose for now that ψ_p is known, and consider the form of 5.1 following the first claim. Because G_t is not known, this is an error term with two points of support – for satisfactory or unsatisfactory service -- $\alpha_2(x_t)(\frac{1-\mu_p}{\psi_p + 1})$ or $\alpha_2(x_t)(\frac{-\mu_p}{\psi_p + 1})$. The actual value of this term will "play out" in the periods following this claim. As consumers who received poor quality depart, the proportion of those who received good service grows, and thus the departure probability falls. Crucially, a long line of research (see Heckman, 1991, for a survey of this

---

[23] In addition, we have tested robustness by using alternate forms for the market distribution of prices, including logistic and chi-square, with no changes in the substantive results.

[24] Since λ_t is identified directly from the occurrence of chargeable and non-chargeable claims, it is not included in this discussion.

work) establishes that we can recover the heterogeneity distribution from this pattern of "post-claims tenure dependence."[25]

Here, this means that we can recover the probability of satisfactory service, q_c, our first parameter of interest, plus the two points of support. For any given x_t, these two points of support map directly into the other two parameters of interest, $\alpha_2(x_t)$ and μ_p, as they are the only other free parameters. In essence, μ_p (the prior expectation) determines the *relative* impact of good and bad claims. If it is close to 1 then seeing a good claim has little effect while seeing a bad claim leads to a substantial downward revision, and as it moves toward 0 this pattern reverses. So, μ_p is identified by the relative magnitude of the two points of support. $\alpha_2(x_t)$ is a multiplier on the error term, increasing the *absolute* value of its effect, as we'd expect since it measures the value consumers put on good claims service.

This leaves only ψ_p. This is the model's "speed of learning" parameter -- the lower ψ_p the less certain consumers are about their prior expectation and thus the more willing they are to learn with experience. Here, identification relies on the second helpful feature of the data, observation of a series of claims. Low certainty about the prior expectation implies that the first few claims are particularly informative relative to later claims. And we see this in 5.1 – if ψ_p is close to 0 then the first claim has a denominator of 1 and thus has 2 times the impact of the second claim, which has 1.5 times the impact of the third claim, etc. As ψ_p increases, the relative impact of successive claims becomes much flatter. So, while previous work has relied on the relative impact of information observed before and after joining a firm to measure faith in the prior, this paper relies on the impact of early claims relative to later ones.

The variation in both μ_p and ψ_p with age is directly identified from corresponding variation in the data. In particular, the impact of age can be identified from the slowing of the learning process – and thus the flattening of the impact of successive claims – alone. With this in hand, the impact of age on expected service quality from other firms, μ_Ω, is identified by changes in the overall departure probability.

If consumer learning were the only channel through which claims could impact departures, this argument would be complete. However because identification proceeds by backing out learning

[25] To be precise, it indicates that we can recover either this heterogeneity distribution **or** the "true tenure dependence," through which the behavior of a given individual, who has received a particular quality of service, depends on the number of periods since a claim. Since we have no reason to include any such "true post-claim tenure dependence," given that we are already controlling for the overall tenure with the firm and the tenure with no chargeable claims, we can directly recover the heterogeneity distribution.

parameters from the observed impact of claims, it is crucial that we control for all other potential impacts. The most obvious alternate impact is that claims cause consumers to "pay attention" to their auto insurance policies, so consumers may be most likely to search and depart in the periods following a claim, even with no learning. This is why the search event indicator, Δ_t, is included in the offer arrival function. As a result, learning is identified from the extra impact of claims relative to other search events. Most importantly, learning implies that consumers update their preferences, an effect that will have an impact on departure rates long after the increased search (which is estimated to last only 1 period) has subsided.

In addition, the model explicitly allows for four other potential impacts of claims: changes in prices from the study firm, changes in market average prices, proximity to price discounts based on τ_t^{nc}, and correlation between chargeable claims risk and departures. These can be separately identified from learning in two ways – by relying on non-chargeable claims which do not contain information on a consumer's underlying claims risk and thus do not impact prices, and by using the known functional forms for the impact of chargeable claims.

VI. Results

Table 6.1 reports all parameter estimates and standard errors, along with the functional forms for $f_c(x_t, \tau_t, \tau_t^{nc}, \eta)$ and $\alpha_2(x_t)$. Before turning to the learning process, note that our focus on service quality appears justified by the high estimated consumer valuation. We specify the cost of a claim (without good service) as $\alpha_2(x_t) = \alpha_2^0 + \alpha_2^1 * EX_t$, where EX_t is an indicator for the purchase of "extra insurance" above the minimum liability policy. This allows for two types of consumers, those who simply purchase enough insurance to meet the legal requirement (18.9% of observations) and those who appear to actually be seeking additional services (81.1% of observations). The parameter estimates support this distinction, with a base value of 823.41 plus 232.23 for those purchasing extra insurance. Because the sample average price coefficient is 0.998, we can effectively interpret these directly as dollar values. So, the average consumer faces $1011.75 in costs from a claim unless he receives good service.

In fairness, the actual effect on preferences across firms is multiplied by λ_t – the expected number of claims per period – whose sample average value is only 0.084. So, ex-ante, a typical consumer would pay $84.99 for certain good service rather than certain poor service. And more importantly for observed demand, he'd pay an extra $8.50 for a 10 percentage point increase in the probability of receiving satisfactory service. For comparison, the sample average policy price is $148.11, so a typical consumer would pay 5.7% of the policy price for a 10 point increase in firm quality.

