Cellular Service Demand:
Biased Beliefs, Learning, and Bill Shock*

Michael D. Grubb[†] and Matthew Osborne[‡]

October 27, 2011

Abstract

We estimate a model of consumer plan choice, usage, and learning in cellular-phone services on a detailed panel data set of individual bills. Our model allows consumers to learn about how much they value cellular services. We infer consumers' predictions of their future cellular usage from plan choices, and compare these predictions to actual usage. We find that on average consumers underestimate their average tastes for calling (mean bias), underestimate their own uncertainty about their average tastes (overconfidence), and underestimate the monthly variation (projection bias) in their tastes for usage. Counterfactual experiments show these biases cost consumers $49 per year. Our paper also advances structural modeling of demand in situations where multipart tariffs induce marginal price uncertainty at the time consumers make consumption choices. Our approach is based on novel evidence that consumers are inattentive to past usage in such settings. Holding prices fixed, we find that the FCC's proposed bill-shock regulation requiring users be notified when exceeding usage allowances would cut revenues 8% and increase average consumer welfare by about $21 per year. We find that bill-shock regulation is particularly effective because consumers are biased. Absent consumer bias, the regulation would only increase average consumer welfare by less than $2 per year. These findings change when we allow firms to optimally respond to the bill shock regulation: in that case, firms adjust prices so that the regulation has almost no effect on profits, leading to an overall reduction in consumer welfare of about $25 per consumer.

---

*A previous version circulated under the title "Cellular Service Demand: Tariff Choice, Usage Uncertainty, Biased Beliefs, and Learning". We thank Parker Sheppard and Mengjie Ding for research assistance. The views expressed herein are those of the authors and not necessarily those of the Bureau of Economic Analysis or the U.S. Department of Commerce. Work on this paper began when Osborne worked at the Department of Justice.

[†]Massachusetts Institute of Technology, Sloan School of Management. mgrubb@mit.edu.

[‡]Bureau of Economic Analysis. Matthew.Osborne@bea.gov.

1 Introduction

This paper addresses three important questions relating to consumer demand for new products and demand for services sold via multipart tariffs. The first question relates to the extent to which consumer beliefs are unbiased: in particular do consumers make predictable mistakes when using new products, and if so, what sort of biases do they display? Moreover, how quickly are initial mistakes corrected via learning and switching? The second question concerns the marginal price uncertainty that arises in markets for services such as cellular phones, electricity, and health care, where multipart tariffs cause marginal price to vary with usage. When making any particular consumption choice, consumers are typically uncertain about their future consumption choices, and hence are uncertain about the marginal price of usage. Our second question has two parts. The first is economic: how do consumers make usage decisions given marginal price uncertainty? The second is methodological: how can we incorporate consumer uncertainty about the marginal price into our demand models in a tractable way? The third question ties the first two questions together: how do biased beliefs and marginal price uncertainty affect contracts and regulatory intervention? For instance, how will the *bill-shock* agreement recently reached between the FCC and cellular carriers affect offered contracts and welfare when it is implemented? By April 2013, this agreement commits cellular service providers to inform consumers when they exceed their allowance of included minutes so that consumers always know when the marginal price increases to the overage rate (CTIA - The Wireless Association 2011a). As shown by Grubb (2011), the effect of this disclosure on welfare is theoretically ambiguous and can depend importantly on consumer biases.

To answer these questions, we develop and estimate a dynamic model of plan choice and usage that makes use of detailed cellular phone data described in Section 3.1. The data was obtained from a major US university that acted as a reseller for a national cellular phone carrier, and covers all student accounts managed by the university from 2002 to 2004. At the time this data was collected, cell phones were a relatively new product, having 49% penetration in 2002 in the United States, compared to 98% in 2010.[1] This feature of our data makes it ideal for investigating consumer beliefs about new products.

Our modeling approach is shaped by six stylized facts in the data documented in Section 3.2. First, (1) consumers' usage choices are price sensitive. Second, (2) subscribers to the three-part tariffs had an overage 16 percent of the time and made usage choices while uncertain about the ex post marginal price. These two features make our data set a good candidate for examining usage

[1] Penetration rates are calculated as estimated total connections (CTIA - The Wireless Association 2011b) divided by total population (U.S. Census Bureau 2011).

decisions under marginal price uncertainty. Third, (3) consumers are inattentive to the balance of remaining 'free' minutes during the billing cycle. Fourth, (4) consumers are uncertain about future usage choices when choosing calling plans. Fifth, (5) consumers learn about their own usage patterns over time and switch plans in response. Sixth, (6) consumers make predictable mistakes indicative of biased prior beliefs.

Consumers in our data are heterogeneous in their average taste for cellular-phone usage. Consumers do not know their own average tastes when they initially choose a calling plan. Rather, they are initially uncertain about their own average taste for usage. Consumers then learn about their own tastes over time, and switch to more appropriate plans if an initial plan choice was not a good match. A consumer's initial plan choice is determined not by his true average taste for usage but by his beliefs about his average taste for usage. The fact that consumers make different initial plan choices reflects the fact that initial consumer beliefs are heterogeneous. We call a consumers' average taste for cellular-phone usage his true type. We assume that each consumer's prior consists of a point estimate of her own true type and a level of perceived uncertainty about this point estimate.

Our data is informative both about consumers' actual average tastes for cellular phone usage and about their prior beliefs about their own tastes. Consumers' usage choices identify the distribution of consumers' true types, while consumers' initial plan choices and subsequent switching decisions identify beliefs. The joint distribution of beliefs and true types determines whether beliefs are biased in the population. For instance, suppose that we consider the subset of consumers that all share a particular prior belief about their own types. A common assumption (often labeled rational expectations) is that this belief coincides with the distribution of true types within this subset of the population. We relax this assumption, separately identify both beliefs and the distribution of true types conditional on beliefs, and then compare the two distributions. We label differences between these distributions as biases.[2]

We assume that consumers are Bayesian learners, as is standard in the literature which estimates learning models from consumer level data (Erdem and Keane 1996, Ackerberg 2003, Crawford and Shum 2005).[3] Thus even in the presence of biased prior beliefs, consumers will eventually learn

[2]An alternate interpretation is that unmeasurable prior beliefs were unbiased at some previous time, but are now measurably and systematically different from reality at the population level (although consistent with rational expectations) due to the arrival of a correlated shock or signal at the population level. The distinction is pedantic as it does not matter for optimal firm pricing, consumer welfare, policy counter-factuals or other issues of interest.

We do not allow for other types tariff choice biases such as the flat-rate bias documented by Lambrecht and Skiera (2006) in internet service choice. For our purposes this bias is not likely to be central, though, since none of the phone plans we analyze are flat rate.

[3]One exception we are aware of is Camacho, Donkers and Stremersch (2010), who develop a modified Bayesian

their true types by observing their own usage each month and updating their beliefs. The rate at which learning occurs depends on the monthly volatility in tastes for usage: the higher is volatility the slower consumer learning will be. A commonly made assumption is that consumers know the volatility of their own tastes, which is another aspect of rational expectations. We relax this assumption, allowing for another way in which consumers can be biased: by over- or under-estimating this variance.

Our first contribution is to identify two substantial biases causing predictable mistakes. The first we label *overconfidence*, which arises when a consumer underestimates her own uncertainty surrounding her point estimate of her true type.[4] We find that consumers underestimate their own uncertainty about their true type by 82%. Overconfident consumers initially choose plans that are too risky. Moreover, they place too much weight on their prior point estimates when updating beliefs and will be slow to learn and switch plans based on experience.

The second type of bias that we focus on we call *projection bias*, which arises when consumers underestimate the monthly volatility in their tastes for usage. Our estimates imply that consumers underestimate the volatility in their taste for usage by 47%. Consumers who exhibit projection bias underestimate the extent to which their tastes will change over time, a prevalent behavior that has been documented in a variety of experiments, surveys, and field studies (Loewenstein, O'Donoghue and Rabin 2003, Conlin, O'Donoghue and Vogelsang 2007). Similar to overconfidence, projection bias causes consumers to underestimate the uncertainty in their usage predictions when making plan choices, and choose plans that are too risky. However projection bias has the opposite effect of overconfidence on the rate of learning: projection bias causes consumers to underweight their priors relative to past usage when updating their beliefs about their average tastes for usage. This leads to faster learning and more frequent plan switching. Note that both overconfidence and projection bias cause consumers to choose plans that are too risky, but the rate of plan switching allows us to separate the two biases.[5] Because we find that overconfidence is stronger than projection bias,

---

learning model of physician learning about prescription drugs where physicians place more weight on information from patients who switch prescriptions as opposed to those who do not.