A. Is There Evidence for Learning at this Well-Established Firm?

Table 6.2 details consumers' priors on the firm's quality, along with their expectation of the quality available in the broader market, both as functions of age. The general picture which emerges is one of very limited information. For example, upon joining the firm, a 40 year old's expectation is 16 percentage points away from the true probability of good service, 0.739. And his confidence in this prior is quite limited, equivalent to just more than 1 direct claim observation.[26]

The estimates do suggest some pre-entry learning, perhaps through word of mouth or general market experience. Young drivers joining the firm have prior information equivalent to less than 4/5 of a single claim, yielding expectations which are off by more than 25 percentage points. Over the course of their lives, they update this expectation with information equivalent to approximately ¾ of a claim. Overall, though, the amount of pre-entry learning is limited -- by age 80 a consumer's prior expectation remains nearly 14 points away from the true quality.

Most importantly, the low value of ψ_p suggests the potential for rapid post-entry learning with claims experience. Table 6.3 illustrates this point by considering a 40 year old consumer's expectation of the firms' quality as a function of total claims experienced, C_t, and claims which received good service, G_t. After one claim, the expected probability of receiving good service ranges from 50% to 95%, depending on the service received. After 5 claims, the range is 18% to 98%.

This table also displays the relative impact of successive claims. For any given experience, represented by a cell in the table, the next claim can have two effects -- bad service pushes the individual straight down one row, while good service pushes him diagonally down and to the right. So the impact of the service received on each claim is the difference in the expected probability of adjacent cells. For the first claim, this impact is roughly 45 percentage points; it declines to roughly 29 points for the second claim, and continues to decline to 16 points for the fifth claim. This steeply declining path is a characteristic learning pattern, which is harder to justify with alternate explanations for the impact of claims. And the fact that ψ_p increases with age means that this path flattens for older consumers, suggesting that they know more at entry and thus further supporting the argument for learning.

Still, the most compelling alternative explanation for the impact of claims is that they remind consumers to pay attention to their auto insurance policies, inducing "post-claims search" for alternate firms. Recall that the model controls for this by allowing the arrival of offers to increase following a search event, defined as a claim or any change in a consumer's rating class. The results in Table 6.1 show

[26] One possible explanation for this limited information is that the firm's quality is changing over the study period. We tested this by allowing variation in q_c through time, but no significant time effects were found.

a small, significant increase in offer arrival in periods following such a search event – with a base of 0.431 offers per period increasing to 0.516 following a search event.

Comparing the impact of search events with the impact of claims, as is done in Figure 1, yields the paper's strongest evidence for learning. The horizontal access is the number of periods that have passed since a search event or claim – so 0 is the period prior to the event, 1 is the first period following the event, etc. The figure graphs four series using all observations in the dataset. The first series is the average of the predicted departure probabilities for periods following any search event *other than a claim* -- meaning the average of all observations 1 period after a search event, the average for all observations 2 periods after a search event, etc. The second series is the actual departure rate for these observations The third series is the average of the predicted departure probabilities for periods following a *non-chargeable claim*, and the fourth is the actual departure rate for these observations. All predicted and actual values are adjusted to represent a consumer in her first period with the firm, in order to eliminate the confounding impact of tenure.[27]

The patterns displayed in Figure 1 are striking on several dimensions. First, they show that the model fits quite well, as the predicted and actual lines are quite similar. Second, they demonstrate the difference between search effects and learning effects. The pattern following a search event is simple – consumers' search intensity and thus departure probability increases in period 1, but if consumers do not leave in that period, the probability returns to its original level form period 2 on. The pattern following a non-chargeable claim is quite different. First, the initial spike in departures is higher, which may reflect the impact of learning in addition to increased search. But since claims may be uniquely powerful search events – if consumers are more likely to pay attention to their policies following a claim then a price change – we shouldn't read too much into this spike. Much more telling is the pattern from period 2 on, which demonstrates a lasting impact of learning on consumer preferences, consistent with learning. In periods 2-4, those consumers who were unsatisfied, and thus have lowered their expectations of firm quality, continue to depart in greater numbers. But over time, as these consumers leave, the remaining consumers are disproportionately those who were satisfied. So, the departure rate falls steadily, and is ultimately lower than its original value, reflecting the ongoing positive effect of satisfactory claims experiences on consumer expectations.

[27] For predicted values, this simply means that tenure with the firm is set to 1 for all calculations. For actual departure rates, it means that the departure rates by tenure with the firm are converted to what they would have been in period 1, using the model's estimated tenure effects.

B. Does Consumer Satisfaction Increase or Decrease with Experience?

As noted above, most studies of experience goods assume that risk averse consumers prefer products about which they have more information. This paper suggests instead that, if consumers do not all agree about the quality of a given firm, then those consumers who select into a particular firm will surely be those who are optimistic about its quality, which means that they will ultimately be disappointed by direct experience. While in the interest of simplicity, the model used here does not explicitly account for risk aversion or heterogeneous priors, it does allow for a direct comparison of these two hypotheses. If consumer preferences increase with experience, the estimated prior will be below the true quality, and vice versa if optimism causes preferences to fall with experience.

The results provide strong evidence for optimism of those joining the firm. A 40 year old joining the firm expects good service on more than 90% of claims, vs. the actual rate of 73.9%. Table 6.3 demonstrates the rapid disappointment of such a consumer with actual experience: 1 bad draw reduces this expectation to 49.6%; 2 bad draws push it to 34.2%. For a consumer who has experienced 4 claims, the firm's true quality predicts that she has received bad service once, in which case her expected probability of good service is reduced to 78.6%.

Note that, while consumers' optimism about the study firm declines with age – as predicted by pre-entry learning -- their optimism about the broader market declines even more rapidly. So, by age 40, consumers expect the average firm in the market to provide satisfactory service just over 80% of the time, vs. 90% for the study firm; by age 80, this falls to 66.3% vs. 87.1% at the study firm. This means that, even as consumer expectations about the study firm become more accurate, those choosing to join are still optimistic about its relative quality.