[4]Overconfidence could more broadly be interpreted to include projection bias, however we seek to draw a distinction between two different biases and define overconfidence more narrowly to do so. A significant body of experimental evidence shows that individuals are overconfident about the precision of their own predictions when making difficult forecasts (e.g. Lichtenstein, Fischhoff and Phillips (1982)). In other words, individuals tend to set overly narrow confidence intervals relative to their own confidence levels. A typical psychology study might pose the following question to a group of subjects: "What is the shortest distance between England and Australia?" Subjects would then be asked to give a set of confidence intervals centered on the median. A typical finding is that the true answer lies outside a subject's 98% confidence interval about 30% to 40% of the time.

[5]Our model includes a price consideration parameter that plays a similar role to a switching cost. This is separately identified from the learning rate by the rate at which consumers fail to switch away from strictly dominated plans.

consumers overweight their prior beliefs relative to new information and learn and switch plans relatively slowly. Thus initial plan choice mistakes are especially costly. Together, we find that overconfidence and projection bias reduce consumer welfare by about \$29 per student over a one year period.

There are other types of biases which could result in consumer behavior that is similar to that caused by overconfidence and projection bias. To ensure we do not misattribute other errors as overconfidence or projection bias, we estimate a flexible distribution of initial beliefs which captures (at least) two other potential sources of bias.[6] The first is aggregate mean bias, which allows the average consumer to under or overestimate his true type, choosing plans that are predictably too small or too large. The second is conditional mean bias, which allows consumers to overreact or underreact to private information. If consumers overreact to private information, they will predictably choose plans that are too extreme. Consumers who choose the largest calling plans would predictably benefit by moderating choices and choosing a smaller plan, while consumers who choose the smallest calling plans would predictably benefit by moderating choices and choosing a larger plan. We estimate significant negative aggregate mean bias and positive conditional mean bias. These biases have a smaller impact on consumer welfare than overconfidence and projection bias, as they reduce consumer welfare by an additional \$20 per student.

The second contribution of our paper is to provide new evidence on how consumers make consumption choices under marginal-price uncertainty and estimate a tractable model incorporating realistic behavior with marginal price uncertainty. The issue arises in cellular phone service, electricity, health care, and whenever a consumer must make a series of small purchase choices that are aggregated and billed under a multipart tariff. The current state-of-the-art approach to modeling marginal-price uncertainty is typically to assume it away. (Notable exceptions are Yao, Mela, Chiang and Chen (2011) and Jiang (2011).) Models typically either assume that consumers can perfectly predict their future usage (Cardon and Hendel 2001, Reiss and White 2005), or that consumers believe they can perfectly predict their usage up to an implementation error which they ignore (Iyengar, Ansari and Gupta 2007). The first assumption predicts that the distribution of usage will include bunching at contract kink points where marginal prices increase. We reject this in our data (Section 3.2), as do Saez (2002) and Borenstein (2009) in the contexts of labor supply and electricity consumption respectively. The second assumption conflicts with our finding that overconfidence and projection bias, while severe, are not complete: consumers are aware that they are uncertain about their future usage.

[6]We are able to separately identify these biases due to the rich choice set of plans in our data that importantly include both three-part tariffs and a two-part tariff.

If consumers do face marginal-price uncertainty and are (at least partially) aware of it, then how do they make consumption choices? Attentive and unboundedly rational consumers would solve a complicated dynamic programming problem. At each calling opportunity, consumers would place or answer a call only if its value exceeded some threshold v^*, where v^* would be conditioned on the balance of include minutes and days remaining within the billing cycle. Our paper provides novel evidence in Section 3.2 testing this hypothesis, which is made possible because we observe each phone call made rather than only monthly totals. The primary testable prediction of the unboundedly rational model is that consumers should cut back calling following a period of high-usage (and vice versa) at the end of a billing cycle, but not at the start of a billing cycle. We find no evidence of such behavior and conclude that consumers are inattentive to their remaining balance of included minutes within a billing cycle.[7]

Building on these findings, we model consumers who are aware of their own uncertainty about ex post marginal price when making usage decisions. We assume calling opportunities arise exogenously and consumers choose a calling threshold, accepting calls that are more valuable than the threshold but rejecting those that are less valuable. Consumers choose their threshold to maximize their expected utility conditional on their beliefs. This is optimal behavior for an inattentive consumer who does not keep track of past usage within the billing cycle, and hence cannot condition calling choices on this information (Grubb 2011).

The type of threshold model we implement has been proposed in earlier work, but has not been implemented in a structural model. In the context of electricity demand, Borenstein (2009) independently proposes that consumers choose *behavioral rules*, such as setting the thermostat, that determine consumption. The calling threshold chosen by consumers in our model is similar to Borenstein's (2009) behavioral rule. Borenstein (2009) uses the behavioral rule assumption to motivate using expected marginal price rather than realized marginal price in reduced form estimates of electricity price elasticities. Saez (2002) also suggests a very similar model for labor choice by income tax filers. Note that in the approaches taken in both of these papers, consumer beliefs about the distribution of the idiosyncratic error must be modeled. An advantage of our approach, which embeds the usage rule into a structural model, is that we can estimate consumer beliefs.

Our third contribution is a counter factual evaluation of the *bill-shock* agreement recently reached between the FCC and cellular carriers that will require firms to begin informing consumers

[7] Due to the way the university contracted with the carrier, students could not easily check how many minutes they had used during the course of a billing period. This means that it was very difficult for students to keep track of minutes used, making consumer inattention an especially plausible assumption.

when their included minutes are exhausted before April 2013 (CTIA - The Wireless Association 2011a).[8] Our usage model allows us to examine the welfare implications of this and other interesting regulatory interventions, which would not have been possible with earlier usage models. In particular, our usage model allows us to forecast consumer response to the introduction of such regulation, and to compute the welfare implications of the new rule. Absent price changes, we find that this regulation increases consumer welfare: it would reduce operator revenue by about 8 percent and increase consumer welfare by about \$22 per customer per year, if customers are faced with the same set of plans as the university students. Under the assumption that consumers only face the set of publicly available plans, where no fixed rate plan was available, the increase in consumer welfare rises to \$43. The presence of overconfidence and projection bias has a strong influence on the effectiveness of the bill-shock regulation. When these biases are removed, the regulation only increases consumer welfare by about \$4 per customer per year. When all biases are removed, this effect drops to less than \$2 per customer per year.

The fact that bill-shock regulation raises consumer surplus is a foregone conclusion when prices are held constant. However, we would expect firms to adjust their prices in response to the introduction of the rule, which may actually end up making consumers worse off. We conduct a counterfactual simulation where we allow firms to adjust prices in response to the regulation. To do so, we add additional supply side structure to our model and add a parameter λ measuring the amount of differentiation across firms. This firm differentiation parameter λ is omitted from our estimated demand model because our demand data is from a single carrier and does not identify λ. To complete our endogenous price counterfactual simulations, we therefore first calibrate the firm differentiation parameter λ conditional on our demand estimates using observed prices. We find that firms respond to bill-shock regulation by raising fixed fees and increasing included minute allowances on three-part tariffs. By doing so, firms maintain profits close to unregulated levels (we find a slight increase in firm profits of about \$3.80 per person annually). This means that consumers are essentially residual claimants on total welfare and hence consumer welfare drops by about \$32 per person. Absent consumer biases, bill-shock regulation has no effect. Consistent with Grubb (2009), we find that firms only offer two-part tariffs when consumers have no biases, which leaves no scope for bill-shock regulation.

Section 2 discusses related literature. Section 3 describes our data and outlines six stylized facts in our data that shape our modeling approach. Section 4 describes our model and explains identifi-

[8] At the announcement of the agreement, President Barack Obama explained: "Far too many Americans know what it's like to open up their cell-phone bill and be shocked by hundreds or even thousands of dollars in unexpected fees and charges. But we can put an end to that with a simple step: an alert warning consumers that they're about to hit their limit before fees and charges add up."

cation. Sections 5, 6, and 7 discuss estimation, present results and conclude. Additional details are in the Online Appendix available at \url{www.mit.edu/˜mgrubb/GrubbOsborneAppendix.pdf}.

2 Related Literature

We model consumers who are aware of their own uncertainty about ex post marginal price when making usage decisions and do not condition calling choices on their remaining balance of included minutes because they are inattentive to this information. Our consumers choose a constant calling threshold and only make calls that are more valuable. This represents an advance over the existing literature which typically assumes away marginal-price uncertainty (Cardon and Hendel 2001, Reiss and White 2005, Iyengar et al. 2007).[9] Yao et al. (2011) and Jiang (2011) are two recent exceptions that explicitly model marginal-price uncertainty.

Yao et al. (2011) assume that consumers are attentive and condition calling choices on their remaining balances of included minutes. We reject this model because it is inconsistent with consumer behavior in our data (Section 3.2). Our results are supported by Leider and Şahin's (2011) experimental work, which suggests that consumers who receive feedback about past usage do not follow an optimal dynamic program but instead use a constant calling threshold until all included minutes are used up and then adjust to the overage rate. This finding is consistent with our model of consumer behavior under our bill-shock counterfactual in which consumers are alerted when exceeding their allowance. In contrast, Yao et al. (2011) reject our static calling threshold model in favor of attentive dynamic behavior using Chinese cellular phone data.[10] The discrepancy between Leider and Şahin's (2011) finding and our own may be due in part to the fact that, unlike consumers in our data, the Chinese consumers could check their minute balance. Moreover, results in all three papers can be reconciled by the fact that the financial incentives to pay attention were likely stronger for Chinese consumers than for American consumers and lab subjects.

Complementary work by Jiang (2011) assumes that each consumer chooses a target quantity implemented with error and anticipates this error when choosing plans and target quantities. Jiang (2011) also evaluates the new bill-shock agreement between the FCC and cellular carriers via counterfactual simulation, predicting a $370 million welfare improvement. In contrast to our own

[9]Narayanan, Chintagunta and Miravete (2007) model consumer usage decisions in telephone plan choice where consumers anticipate ex-ante uncertainty; however, in their application consumers always face constant marginal prices, meaning that there is no marginal-price uncertainty.

[10]Yao et al. (2011) show that a scatter plot of cumulative weekly usage within a billing cycle against its lag is concave. In contrast we find the relationship is linear in our data which is consistent with our constant calling threshold. Because such linearity does not rule out dynamic behavior we conduct the additional analysis reported below.

approach, Jiang (2011) imposes rational expectations rather than estimating consumer beliefs and has cross sectional data so cannot address learning. (A strength of Jiang's (2011) data is that it is nationally representative and covers all carriers.) Finally, Jiang's (2011) implementation error does not enter consumption utility and hence the model is isomorphic to one without usage uncertainty in which marginal price uncertainty arises from exogenous billing errors that make usage allowances stochastic. Jiang's (2011) bill-shock counterfactual corresponds to removing these exogenous billing errors from the model. In contrast, a strength of our approach is that usage uncertainty arises endogenously from shocks to tastes and we explicitly model the information disclosure required by bill-shock regulation.

Liebman and Zeckhauser (2004) suggest that individuals may respond to average prices rather than marginal prices, a behavior they dub *ironing*, and find supporting evidence that this is true of labor choices following introduction of the child tax credit in 1998. Similarly, Ito (2010) shows that electricity consumers respond to average price rather than marginal price. It is not surprising that ironing arises in these settings because electricity tariffs and the income tax code are both very complex and often not well understood by consumers. A typical electricity consumer may not even realize electricity pricing is nonlinear, in which case average price would be a good estimate of marginal price. However, in the context of cellular phones this model is not appealing because consumers are fully aware that contracts include an allowance of 'free' minutes.

Empirical models with consumer beliefs typically impose rational expectations. Examples include Erdem and Keane (1996), Ackerberg (2003), and Osborne (2011) in consumer packaged goods, and Miravete (2002), Gaynor, Shi, Telang and Vogt (2005), Narayanan et al. (2007), Iyengar et al. (2007), and Jiang (2011) in telephone service, and Chintagunta, Manchanda and Sriram (2009) in video-on-demand service. Growing evidence shows that consumers are often biased.[11] A small number of papers including Crawford and Shum (2005) and Goettler and Clay (2010) relax the rational expectations assumption and estimate mean biases. Due to the richness of the tariff choice-set in our data, we are able to rely less on the rational expectations assumption and identify more about prior beliefs from choice data than such earlier work. For instance, the paper most similar to ours is Goettler and Clay (2010) but Goettler and Clay's (2010) relaxation of rational expectations is limited to mean biases, while we also measure (rather than assume away) projection bias and overconfidence. Goettler and Clay (2010) cannot identify higher moments of beliefs because the choice set in online grocery-delivery service is limited to two-part tariffs.

[11] For instance, using the same data as this paper, Grubb's (2009) static analysis suggests consumers underestimate uncertainty about future usage but cannot measure the bias or distinguish overconfidence from projection bias. Bar-Gill and Stone (2009) make similar findings with alternate cellular billing data.

Our work is related to a sequence of papers about Kentucky's 1986 local telephone tariff experiment (Miravete 2002, Miravete 2003, Miravete 2005, Narayanan et al. 2007, Miravete and Palacios-Huerta 2011). A theme of this work is that the standard model of consumer choice does well at explaining behavior. Most consumers choose the right plan initially (over 80%) and, of those who do not, many (over 15%) switch plans within three months to lower their bills, even though typically savings from switching were less than four dollars per month (Miravete 2003, Miravete and Palacios-Huerta 2011). While our results emphasize the presence and value of modeling of systematic consumer biases, our results are nevertheless consistent with the Kentucky experiment. First, although the standard model of consumer choice does well at explaining behavior in the Kentucky experiment, our estimates of negative aggregate mean bias and positive conditional mean bias are consistent with evidence in Miravete (2003) which documents that on average all consumers who chose a small metered plan would have saved money on a larger flat rate plan.[12] (Consumers were not offered three-part tariffs in the Kentucky experiment so their choices do not shed light on overconfidence or projection bias.) Second, although we document several systematic biases, consumers also do a lot right in our setting just as they did in the Kentucky experiment. In particular, we also find that most consumers initially choose the tariff that turns out to be optimal ex post (55 to 71 percent, Tables 3-4).[13] Moreover, 15 percent of consumers switch plans at least once and 60 to 75 percent of switches appear to be in the right direction to lower bills (Section 3.2).

In Section 6.3 we conduct counterfactual simulations with endogenous prices. This excercise is related to the literatures with standard consumers on monopoly sequential-screening (surveyed by Rochet and Stole ((2003), Section 8), including Baron and Besanko (1984), Riordan and Sappington (1987), Miravete (1996), Courty and Li (2000), Miravete (2005), and Grubb (2009)) and competitive static-screening (surveyed by Stole (2007), including Armstrong and Vickers (2001) and Rochet and Stole (2002)). Moreover, it is related to the growing literature on optimal contracting with non-standard consumers (for which Spiegler (2011) provides a good guide). Of particular relevance are DellaVigna and Malmendier (2004), Uthemann (2005), Eliaz and Spiegler (2006), Eliaz and Spiegler (2008), Grubb (2009), Herweg and Mierendorff (Forthcoming), and Grubb (2011).

Finally our paper is about the cellular phone industry, about which there is a small literature.

---

[12]Interestingly, in Miravete (2003) the bias that can be inferred from elicited expectations differs from that inferred from choices.