C. How Fast is Learning in Practice?

Table 6.4 examines the rate of consumer learning, as a function of tenure with the firm, again for a 40 year old. The first column details the learning that would occur if this were a standard experience good with a learning draw each time the consumer buys from the firm, which in this case means one claim per period. The figures reported are the expected values of consumers' expected claims service, accounting for the randomness in service received. In this case, we expect a consumer's learning to be rapid, so that her expectation – which is initially 16.5 percentage points too high -- is within 5 percentage points of the true probability after 3 periods, and then gradually improves to within 2 points after 10 periods (5 years). In contrast, the second column details the learning which actually occurs, given the slow arrival of claims, meaning that the expected value is taken with respect to both the randomness in claim arrival and service. Here we see that learning is much slower. After 3 periods, a consumer's expectation is still nearly 15 points too high, improving only to 12 points after 5 years.

D. How Much Impact Does Learning Have on Departures?

A final important characteristic of the learning process is the impact on departure decisions. The base impact – determined by the value on good service – is mitigated by several other features of the model. First, the estimated value of $\overline{\sigma}_p$ indicates that the standard error of price in the market distribution is roughly 40% of the mean, or \$67.39 for the average consumer. This is higher than the generally accepted figure in the industry -- which is closer to 25% -- and estimated imprecisely, but it suggests that a very large change in price or expected quality is required to induce much of a change in an individual's position in the market distribution and thus his departure probability. Second, the sample average value of Ω_t is0.448, which indicates that consumers receive far less than one draw from the market each period. Such limited exposure to the broader market -- which is a generally accepted feature of this market -- further reduces the impact of price and quality changes at one's current firm.

The first section of Table 6.5 simulates the impact of learning on departures for a "base consumer" -- a 40 year old, facing a current price equal to the market average, with all other variables at their market averages. We know that this individual's prior expectation of the firm's quality is 0.904, while the true value is 0.739, so increasing the rate of learning will hasten departures. To quantify this effect, we consider 3 successively higher rates of learning – the actual random arrival of claims; one claim per period; and full information immediately upon joining the firm – relative to the case with no learning. Under these assumptions, we simulate claims arrival, service received, and thus departures, for 10,000 consumers using the estimated parameters. The percentage surviving through 1, 3, 5, and 10 years – and, in parentheses, this survival probability as a percentage of the no information case -- are reported.

Focusing on the 5 year figure, we see that shifting from no learning to the actual, random arrival of claims only reduces the survival probability by 4.2%. Higher rates of learning -- one claim per period or full learning immediately upon joining -- reduce the survival probability more, by 21.5% and 28.1%, respectively. Still, the chance of remaining with the firm for 5 years, among consumers who know that the firm's probability of good service is nearly 17 points below the prior, is more than 70% as high as the chance among those who never learn this fact.

The impact on departures is further limited by consumer "lock-in", included via $f_c(x_t, \tau_t, \tau_t^{nc}, \eta)$ which is parameterized as:

$$f_c(x_t, \tau_t, \tau_t^{nc}, \eta) = \beta_0 + \beta_1 MCAR_t + \beta_2 FP_t + \beta_3(\tau_t) + \beta_4 \lambda_t^l + \beta_5(\tau_t^{nc}) \qquad (6.1)$$

$MCAR_t$ is an indicator for consumers with multiple insured cars – 78.1% of all observations – and FP_t is an indicator for consumers who purchase first-party coverage for their own vehicles – 65.2% of all observations. They are included as proxies for switching costs, since it's clearly more difficult to search for and switch to alternate firms when one's policy is more complex.[28] $\beta_3(\tau_t)$ allows lock-in to grow with tenure. It's left as a general function of periods with the firm, subject to the obvious normalization that $\beta_3(1) = 0$, along with the restriction that $\beta_3(\tau_t) = \beta_3(13)\ \forall\ \tau > 13$ because we have limited data for tenures above 13. Finally, $\beta_4 \lambda_t^l + \beta_5(\tau_t^{nc})$ is included to control for alternate impacts of claims, as discussed above.

Note that the estimate of β_0 indicates that, as a baseline, consumers have a preference for their current firm which is more than \$25 above any alternate firm. This is an additional reason for the fairly small departure effects reported above. This preference is increased for those with more complex policies – by more than \$20 for a consumer with multiple cars and by nearly \$5 for a consumer buying first party coverage – which are fairly large relative to the average policy price of \$148.11. To examine the effect of these switching costs, we consider consumers for whom they are minimized, namely those purchasing a simple liability policy for a single car, $FP_t = 0$ and $MCAR_t = 0$. The second section of Table 6.5 reports the departure effects for these consumers. In this case, the actual learning rate reduces the 5 year survival probability by 5.6%, up from 4.2% with the sample average switching costs, and immediate learning reduces it by 36%, up from 28.9%.

It is also interesting to consider the impact of the increase in lock-in over tenure. According to the estimates of $\beta_3(\tau)$, a consumer entering his 4th year has preference for the current firm more than \$16 above a new consumer, and by year 7 the gap is nearly \$40. This creates the real possibility that consumers may be too locked-in to react by the time learning occurs. We explore this by setting $\beta_3(\tau) = 0\ \forall\ \tau$. This experiment can be viewed as eliminating any sources of lock-in with tenure, or – perhaps more compellingly – speeding up the learning and offer arrival so that everything happens within one period, before tenure effects have a chance to set in. The third section of Table 6.5 shows these results. Here, the actual learning rate reduces the 5 year survival probability by 8.6% over no-information, a substantially larger reduction than we saw in the minimal switching cost case. In contrast, while larger than the base case, the reductions for 1 claim per period and immediate learning – 26.3% and 33.1%, respectively – are not as large as in the minimum switching cost case. So, this indicates that, while both

[28] The best measure of this might be whether the consumer purchases multiple lines of insurance – say homeowners or life in addition to auto – at the study firm. However, this variable is only available for the last 2 years of data.

switching costs and tenure effects limit the impact of learning on departures, the impact of tenure effects is particularly important when claims arrive slowly and thus may be "too late" to have an effect.