[13]It is true that consumers in our data make more mistakes those in the Kentucky experiment, however this is likely simply due to the fact that our consumers faced a larger and more complex choice set. Within our own sample, for instance, we see the incidence of ex post mistakes increases from 29 percent to 45 percent between fall 2002 and fall 2003 when price changes made the choice problem more difficult (Tables 3-4). Moreover, few consumers in our data persist in obviously bad choices: Of those who choose a three-part tariff, only 6 percent fail to immediately reduce usage, quit, or switch plans after making 3 consecutive overage payments in excess of $10.

Beyond work already mentioned, other work on the cellular phone industry examines carrier switching costs (Kim 2006), the effect of entry on pricing (Seim and Viard 2010, Miravete and Röller 2004), the effect of number portability regulation on competition (Park 2009), the role of multi-market contact in competition (Busse 2000), and demand (Iyengar, Jedidi and Kohli 2008, Huang 2008).

3 Data and Stylized Facts

3.1 Data

We use two sets of data: First, we use a panel of individual monthly billing records for all student enrollees in cellular-phone plans offered by a national cellular carrier in conjunction with a major university from February 2002 to June 2005. This data set includes both monthly bill summaries as well as detailed call-level information for each subscriber.[14] Second, we acquired EconOne data on the prices and characteristics of all cellular-phone plans offered at the same dates in the vicinity of the university. The price menu offered to students differed from that offered by the carrier directly to the public. First, relative to public prices, the university negotiated that the carrier offer a 15% discount, the option of choosing a two-part tariff not available to the public, and other favorable terms such as a limited three-month contractual commitment. Second, the carrier offered different monthly promotions of additional *bonus* minutes to students than to the public. Third, the university levied an additional $5 per month surcharge on top of carrier charges to cover its administrative costs.

The bulk of our work makes use of the monthly bill summaries. For reasons discussed in Appendix A, we restrict attention to the period June 2002 to October 2004 and exclude individuals who are left censored (those who are existing subscribers at the start of the panel).[15] We focus on customer choice between four *popular* plans, that account for 89% of bills in our data. We group the remaining price plans with the outside option, and hence drop the 11% of bills with unpopular

[14] Students received an itemized phone bill, mailed by default to their campus residence, which was separate from their university tuition bill. The sample of students is undoubtedly different than the entire cellular-phone-service customer-base. However, a pricing manager from one of the top US cellular phone service providers made the unsolicited comment that the empirical patterns of usage, overages, and ex post "mistakes" documented in Grubb (2009) using the same data were highly consistent with their own internal analysis of much larger and representative customer samples.

[15] When we estimate the structural model of plan choice and usage, we use a somewhat shorter period of August 2002 to July 2004. We restrict the sample further because we have to infer the number of included minutes for each plan-month pair, and we felt we could only reliably do this for plans offered during these months. Appendix A describes the procedure we used to infer plan characteristics and provides an overview of plan offerings on a monthly basis.

price plans.[16] Finally, rate plan codes are frequently miss-coded as a default value on a customers initial bill, in which case we remove the first bill. Our final data set contains 1366 subscribers and 16,283 month-subscriber observations. Note that for much of our analysis, we also exclude pro-rated bills during months of partial service, or customer switching between plans (however, pro-rated bills are included in the sample we use to estimate the structural model).

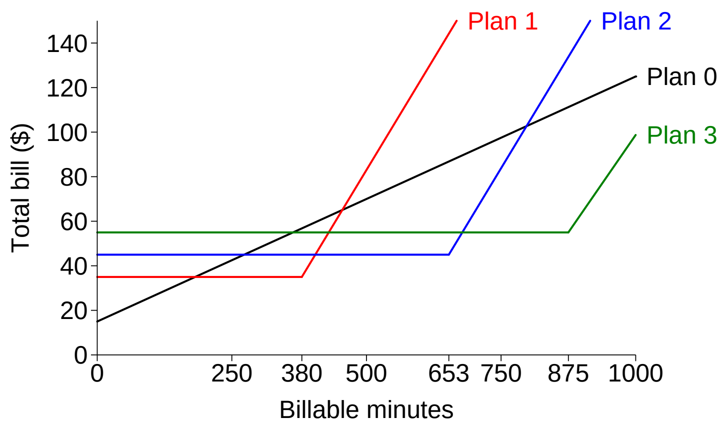
Every month, new subscribers were offered a choice of calling plans. There are four classes of calling plan: business, standard local, local with free-long-distance, and national. Business plans are two-part tariffs: a subscriber pays a monthly fee (typically \$14.99) and a flat per minute rate of 11 cents. All other plans are three-part tariffs: customers paid a monthly fee (M_j), received unlimited off-peak minutes (nights & weekends) and a number of free peak-minutes (Q_j), and paid an *overage* charge (p_j) of 35 to 45 cents per peak minute once the free minutes were used up. Local plans require calls to be made within the subscriber's calling area (the neighboring states within which they live) to avoid roaming charges of 66 cents per minute or more. Standard local plans are charged an additional 20 cents per minute for long distance. National plans offered both free long-distance and no roaming fees for all calls made within the United States. Table 1 summarizes plan shares. Business and standard local-plans account for over 90% of bills. Within these two plan classes, the most popular plans are the 14.99 business plan (plan 0), and the 34.99, 44.99, and 54.99 local plans (plans 1-3), which we refer to as the four *popular* plans and label plan 0 through plan 3 respectively. Shares of these four popular plans are highlighted in bold in Table 1. Prices of the four popular plans are shown graphically for the Spring of 2003 in Figure 1.

Table 1: Shares of Plan Types, By Monthly Fee and Class

Monthly Fixed Fee	Plan Class: Business	Local	Local, Free LD	National
14.99	**44.19**	0.00	0.00	0.00
34.99	0.00	**27.88**	1.28	1.88
44.99	0.00	**15.25**	0.38	3.46
54.99	0.00	**1.83**	0.11	0.60
other	0.75	0.64	0.00	1.76
	44.93	45.60	1.77	7.70

Plan shares are the percent of bills observed for each different access fee and plan class. Four *popular* plan shares are highlighted in bold. Together, these account for 89% of bills.

[16] In fact, we treat switching to an unpopular plan the same as quitting service, hence we also drop all remaining bills once a customer switches to an unpopular plan, even if they eventually switch back to a popular plan.



Plan	M_j	Q_j	p_j
Plan 0	\$14.99	0	\$0.11
Plan 1	\$34.99	380	\$0.45
Plan 2	\$44.99	653	\$0.40
Plan 3	\$54.99	875	\$0.40

Figure 1: Popular Plan Prices, Spring 2003.

Once a customer chose a plan, the plan terms remain fixed for that customer, regardless of any future promotions or discounts, until they switched plans or terminated service. However, the terms of any given plan vary significantly with promotions available at the date a customer chooses the plan. Plan terms varied significantly on three important dimensions. First, plan 0 included free off-peak (nights & weekends) calling for those who chose the plan in the 2002-2003 academic year, but the promotion was not offered to those who chose the plan in the 2003-2004 academic year. Second, some plans, such as plan 2, offered free in-network calling at some dates but not others. Finally, the number of free peak minutes included with plans 1-3 varied over time.

Prices of the four popular plans are described for all dates in Appendix A Table 13. In addition, the important price changes are highlighted in Appendix A Figure 9 along with a monthly tabulation of the total number of subscribers in the data set, the number of new subscribers, the number of existing subscribers switching plans, and the number of existing subscribers quitting (or switching to a non-popular plan). This price series was inferred from billing data rather than directly observed, as discussed in Appendix A.

3.2 Stylized Facts

3.2.1 Three stylized facts relevant to modeling usage choices

There are three important features of the data that are important to accurately model usage choices by customers of cellular phone service. First, consumers' usage choices are price sensitive. Second, consumers' usage choices are made while consumers are uncertain about the ex post marginal price. Third consumers are inattentive to the remaining balance of included minutes during the course of a billing cycle.

Consumer price sensitivity is clearly illustrated by a sharp increase in calling volume on weekday evenings exactly when the off-peak period for free night and weekend calling begins (Figure 2). This is not simply a 9pm effect, as the increase occurs only on weekdays, and at 8pm for plans with early nights-and-weekends.[17]

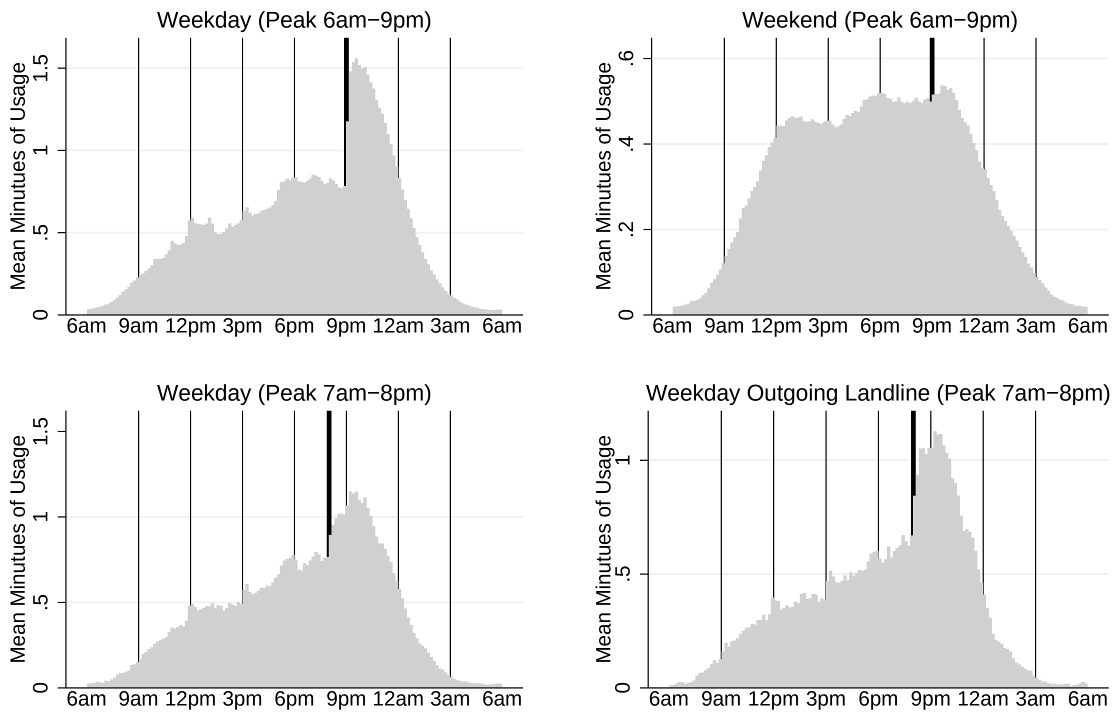


Figure 2: Daily usage patterns for subscribers with free nights and weekends. Top row: weekday (Panel A) and weekend (Panel B) usage patterns for subscribers with 6am-9pm peak hours. Bottom row: weekday usage patterns for subscribers with 7am-8pm peak hours. Panel C shows all weekday calling, while Panel D is restricted to outgoing calls to land-lines (recipients for whom the cost of receiving calls was zero). The patterns are qualitatively similar for bills with peak usage strictly below the free allowance.

Given clear sensitivity to marginal price, if consumers could anticipate whether they would be under their allowance (zero marginal price ex post) or over their allowance (35 to 45 cents per minute marginal price ex post) we would expect to see substantial bunching of consumers consuming their entire allowance but no more or less. Figure 3 shows this is not the case. Moreover, consumers who anticipate being strictly under their allowance (zero marginal price ex post) should exhibit no price response at the commencement of off-peak hours. However, as noted in the caption of

[17]For plans with free weeknight calling starting at 8pm, there is still a secondary increase in usage at 9pm (Figure 2 panel C). Restricting attention to outgoing calls made to land-lines almost eliminates this secondary peak (Figure 2 panel D). This suggests that the secondary peak is primarily due to calls to and from cellular numbers with 9pm nights (the most common time for free evening calling to begin) rather than a 9pm effect.

Figure 2, there is a sharp increase in calling at 9pm, even in months for which the peak allowance is under-utilized. This is a natural consequence of usage choice under uncertainty about ex post marginal price.

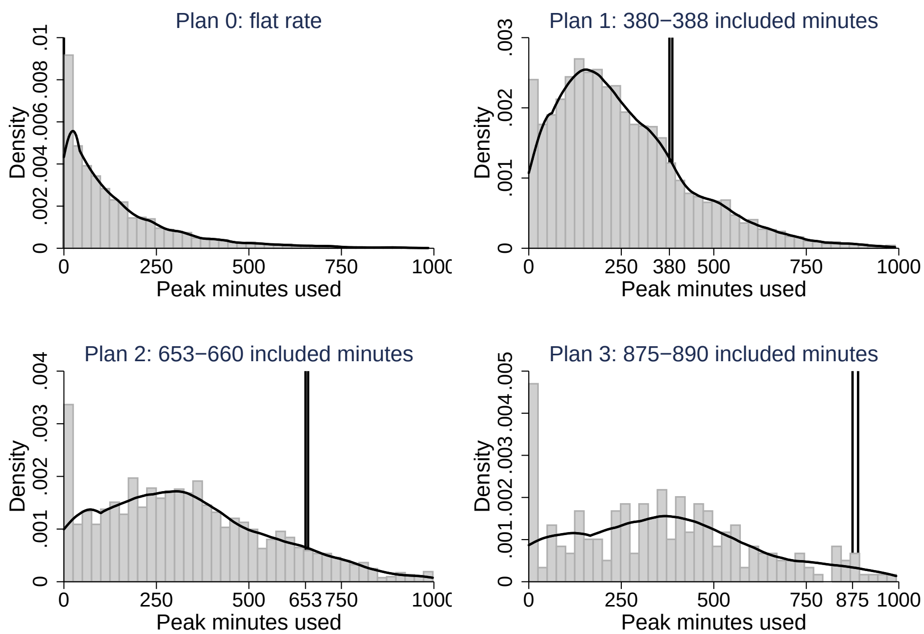


Figure 3: Usage densities for popular plans are constructed with 9,080, 5,026, 2,351, and 259 bills for plans 0-3 respectively. The sample for plans 1-3 is selected to only include bills for which in-network calls were costly and for which included peak minutes were within a narrow range, as indicated above each plot. Vertical lines bound the range of included free minutes for each plan.

If consumers are attentive to the remaining balance of included minutes during the billing cycle they should use this information to continually update their beliefs about the likelihood of an overage and a high marginal price ex post. Following an optimal dynamic program, an attentive consumer should (all else equal) reduce her usage later in the month following unexpectedly high usage earlier in the month. This should be true for any consumers who are initially uncertain whether they will have an overage in the current month. For these consumers, the high usage shock early in the month increases the likelihood of an overage, thereby increasing their expected ex post marginal price, and causing them to be more selective about calls. If calling opportunities arrived independently throughout the month, this strategic behavior by the consumer would lead to negative correlation between early and late usage within a billing period. However looking for negative correlation in usage within the billing period is a poor test for this dynamic behavior, because it is likely to be overwhelmed by positive serial correlation in taste shocks.

To test for dynamic behavior by consumers within the billing period, we use our data set of individual calls to construct both fortnightly and weekly measures of peak usage.[18] A simple regression of usage on individual fixed effects and lagged usage shows strong positive serial correlation. However, we take advantage of the following difference: Positive serial correlation between taste shocks in periods t and $(t-1)$ should be independent of whether periods t and $(t-1)$ are in the same or adjacent billing cycles. However, following unexpectedly high usage in period $(t-1)$, consumers should only cut back usage in period t if the two periods are in the same billing cycle. Thus by including an interaction effect between lagged usage and an indicator for the lag being in the same billing cycle as the current period, we can separate strategic behavior within the month from serial correlation in taste shocks.

Equation (1) describes our first specification, which appears in Table 2.

$$\ln(q_t) = \beta_{0,i,t} + \beta_1 \ln(q_{t-1}) + \beta_2 d_{t-1} \ln(q_{t-1}) \tag{1}$$

We include time and individual fixed effects ($\beta_{0,i,t}$) and use the Stata procedure xtabond2 to correct for bias induced by including both individual fixed effects and lags of the dependent variable in a wide but short panel (Roodman 2009). The indicator d_{t-1} is equal to 1 if period $(t-1)$ is in the same billing cycle as period t. If there is both positive serial correlation in demand shocks, and strategic behavior by the consumer within the billing cycle, then we expect β_1 to be positive (capturing serial correlation in shocks) and β_2 to be negative (capturing the strategic behavior). Reported analysis are for plan 1, the most popular three-part tariff. In our first specification, Column (1) of Table 2, β_2 has a negative point estimate, but is not significantly different from zero. This suggests that consumers are not attentive to past usage during the course of the month.