Finally, the 4th section of the table considers the combined minimum switching cost/no-tenure effect case. This provides a nice summary of the results thus far. With no tenure effects and minimum switching costs, immediate learning about the firm's quality reduces the 5 year survival probability by 42.1% vs. no-information. However, in reality, the learning impact is reduced by the slow arrival of claims – with random claim arrival, the 5 year survival probability is reduced by only 8.9%. Further, Section 1 of the table demonstrates that the learning impact is reduced by lock-in -- in the presence of tenure effects and sample average switching costs, full information reduces the 5 year survival probability by only 28.1%. Finally, combining slow information arrival and lock-in, the 5 year survival probability is reduced by only 4.2%, vs. the full "potential reduction" of 42.1%.

E. How Much Benefit Do Firms Receive from Increases in Service Quality?

The results thus far raise obvious questions about firms' benefits from investing in service quality. While we can not fully address this issue without data on consumer entries as well as exits, along with a complete model of market competition, we can provide some evidence on the size of these benefits in generating repeat business. To examine this issue, we repeat the simulation method described above, again for a "base consumer." In this case, however, we consider the impact of increasing the probability of good service, q_c, from 0 to 1/3, 2/3, or 1.

As a simple way to summarize the incentive effect, assume that the regulatory process allows a fixed markup M, for all consumers in all periods, and that if a consumer departs the firm there is 0 probability that she will ever return. With these assumptions, we can compute the expected profit from a consumer – following initial purchase from the firm – as the sum of these markup terms, discounted by the probability that the consumer remains with the firm long enough to pay that markup and by an assumed discount factor, equal to 0.97 for each 6-month policy period. Denoting the 10-year expected, discounted profit as Π_{10}, we have:[29]

$$\Pi_{10} = \sum_{t=1}^{20} \beta^t \prod_{s=1}^{t} (1 - D_s) M \qquad (6.2)$$

where D_s is the departure probability defined in (4.14), accounting for all learning, tenure effects, etc. accrued through period s.

[29] It is not feasible to compute the infinite sum here because claims experience can accrue indefinitely, so this series is not stationary. However, using values larger than 10 years has no qualitative effect on the results.

Table 6.6 contains the average of 10,000 simulations of these profit terms, reported in markup units. Of most interest is the number in parentheses in each cell, which is the increase in Π_{10} achieved by increasing q_c from the next lower value to the value in that column.[30] So, for example, increasing quality from 2/3 to 1 in the base model increases this discounted profit stream by roughly 1.5 single period markups per consumer. While it's hard to evaluate this in absolute size, we can compare it to the benefit in the immediate learning case, where the same increase in quality would increase the profit stream by just over 3.2 markups per consumer, more than twice the benefit in the slow learning case. While it is not possible to determine the exact effect of this change without a model of quality competition, such a large gap suggests that concerns about under-provision of quality due to slow learning may be well-founded and deserve further attention.

What about the impact of consumer lock-in? Here the results run counter to the intuition often expressed in the industry that, because consumers are locked in, firms have little incentive to provide quality service. In fact, eliminating the tenure effects reduces the benefit to increasing quality from 2/3 to 1 by slightly more than 0.2 markups per consumer. And, for consumers with minimal switching costs, this is again reduced by slightly more than 0.2 additional markups. Why? While it's true that lock-in reduces the impact of expected service quality on the departure probability in any given period, it also increases the probability that consumers will remain with the firm into later periods, and thus the weight on these later markups. That is, firms' incentive to increase quality comes from increases in lifetime profits, and these increases are larger if that lifetime is expected to be longer.

Note that this result turns critically on the assumption that once a consumer leaves the firm, he is unlikely to return. But that assumption is highly consistent with observed patterns in this industry and service markets in general. And while we can't make precise statements about benefits from increasing quality without a more complete model, this result highlights an important feature that such a model needs to address – in a market where most sales and profits come from repeat business and where learning is slow, increasing lock-in may increase the return to investments in service quality.

VII. Model Extensions

Perhaps the model's most restrictive assumption is that q_c is fixed across consumers. It is certainly possible that the firm either treats consumers differently or that some consumers are simply

[30] We could compute a more precise measure of the benefit to an existing firm by simulating the model for, say, 10 years to generate an established base of consumers, and then simulating the impact of changes in quality. However, this substantial complication leads to no change in the conclusions drawn here.

more likely to be unsatisfied. The biggest concern this presents is that unobserved heterogeneity in q_c may be confounded with the unobserved heterogeneity in service outcomes that we rely on to identify learning.

To test the robustness of results to this possibility, we allow q_c to vary as follows:

$$q_c = q_0 + q_1 CHG + q_2 CLM\_PER_3 + q_3 ATL + q_4 YNG + \varepsilon \tag{7.1}$$

where *CHG* is an indicator for a chargeable claim, included to allow the possibility that the firm provides lesser service on such claims to "weed out" high risk consumers. CLM_PER_3 measures the number of chargeable claims per period over the last 3 years, in case the firm bases service on this cumulative measure. *ATL* is an indicator for residence in Atlanta and *YNG* is an indicator for consumers 25 or under – both of these are groups that insurance companies have been accused of neglecting. Finally, ε_q is a fixed, individual specific error term, allowing for the possibility that the firm treats consumers differently based on unobservables, or that certain individuals are simply harder to satisfy. It takes on two values, $\underline{\varepsilon}$ and $\overline{\varepsilon}$, with $pr(\varepsilon = \overline{\varepsilon}) = \rho_\varepsilon$, and is subject to the restriction that $\underline{\varepsilon} = -(\rho_\varepsilon \overline{\varepsilon} /(1-\rho_\varepsilon))$, so that $E(\varepsilon) = 0$.