Consumers who either never have an overage (43% of plan 1 subscribers) or always have an overage (3% of plan 1 subscribers) should be relatively certain what their ex post marginal price will be, and need not adjust calling behavior during the month. For instance, consumers who always make overages may only make calls worth more than the overage rate throughout the month. For such consumers we would expect to find $\beta_2 = 0$, and this may drive the result when all consumers are pooled together as in our first specification. As a result, we divide consumers into groups by the fraction of times within their tenure that they have overages. We repeat our first specification for different *overage-risk groups* in Columns (2)-(6) of Table 2. The coefficient β_2 is indistinguishable from zero in all overage risk groups. Moreover, in unreported analysis, more flexible specifications that include nonlinear terms[19] and a similar analysis at the weekly rather than fortnightly level

[18]We divide each month into four weeks or two fortnights, and drop the extra 2-3 days between weeks 2 and 3.

[19]Average q_t will vary with expected marginal price, which is proportional to the probability of an overage. The

Table 2: Dynamic usage pattern at fortnightly level.

	(1)	(2)	(3)	(4)	(5)	(6)
Overage Percentage	0-100%	0	1-29%	30-70%	71-99%	100%
$\ln(q_{t-1})$	0.649***	0.607***	0.535***	0.499***	-1.046	0.958***
	(0.0258)	(0.0529)	(0.0431)	(0.0683)	(1.065)	(0.0441)
SameBill*$\ln(q_{t-1})$	0.0133	0.0245	0.0193	-0.0149	-0.0837	3.685
	(0.0107)	(0.0183)	(0.0181)	(0.0222)	(1.180)	(4.745)
Observations	9068	3727	3218	1830	217	76
Number of id	386	167	130	87	11	6

Standard errors in parentheses. Time and individual fixed effects, xtabond2. Key: *** p<0.01, ** p<0.05, * p<0.1

all estimate β_2 indistinguishable from zero. There is simply no evidence that we can find that consumers strategically cut back usage at the end of the month following unexpectedly high initial usage. We conclude that consumers are inattentive to their remaining balance of included minutes during the billing cycle.[20]

3.2.2 Three stylized facts relevant to modeling plan choices

There are three important features of the data which are important to accurately model plan choice by cellular phone service customers. First, consumers are uncertain about future usage levels when making plan choices. Second, consumers learn about their own usage levels over time, and switch plans in response. Third, consumers' prior beliefs are biased: in the short run, before learning and switching plans, consumer plan-choice mistakes are predictable and can be exploited for profit. (We assume that consumers always make optimal plan-choices conditional on beliefs. When initial

probability of an overage in a billing period which includes periods t and $(t-1)$ clearly increases nonlinearly in q_{t-1}. In one specification, we first fit a probit on the likelihood of an overage as a function of the first fortnights usage, and then used the estimated coefficients to generate overage probability estimates for all fortnights. We then included these (lagged) values as explanatory variables. In an alternative unreported specification we simply added polynomial terms of lagged q_{t-1}.

[20]It is perhaps not surprising that we found no evidence for consumers dynamically updating their usage plan during the month. To follow such a sophisticated dynamic optimization, consumers need to be very attentive. To respond to past usage, one must be aware of past usage. In a normal situation this requires calling an automated phone system for account information, or logging into a webpage, or keeping close mental track of calls. In this case, due to the fact that service was provided through an intermediary, the university, such account information was actually not available in the middle of a billing period.

choices are suboptimal in a predictable way, we refer to consumers' prior beliefs as biased.)

Consumers must be uncertain about their future usage when choosing calling plans, because calling plan choices frequently turn out to be suboptimal ex post. Figure 1 shows prices of the four most popular calling plans in our data (plans 0-3). Table 3 cross tabulates consumers' actual plan-choices (among popular plans 0-3) from October 2002 to August 2003 against the plan which would have been cheapest (holding actual usage fixed) over the duration of the customer's subscription to the chosen plan. The diagonal shows the number of consumers whose ex ante choices were optimal ex post. If consumers' faced no uncertainty about their own future usage, all consumers would lie on the diagonal. Instead, Table 3 shows that 29% of consumers made ex post plan-choice "mistakes" between October 2002 and August 2003. Table 4 shows even higher levels of ex post mistakes (45%) for the period September 2003 through July 2004. (The level of mistakes is lower in the earlier period because plan 0 initially offered free nights-and-weekends. As a result, for most consumers plan 0 dominated the other options by a large margin, which made making the ex post optimal choice relatively easy for most customers. A subset of these ex post mistakes are already documented in Grubb (2009).)

Table 3: Ex Post Plan Choice "Mistakes", 10/02-8/03

		Best Plan				
		Plan 0	Plan 1	Plan 2	Plan 3	Total
Chosen Plan	Plan 0	**464**	3	12	2	481
	Plan 1	61	**12**	21	7	101
	Plan 2	66	1	**39**	23	129
	Plan 3	9	0	7	**4**	20
	Total	600	16	79	36	731

Dates: 10/02-8/03, when Plan 0 included free nights & weekends. The "best" plan is that offered at the time of original choice which minimizes average expenditure holding usage fixed over the entire period the subscriber maintained their initial choice. 29% of subscribers made ex post plan-choice "mistakes".

Consumers switch plans over time. In some cases this may be in response to changes in tastes, or to price decreases which make previously unattractive plans more attractive. However the pattern of plan switches shows that they are also made in response to learning. There are 1366 customers in our data set, who we observe for an average of 12 months before either the data set ends or the customer quits.[21] Among all customers, 207 (15%) switch plans at least once, and 28 (2%) switch plans more than once, leading to a total of 246 plan switches (Table 5). Of these switches, 85 (35%)

[21] In our sample, 31 percent of customers are observed for more than 12 months. Standard cellular phone contracts

Table 4: Ex Post Plan Choice "Mistakes", 9/03-7/04

		Best Plan				
		Plan 0	Plan 1	Plan 2	Plan 3	Total
Chosen Plan	Plan 0	**129**	29	7	3	168
	Plan 1	66	**229**	123	27	445
	Plan 2	11	56	**81**	19	167
	Plan 3	4	3	17	**11**	35
	Total	210	317	228	60	815

Dates: 9/03-7/04, when Plan 0 did not include free nights & weekends. The "best" plan is that offered at the time of original choice which minimizes average expenditure holding usage fixed over the entire period the subscriber maintained their initial choice. 45% of subscribers made ex post plan-choice "mistakes".

are to plans that have either dropped in price or been newly introduced since the customer chose their existing plan. These switches could be motivated by price decreases rather than learning. However, the remaining 161 (65%) switches are to plans that are weakly more expensive than when the customer chose his or her existing plan. These switches must be due to learning or taste changes.

Table 5: Plan Switching

		New Plan				
		Plan 0	Plan 1	Plan 2	Plan 3	Total
Old Plan	Plan 0	0	27	25	6	58
	Plan 1	71	1	55	16	143
	Plan 2	9	16	7	6	38
	Plan 3	2	2	3	0	7
	Total	82	46	90	28	246

Switches on the diagonal represent an active switch to take advantage of an increase in the number of included minutes currently offered for the same plan.

Not only do consumers switch plans, but they switch in the "right" direction. To substantiate this claim we make two calculations. First we calculate how much the customer would have saved

often include switching costs (such as extension of commitment and delay of new phone subsidy) for switching plans prior to the expiry of one or two year contracts. In such a setting, more than 12 months of data would be needed to observe switching and learning. The students in our sample, however, could switch plans at any time and cancel after only three months, without any cost except hassle costs. As a result, we are able to observe active switching and learning over shorter time periods.

had they signed up for the new plan initially, holding their usage from the original plan fixed. By this calculation, 60 to 61 percent of switches which can not be explained by price decreases saved customers money. (Switches that can not be explained by price decreases are those to plans which are weakly more expensive at the switching date than at the initial choice date.) Average savings, across money saving and money losing switches, are \$11.03 to \$15.44 per month.[22] The savings estimates of \$11.03 to \$15.44 per month underestimate the benefit from switching plans, since they do not take into account the fact that consumers can re-optimize usage choices upon switching plans. For instance, when switching to a plan with more included minutes consumers may optimally choose to talk more in response to the lower marginal price. An upper bound on the value of these additional calls is their price under the old plan. Hence our second calculation is the money that would have been lost had the customer not switched plans and remained on their original plan, again holding usage fixed. By this calculation average savings for switching are \$24.42 to \$31.84 per month, and 68 to 75 percent of switches saved money.[23] Hence consumers' expected benefit is between \$11.03 and \$31.84 per month when switching to plans that have not decreased in price since their previous choice, and 60 to 75 percent of switches are in the "right" direction.

Additional evidence of plan switching due to learning is presented in Appendix H: (1) the likelihood of switching declines with tenure (Appendix H Figure 11), and (2) the likelihood of switching to a larger plan increases after an overage (Appendix H Table 15). Narayanan et al. (2007) estimate that consumers in the Kentucky experiment learn to switch up from overuse faster than they learn to switch down from underuse. For simplicity we implement symmetric learning in our structural model.

The presence of ex post mistakes alone shows only that consumers face uncertainty ex ante at the time of plan choice. However, ex post mistakes are not only present, they are also predictable. This implies that consumers' prior beliefs are biased and differ from average posteriors. Two plan-level savings opportunities demonstrate that customer mistakes are predictable and show how such predictability can be exploited by firms. The university acts as a reseller and charges students a fixed five dollar fee per month to cover administrative costs. Although the university did not do so,

---

[22] We calculate bounds because we cannot always distinguish in-network and out-of-network calls. Both figures are statistically greater than zero at the 99% level. The 60-61 percent rates of switching in the "right" direction are statistically greater than 50 percent at the 95% level. This calculation is based on 98 of the 161 switches which can not be explained by price decreases. The remaining 63 switches occur so soon after the customer joins that there is no usage data prior to the switch that is not from a pro-rated bill.

[23] This calculation is based on 157 of the 161 switches which can not be explained by price decreases. The calculation cannot be made for the remaining 4 switches since there is no usage data following the switch that is not from a pro-rated bill. Figures are significant at the 99% confidence level.

Table 6: Predictable Customer Mistakes Yield Savings Opportunities

	First Opportunity	Second Opportunity
Dates	10/02-8/03	9/03 onwards
Enrollment Change	plan 1-3 → plan 0	plan 1 → plan 2
Affected Customers	251 (34%)	445 (55%)
Savings		
Total	\$20,840 (47%)	\$7,942 (28%)
Per Affected Bill	\$8.76	\$2.64
Per Affected Cust.	\$83.03 (149%)	\$17.85 (46%)

The University acts as a reseller and could bill students for their chosen plan, sign them up for an alternative plan, and save the difference in charges. These plan-level savings opportunities indicate that consumers choose overly risky plans (overconfidence or projection bias). Savings estimates are a lower bound because we cannot always distinguish in and out-of-network calls.

they could have billed students based on the terms of their chosen calling plan, but signed them up for a predictably cheaper plan and saved the difference in charges. Table 6 illustrates two substantial opportunities. In the 2002-2003 academic year, when plan 0 offered free nights-and-weekends, by signing the 248 students who selected plans 1-3 up for plan 0, the university would have saved at least \$20,731, or \$83.59 per affected student. In the following year, the cellular company closed this opportunity by ending free nights-and-weekends on plan 0. However, an alternative was to sign up the 439 students who chose plan 1 onto plan 2, which would have saved at least \$7,934, or \$18.07 per affected student. These plan-level savings opportunities indicate that consumers choose overly risky plans (overconfidence or projection bias).[24]

4 Model

At each date t, consumer i first chooses a plan j and then chooses peak and off-peak quantities summarized by the vector $\mathbf{q}_{it} = (q_{it}^{pk}, q_{it}^{op})$. (The text suppresses the distinction between in-network and out-of-network calling, which is covered in Appendix D.) Total billable minutes for plan j are

$$q_{itj}^{billable} = q_{it}^{pk} + OP_j q_{it}^{op},$$

[24] Aggregate and conditional mean biases could explain one or other plan-level savings opportunity but only overconfidence and projection bias can simultaneously explain both savings opportunities. Note that the first savings opportunity is robust to dropping the top 30 percent of customers with the highest average savings, while the second savings opportunity is robust to dropping the top 2 percent of customers.

where OP_j is an indicator variable for whether plan j charges for off-peak usage. At the end of period t, consumer i is charged

$$P_j\left(\mathbf{q}_{it}\right) = M_j + p_j \max\{0, q_{itj}^{billable} - Q_j\},$$

where pricing plan j has monthly fee M_j, included allowance Q_j, and overage rate p_j.

We assume consumers are risk neutral, consumers have quasi-linear utility, and peak and off-peak calls are neither substitutes nor complements. Consumer i's utility in month t from choosing plan j and consuming $\mathbf{q}_{it}$ units is

$$u_{itj} = \sum_{k\in\{pk,op\}} V\left(q_{it}^k, \theta_{it}^k\right) - \alpha P_j\left(\mathbf{q}_{it}\right) + \eta_{itj},$$

where

$$V\left(q_{it}^k, \theta_{it}^k\right) = \frac{1}{\gamma}\left(\theta_{it}^k \ln\left(q_{it}^k/\theta_{it}^k\right) - q_{it}^k\right) \tag{2}$$

is the value from category $k \in \{pk, op\}$ calling, which depends on a pair of non-negative taste-shocks $\boldsymbol{\theta}_{it} = (\theta_{it}^{pk}, \theta_{it}^{op})$, and η_{itj} is an i.i.d. logit error.[25]

Define $q(p, \theta_{it}^k) \equiv \arg\max_q \left(V\left(q, \theta_{it}^k\right) - \alpha pq\right)$ to be a consumer's demand for category-k calls given a constant marginal price p. (This differs from a consumer's actual demand when marginal price varies with usage.) Define $\beta \equiv \gamma\alpha$. Then given equation (2),

$$q\left(p, \theta_{it}^k\right) = \theta_{it}^k / \left(1 + \beta p\right).$$

Note that $q\left(p, \theta_{it}^k\right)$ is multiplicative in θ_{it}^k, and can be expressed as the product

$$q(p, \theta_{it}^k) = \theta_{it}^k \hat{q}(p), \tag{3}$$

where $\hat{q}\left(p\right) = 1/\left(1 + \beta p\right)$ and $\hat{q}\left(0\right) = 1$.[26] The interpretation is that θ_{it}^k is the volume of category-k calling opportunities that arise and $\hat{q}(p)$ is the fraction of those calling opportunities worth more than p per minute.

[25] We model consumers' choice between the four most popular pricing plans (plans 0-3), comparable plans from other carriers, and an outside option. For plans other than the four popular university plans, the logit error η_{itj} has a clear economic interpretation: it includes all unmodeled plan heterogeneity including network quality, available phones, and roaming charges. Within the four popular plans, the logit error η_{itj} has no satisfactory economic interpretation, as these plans only differ in price, and in the complete model we capture all the dimensions on which prices differ. All initial plan choices could be explained without including the logit error, but they are required to explain switches that appear to be in the "wrong" direction.

[26] The fact that $\hat{q}(0) = 1$ simply reflects the chosen normalization of θ_{it}^k.

There are two price coefficients in the model, a contract price-coefficient α and a calling price-coefficient β. The contract price-coefficient α determines how sensitive plan choice is to overall plan cost including the plan fixed fee. The calling price-coefficient β determines how sensitive calling choices are to the marginal price of an additional minute of calling time.