With this specification, everything else proceeds as above. In particular, for simplicity, we assume that consumers still try to infer a single value for q_c.[31] The results presented in Table 7.1 can be summarized as follows. First, there is evidence of a slight reduction in the probability of satisfactory service for chargeable claims. This may reflect intentionally reduced service, or it may reflect the additional challenges associated with chargeable claims, which often involves complex liability issues. In any case, it does indicate that differences in service on different types of events is an important issue for future research. Second, however, none of the other estimates are close to statistically significant. So, there is no evidence for heterogeneous treatment by consumer type. This is consistent with the study firm's claim that service personnel know none of the policy characteristics or histories of the consumers they serve. Finally, the other learning parameters are quite similar to the base case, so the results seem robust to the possibility of heterogeneity in claims service.

VIII. Conclusions

Most research on experience goods embodies the notion that, while direct product experience is required to learn about new goods, information is more complete for established products. However, officials in many industries question this view, arguing that limited consumer information impacts

[31] While a more complete model would let them learn the parameters of the q_c function, that's much too complex for a simple robustness test.

demand even for established products. Unfortunately the empirical measurement required to settle this disagreement is quite difficult in the non-durable product markets generally studied – if consumers learn each time they purchase a product, it is nearly impossible to measure the learning process separately from all other sources of state dependence in demand.

Markets for continuously provided services – such as credit cards, telephony, cable television, etc. -- where consumers learn about service quality at distinct interactions with firms, are potentially much better venues for such measurement. Unfortunately, the occurrence of these interactions tends to be either endogenous or unobservable. This paper overcomes these problems by considering automobile insurance, where consumers learn about service quality each time they have a claim, and crucially, the occurrence of claims is completely distinct from a consumer's satisfaction with her firm and fully observable from company records. Using a panel of 18,595 consumers from one well-established auto insurance company, the paper estimates a structural model of consumers' departure decisions with an imbedded Bayesian learning model.

The paper's central result is strong evidence for the importance of consumer learning at a long-standing auto insurer. At entry, consumer's have highly inaccurate priors, which they trust only as much as observation of a single claim. This limited faith in the prior suggests strong impacts from the first claims experience, with rapidly declining impacts for later claims, a characteristic learning pattern. As consumers age, the relative impact of successive claims becomes flatter, suggesting that older consumers enter the firm with more information, presumably due to pre-entry social learning. Most importantly, the pattern of departures following any given claim is quite distinct from the pattern following other "search events" and indicates that claims have a lasting impact on consumer preferences, again suggesting that learning had occurred.

The details of the learning process provide some insights into why information remains so limited for such a well-established firm. First, the slow arrival of claims greatly limits what would otherwise be a rapid learning process. Over the course of her lifetime, a typical consumer may experience 2-3 claims. Quite simply, this means that there's not much information out there, greatly limiting opportunities for social learning. If a consumer calls 5 friends for recommendations, together they may have had 1 claims experience with any given firm, consistent with the faith in the prior estimated here. Second, the results clearly indicate that consumers are optimistic about the firm's quality when they join. This suggests that a firm has little incentive to publicize information about its true quality, particularly since its current consumers may be the most likely to pay attention.

This still leaves open the question of why more formal markets for information have not filled this void. There are some sources – Consumer Reports publishes ratings of auto insurers – but they are incomplete and clearly leave consumers with limited information. The results presented here may partially explain this absence -- after 3 years with a firm consumers are highly reluctant to switch, suggesting that the majority of consumers will have little demand for information.

While suggestive, this leaves one very important open question. If consumers know they will spend most of their lives at one firm, shouldn't they search intensely when they first choose that firm? And while the results suggest that asking friends or firms for information may be ineffective, shouldn't markets arise to aid in that search? One possibility is that, given the small number of consumers who consider switching firms in any given period, the market for information is simply too thin to function effectively. This problem is exacerbated for auto insurance since a consumer's first firm choice – presumably the most important for gathering information – may be to simply join her parent's firm. Alternatively, this may suggest that formal markets for information are less valued by consumers than word-of-mouth from friends and relatives. These issues remain an important topic for future research.

Finally, the results on firms' benefits from increasing service quality are intriguing. The slow rate of learning substantially reduces these benefits, relative to what they would be with faster information arrival, suggesting that concerns about undersupply of quality may be well founded. However, contrary to the views expressed by industry officials, consumer lock-in appears to help here, raising firms' benefits from increasing service quality – if they can keep consumers satisfied in their early days with the firm, they may be able to generate a lifetime stream of profits. While these results are incomplete without a model of quality competition, they suggest important policy considerations, including a potential role for government information provision and a concern about policies designed to reduce lock-in.

References

Ackerberg, D. (2002), "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination," mimeo.

Bolton, R. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction," *Marketing Science*, 17: 46-65.

Erdem, T. and M. Keane (1996), "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15:1-20.

Heckman, J. (1991) "Identifying the Hand of the Past: Distinguishing State Dependence from Heterogeneity," *American Economic Review*, 81: 75-79.

Horstmann, I. and G. MacDonald (1994), "When is Advertising a Signal of Product Quality," *Journal of Economics and Management Strategy*, 3: 561-584.

Horstmann, I. and G. MacDonald (2002), "Is Advertising a Signal of Product Quality? Evidence from the Compact Disc Player Market, 1983-1992," mimeo.

Hubbard, T. (2002), "How Do Consumers Motivate Experts? Reputational Incentives in an Auto Repair Market," *The Journal of Law and Economics*, 45: 437-468.

Israel, M. (2003), "Tenure Dependence in Consumer-Firm Relationships: An Empirical Analysis of Consumer Departures from Automobile Insurance Firms," mimeo.

Joskow, P. (1973), "Cartels, Competition, and Regulation in the Property-Liability Insurance Industry," *Bell Journal of Economics*, 4: 375-427.

Kihlstrom R.E. and M.H. Riordan (1984), "Advertising as a Signal," *Journal of Political Economy*, 92: 427-450.

Jovanovic, B. (1982), "Truthful Disclosure of Information," *Bell Journal of Economics,* 13: 36-44.

Milgrom, P. and J. Roberts (1986), "Price and Advertising Signals of Product Quality," *Journal of Political Economy*, 94: 796-821.

Nelson, P. (1970), "Information and Consumer Behavior," *Journal of Political Economy,* 78: 311-329.

Nelson, P. (1974), "Advertising as Information," *Journal of Political Economy,* 82: 729-754

Shapiro, C. (1982), "Consumer Information, Product Quality, and Seller Reputation," *The Bell Journal of Economics,* 13: 20-35.

Stiglitz, J. (1989), "Imperfect Information in the Product Market," *Handbook of Industrial Organization: Volume 1*, Richard Schmalensee and Robert Willig, eds. Amsterdam: North-Holland.

Tirole, J. (1988), *The Theory of Industrial Organization*, Cambridge: The M.I.T. Press, 106-126.

Wolinsky, A. (1983), "Prices as Signals of Product Quality," *Review of Economic Studies,* 50: 647-658.

Appendix I: Calculation of Market Average Prices

Market average prices are defined as the average price among the member firms of the Insurance Services Organization (ISO) for fiscal 1994-1998 by territory, rating class, and claims history. Average prices are computed separately by coverage. Each average is computed as total earned premiums divided by the total number of policies with the specified coverage in force among the member firms. The average price facing a given consumer is the sum of the prices for the coverages he currently purchases.

Average prices are assumed to be stable for the entire study period, save for growth affecting all firms measured by the motor vehicle insurance CPI. So, premiums were converted to their 1994 equivalents by deflating by this CPI, and all the premium and exposure data were combined for all years. The value for each year was then re-inflated by the appropriate motor vehicle insurance CPI figure. Finally, the insurance price was converted to real dollars by deflating by the overall CPI.

While this measure is better than any used in existing insurance literature, it still faces several limitations. First, ISO firms sell only roughly 25% of the policies in Georgia not sold by the study firm. Second, the computed price refers to all policies in force, rather than just the relevant policies for new consumers. So, for all consumer types, the computed market average price is assumed to be proportional to the average for new consumers in the overall market. The factor of proportionality, μ_0, is estimated as part of the likelihood function, and reported in the table below.

In addition, the available data fail to include several sanctioned rating factors. These are:

1. Whether the vehicle is driven over 7,500 miles per year.
2. Whether the consumer has completed a defensive driving course.
3. Whether the driver is a student who commutes over 150 miles to college.
4. Whether the vehicle is used on a farm.
5. Whether the vehicle is used primarily in the operation of a business.
6. Whether a consumer between 21-24 is 21-22 or 23-24 years old.

Together, these account for less than 1% of the variance in prices at the study firm, so the impact of excluding them should be small. But to be complete, parameters are included to adjust for these omitted classifications. To accomplish this, each excluded factor is assumed to have a multiplicative impact on price that applies to all observed consumer types at all firms. For example, consider the impact of mileage. The observed average price is a combination of high and low mileage consumers:

$$\tilde{\mu}_p = q_1 \mu_1 \underline{p} + (1 - q_1) \underline{p}$$

where $\tilde{\mu}_p$ is the observed average price, q_1 is the percentage of high mileage consumers, μ_1 is the multiplicative price increase for high mileage consumers, and $\underline{p}$ is the average price for a low mileage consumer.[32] Solving for $\underline{p}$ yields the average price for low-mileage consumers:

$$\underline{p} = \frac{\tilde{\mu}_p}{q_1 \mu_1 + (1 - q_1)}$$

The price for high-mileage consumers, then, is simply $\mu_1 \underline{p}$. We assume that the proportion of high-mileage consumers observed at the study firm applies to the market as a whole, so q_1 is observable. This leaves only μ_1 to be estimated as part of the likelihood function. The same logic is followed for $\mu_2 - \mu_6$ as well.

The estimated market price adjustment factors are:

Parameter	Classification	Estimate	Standard Error
μ_0	Proportionality Factor	0.891	0.112
μ_1	High Mileage	1.281	0.231
μ_2	Defensive Driving	0.967	0.441
μ_3	Commuting Student	1.155	0.412
μ_4	Farm	0.821	0.412
μ_5	Business	1.254	0.406
μ_6	23-24	0.929	0.315

[32] In this discussion, we are assuming that the observed average price is already adjusted for the base proportionality factor, μ_0. Also note the implicit assumption that mileage status is uncorrelated with all other rating categories, both observed and unobserved. In this way, the only difference between the average price for low and high mileage consumers results from the multiplicative pricing factor, μ_1.

Appendix II: Claims Arrival Functions

Functional forms for the probability of claims are based on standard industry rating variables. In particular, we specify:

$$\lambda^l(x_t;\eta) = \Phi.(\lambda_0^l(BASE_t) + \lambda_1^l COLL_t + \lambda_2^l ALT_t + \lambda_3^l U21_t + \lambda_4^l WK_t + \lambda_5^l HIGH_t + \lambda_6^l NEW_t + \lambda_7^l OLD_t + \\ .\lambda_8^l COMPACT_t + \lambda_9^l SUV_t + \lambda_{10}^l TRUCK + \lambda_{11}^l ENTRYCLM_t + \lambda_{12}^l VEHDENS + \lambda_{13}^l VEHDENS^2 + \\ .\lambda_{14}^l ATL + \eta)$$

where $BASE_t$ is the consumer's base rating class – defined by age, gender, and marital status; $COLL_t$ is an indicator for whether the consumer purchases collision coverage; ALT_t is an indicator for placement in an "alternate company" that the firm uses for higher risk drivers; $U21_t$ is an indicator for consumers under 21; WK_t is an indicator for vehicles which are used to drive to work; $HIGH_t$ is an indicator for vehicles that are driven more than 7500 miles per year; NEW_t is an indicator for vehicles that are less than 1 year old; OLD_t is an indicator for vehicles that are 8 or more years old; $COMPACT_t$ is an indicator for compact or sub-compact cars; SUV_t is an indicator for SUV's, light trucks, or recreational vehicles; $TRUCK_t$ is an indicator for heavy trucks; $ENTRYCLM_t$ is an indicator for consumers with at least 1 chargeable claim in the 3 years before entering the firm; $VEHDENS_t$ measures vehicles per square mile in the consumer's zip code; and ATL_t is an indicator for residence in Fulton County.

The functional form for $\lambda^n(x_t)$ is symmetric, except that $COMP_t$ – an indicator for comprehensive coverage for storm damage, theft, etc. – replaces $COLL_t$, and η is excluded. Estimates for all parameters are available from the author on request.

Table 3.1: Entry and Exit from Sample

	Count	Percent
Total Policies	18,595	100.0
Year of Entry		
1991	509	2.7
1992	2,336	12.6
1993	2,475	13.3
1994	2,648	14.2
1995	2,709	14.6
1996	2,820	15.2
1997	2,642	14.2
1998	2,456	13.2
Cause of Exit		
Voluntary Departure	5,854	31.5
Right Censored at Dec 1998	7,976	42.9
Randomly Censored before Dec 1998	4,765	25.6

Table 3.2: Claims Experience

Variable	Count	Frequency	Percent
Total Consumer Periods	N/A	85,931	100.0
Claims in Period	0	78,942	91.9
	1	6,319	7.4
	2^+	670	0.8
Chargeable Claims in Period	0	82,632	96.2
	1	3,166	3.7
	2^+	133	0.1
Non-Chargeable Claims in Period	0	82,004	95.4
	1	3,600	4.2
	2^+	327	0.4
Total Claims	0	63,456	73.8
	1	15,283	17.8
	2	4,693	5.5
	3	1,546	1.8
	4	569	0.7
	5	229	0.3
	6^+	155	0.2

Table 3.3: Descriptive Statistics for Price

Statistic	Study Firm Price	Market Average Price
Mean	148.11	162.77
Standard Deviation	73.37	90.50
25^{th} Percentile	105.94	108.79
50^{th} Percentile	126.85	137.04
75^{th} Percentile	158.84	180.39

Table 6.1: Parameter Estimates

Category	Parameter	Description	Estimate	Standard Error
Offers	Ω_0	Base Offers	0.431	0.187
	Ω_1	Search Event	0.085	0.038
Price Stnd. Error	$\overline{\sigma}_p$	Standard Error Factor	0.414	0.212
Price Sensitivity	$\alpha_1(1)$	Limit Class 0	1.029	0.015
	$\alpha_1(2)$	Limit Class 2	0.975	0.013
	$\alpha_1(3)$	Limit Class 3	0.972	0.031
Preference for Current firm, $f_c(\bullet)$	β_0	Constant	25.184	7.131
	β_1	Multiple Cars	21.312	5.333
	β_2	First-Party Coverage	4.964	3.084
	$\beta_3(2)$	Tenure 2	-5.828	0.912
	$\beta_3(3)$	Tenure 3	0.470	0.945
	$\beta_3(4)$	Tenure 4	2.426	0.997
	$\beta_3(5)$	Tenure 5	8.268	1.031
	$\beta_3(6)$	Tenure 6	7.842	1.054
	$\beta_3(7)$	Tenure 7	16.024	1.077
	$\beta_3(8)$	Tenure 8	19.246	1.061
	$\beta_3(9)$	Tenure 9	18.908	1.231
	$\beta_3(10)$	Tenure 10	29.428	1.553
	$\beta_3(11)$	Tenure 11	33.464	2.012
	$\beta_3(12)$	Tenure 12	32.976	2.231
	$\beta_3(13)$	Tenure 13	39.664	2.665
Value on Claims Service	α_2^0	Base Value	823.41	82.213
	α_2^1	Extra Insurance	232.23	93.104
Learning	q_c	Study Firm's Quality	0.739	0.101
	μ_0	Initial Prior	0.991	0.263
	ψ_0	Initial Weight	0.797	0.314
	ψ_1	Age	0.021	0.0091
	ψ_2	Age^2	-0.00015	0.00012
	μ_1	Age	-0.0091	0.0045
	μ_2	Age^2	0.000062	0.000088

Estimation is based on the following functional forms:

- $f_c(x_t, \tau_t, \tau_t^{nc}, \eta) = \beta_0 + \beta_1 * MCAR_t + \beta_2 * FP_t + \beta_3(\tau_t) + \beta_4 * \lambda_t^l + \beta_5(\tau_t^{nc})$

  $MCAR_t$ =0/1 indicator for consumers with multiple insured cars
  FP_t = 0/1 indicator for purchase of first coverages
  β_0, β_1, β_2, and $\beta_3(\tau_t)$ are used in the discussion of results and thus are reported here.

- $\alpha_2(x_t) = \alpha_2^0 + \alpha_2^1 * EX_t$

  EX_t = 0/1 indicator for purchase of extra insurance, beyond the legal mandate.

Table 6.2: Pre-Entry Knowledge

Age	ψ_p	μ_p	μ_Ω
16	0.797	0.991	0.991
20	0.879	0.968	0.956
25	0.974	0.945	0.914
30	1.062	0.928	0.876
35	1.142	0.915	0.840
40	1.214	0.904	0.808
45	1.280	0.896	0.779
50	1.338	0.889	0.753
55	1.388	0.884	0.730
60	1.431	0.879	0.711
65	1.466	0.876	0.694
70	1.494	0.873	0.680
75	1.514	0.872	0.670
80	1.527	0.871	0.663

Table 6.3: Learning With Claims

	G_t					
C_t	0	1	2	3	4	5
0	0.904	**X**	**X**	**X**	**X**	**X**
1	0.496	0.948	**X**	**X**	**X**	**X**
2	0.342	0.653	0.964	**X**	**X**	**X**
3	0.261	0.498	0.735	0.972	**X**	**X**
4	0.211	0.402	0.594	0.786	0.978	**X**
5	0.177	0.338	0.499	0.659	0.820	0.981

All calculations are for a 40 year old consumer.

Table 6.4: Learning Over Tenure

Periods	1 Claim per Period	Random Claim Arrival
0	0.904	0.904
1	0.830	0.898
2	0.801	0.892
3	0.787	0.887
4	0.778	0.881
5	0.771	0.876
6	0.767	0.872
7	0.763	0.867
8	0.761	0.863
9	0.759	0.859
10	0.757	0.855

All calculations are for a 40 year old consumer.

Table 6.5: Impact of Learning On Departures

	Year	No Information	Random Claim Arrival	One Claim per Period	Immediate Learning
Base Consumer	1	0.809	0.805 (99.5%)	0.765 (94.6%)	0.744 (92.0%)
	3	0.572	0.554 (96.9%)	0.488 (85.3%)	0.453 (79.2%)
	5	0.452	0.433 (95.8%)	0.355 (78.5%)	0.325 (71.9%)
	10	0.322	0.289 (89.8%)	0.220 (68.3%)	0.189 (58.7%)
Minimum Switching Costs ($MCAR_t$=0 FP_t=0)	1	0.702	0.700 (99.7%)	0.662 (94.3%)	0.625 (89.0%)
	3	0.395	0.379 (95.9%)	0.320 (81.0%)	0.289 (73.2%)
	5	0.267	0.252 (94.4%)	0.191 (71.5%)	0.171 (64.0%)
	10	0.147	0.129 (87.8%)	0.086 (58.5%)	0.069 (46.9%)
No Tenure Effects ($\beta_3(\tau)=0 \ \forall \tau$)	1	0.825	0.811 (98.3%)	0.773 (93.7%)	0.760 (92.1%)
	3	0.553	0.522 (94.4%)	0.460 (83.2%)	0.437 (79.0%)
	5	0.372	0.340 (91.4%)	0.274 (73.7%)	0.249 (66.9%)
	10	0.140	0.117 (83.6%)	0.088 (62.9%)	0.073 (52.1%)
Minimum Switching Costs/No Tenure Effects	1	0.729	0.722 (99.0%)	0674 (92.5%)	0.653 (89.6%)
	3	0.385	0.370 (96.1%)	0.307 (79.7%)	0.276 (71.7%)
	5	0.202	0.184 (91.1%)	0.133 (65.8%)	0.117 (57.9%)
	10	0.041	0.035 (85.4%)	0.022 (53.7%)	0.016 (39.0%)

All calculations are for a 40 year old consumer. Price at the study firm is assumed equal to the market average. All other variables are at their sample averages, unless otherwise noted.

Table 6.6: Benefits from Increased Claims Service

		$q_c = 0$	$q_c = 1/3$	$q_c = 2/3$	$q_c = 1$
With Tenure Effects & Avg. Switching Costs	Learning with Random Claims	4.502	5.685 (1.183)	7.174 (1.489)	8.681 (1.507)
	Immediate Learning	1.573	2.851 (1.278)	5.419 (2.568)	8.683 (3.263)
Learning with Random Claims	With Tenure Effects & Switching Costs	4.502	5.685 (1.183)	7.174 (1.489)	8.681 (1.507)
	No Tenure Effects	3.900	4.895 (0.995)	6.076 (1.181)	7.351 (1.275)
	No Tenure Effects/ Min Switching Costs	2.666	3.443 (0.777)	4.326 (0.883)	5.379 (1.053)

All calculations are for a 40 year old consumer. Price at the study firm is assumed equal to the market average. All other variables are at their sample averages, unless otherwise noted.

Table 7.1: Heterogeneity in Probability of Good Service

Category	Parameter	Description	Estimate	Standard Error
q_c	q_0	Constant	0.771	0.187
	q_1	Chargeable Claim	-0.073	0.033
	q_2	Claims per Period	0.012	0.010
	q_3	Atlanta	-0.009	0.034
	q_4	Young	-0.007	0.011
	$\overline{\varepsilon}$	High Type	0.011	0.041
	$\underline{\varepsilon}$	Low Type	-0.008	N/A
	ρ_ε	Probability of High Type	0.414	0.365
Learning	μ_0	Initial Prior	0.982	0.291
	ψ_0	Initial Weight	0.765	0.331
	ψ_1	Age	0.022	0.0098
	ψ_2	Age^2	-0.00016	0.00013
	μ_1	Age	-0.0093	0.0041
	μ_2	Age^2	0.000058	0.000101

Figure 1: Predicted vs. Actual Departure Rates Adjusted for Tenure Effects