4.1 Quantity Choices

Recognizing that consumers are uncertain about the ex post marginal price when making usage choices from three-part tariffs is a key feature of our model and where we take a new approach (also suggested independently by Borenstein (2009)). We assume that at the start of billing period t, consumer i is uncertain about her period t taste shock $\boldsymbol{\theta}_{it}$. She first chooses a plan j and then chooses a calling threshold vector $\mathbf{v}^*_{itj} = (v^{pk}_{itj}, v^{op}_{itj})$ based on chosen plan terms $\{Q_j, p_j, OP_j\}$ and her beliefs about the distribution of $\boldsymbol{\theta}_{it}$. During the course of the month, the consumer is inattentive and does not track usage but simply makes all category-k calls valued above v^k_{itj}. Over the course of the month, for $k \in \{pk, op\}$ this cumulates to the choice:

$$q^k_{it} = \theta^k_{it}\hat{q}(v^k_{itj}). \tag{4}$$

Timing is summarized in Figure 4. Figure 5 shows the calling threshold v^{pk}_{itj} and resulting consumption choice $\theta^{pk}_{it}\hat{q}(v^{pk}_{itj})$ in relation to a consumer's realized inverse demand curve for calling minutes, $V_q(q^{pk}_{it}, \theta^{pk}_{it})$.

Choose plan j given prior $\boldsymbol{\theta}_{it} \sim \tilde{F}_{it}$	Choose threshold $\mathbf{v}^*_{itj}$ given plan j and prior $\boldsymbol{\theta}_{it} \sim \tilde{F}_{it}$	Taste θ^k_{it} and usage $q^k_{it} = \theta^k_{it}\hat{q}(v^k_{itj})$ realized for $k \in \{\text{pk,op}\}$. Beliefs updated

Figure 4: Model Time Line

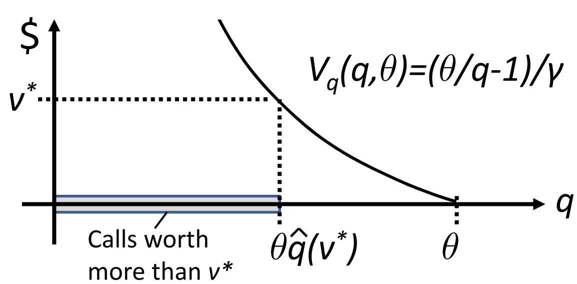


Figure 5: Inverse Demand Curve and Calling Threshold

Making all peak calls valued above the constant threshold v^*_{itj} is the optimal strategy of an inattentive consumer who does not track usage within the current billing cycle and hence cannot update his beliefs about the likelihood of an overage within the current billing cycle. (It is analogous to an electricity consumer setting a thermostat rather than choosing a quantity of kilowatt hours.)

When marginal price is constant, a consumer's optimal calling threshold is simply equal to the marginal price. Thus for plan zero, which charges 11 cents per minute for all billable calls, $\mathbf{v}_{itj}^{*} = (0.11, 0.11OP_j)$. Further, $v_{itj}^{op} = 0$ for plans 1-3 because they offer free off-peak calling.

Conditional choosing one of plans 1-3, which include free off-peak calling and an allowance of peak minutes, consumer i chooses her period t peak-calling threshold v_{itj}^{pk} to maximize her expected utility conditional on her period t information $\Im_{it}$:

$$v_{itj}^{pk} = \arg\max_{v_{itj}^{pk}} E\left[V\left(q(v_{itj}^{pk}, \theta_{it}^{k}), \theta_{it}^{k}\right) - \alpha P_j\left(q(v_{itj}^{pk}, \theta_{it}^{k})\right) \mid \Im_{it}\right].$$

Given allowance Q_j, overage rate p_j, and multiplicative demand (equation (3)), the optimal threshold (derived in Appendix B) is uniquely characterized by equation (5):

$$v_{itj}^{pk} = p_j \Pr\left(\theta_{it}^{pk} \geq Q_j/\hat{q}(v_{itj}^{pk})\right) \frac{E\left[\theta_{it}^{pk} \mid \theta_{it}^{pk} \geq Q_j/\hat{q}(v_{itj}^{pk});\ \Im_{it}\right]}{E\left[\theta_{it}^{pk} \mid \Im_{it}\right]}. \tag{5}$$

Note that the threshold v_{itj}^{pk} will be between zero and the overage rate p_j.

Equation (5) may seem counter-intuitive, because the optimal v_{itj}^{pk} is greater than the expected marginal price, $p_j \Pr(q(v_{itj}^{pk}, \theta_{it}^{pk}) > Q_j \mid \Im_{it})$. This is because the reduction in consumption from raising v_{itj}^{pk} is proportional to θ_{it}^{pk}. Raising v_{itj}^{pk} cuts back on calls valued at v_{itj}^{pk} more heavily in high demand states when they cost p_j and less heavily in low demand states when they cost 0. Note that choosing threshold v_{itj}^{pk} is equivalent to choosing a target calling quantity $q_{it}^{T} \equiv E[\theta_{it}^{pk}]\hat{q}(v_{itj}^{pk})$, which is implemented with error $(\theta_{it}^{pk} - E[\theta_{it}^{pk}])\hat{q}(v_{itj}^{pk})$. Importantly, consumers are aware of their inability to hit the target precisely and take this into account when making their threshold/target choice.

4.2 Plan Choices

We model consumers' choice between the four most popular pricing plans (plans 0-3), comparable AT&T, Cingular, and Verizon plans (Sprint offered no local plans), and an outside option which incorporates all other plans. We adopt Ching, Erdem and Keane's (2009) consideration set model by assuming that consumers make an active choice with exogenous probability P_C and keep their current plan with probability $(1 - P_C)$. We use the frequency of failures to switch away from dominated plans to identify P_C.[27]

Customer i's perceived expected utility from choosing plan j at date t is

$$U_{itj} = E\left[\sum_{k\in\{pk,op\}} V\left(q(v_{itj}^k, \theta_{it}^k), \theta_{it}^k\right) - \alpha P_j\left(\mathbf{q}(\mathbf{v}_{itj}^*, \boldsymbol{\theta}_{it})\right) \mid \Im_{it}\right] + \eta_{itj}, \tag{6}$$

and from choosing the outside option is $U_{it0} = O + \eta_{it0}$. The parameter O will be identified from the frequency at which consumers leave the data set. Conditional on making an active choice, a consumer's consideration set includes plans offered by her current provider, the outside option, and plans from a randomly selected alternative carrier.[28] Consumers myopically[29] choose the plan (or outside option) from their consideration set with the maximum expected utility for the current period.

4.3 Distribution of Tastes

We assume that the non-negative taste-shocks which determines usage are latent taste shocks censored at zero:

$$\theta_{it}^k = \begin{cases} 0 & \tilde{\theta}_{it}^k < 0 \\ \tilde{\theta}_{it}^k & \tilde{\theta}_{it}^k \geq 0 \end{cases}, \; k \in \{pk, op\}.$$

We assume that the latent shock $\tilde{\theta}_{it}^k$ is normally distributed and that consumers observe its value even when censored. This adds additional unobserved heterogeneity to the model but preserves tractable Bayesian updating. Censoring makes zero usage a positive likelihood event, which is important since it occurs for 10% of plan 0 observations.

Usage choices in the data are strongly serially-correlated conditional on customer-plan and date fixed effects. We therefore incorporate simple serial-correlation into our model by assuming that the latent shock $\tilde{\boldsymbol{\theta}}_{it}$ follows a stationary AR1 process with a bivariate normal innovation,

$$\tilde{\boldsymbol{\theta}}_{it} = \boldsymbol{\mu}_i + \varphi\tilde{\boldsymbol{\theta}}_{i,t-1} + \boldsymbol{\varepsilon}_{it},$$

[27] When prices fall consumers often do not switch away from their existing plans even when they are now dominated by plans on the current menu. For instance, most consumers paying \$54.99 for 890 minutes on plan 3 do not switch to plan 2 during the one month promotion in April 2004 when it offered 1060 minutes for only \$44.99. We believe this is because consumers who are not actively making a plan choice do not find out about the price cuts.

[28] We avoid including all plans in the consideration set to reduce computational time.

[29] We assume learning is independent of plan choice, so there is no value to experimentation with an alternative plan. Nevertheless, myopic plan choice is not always optimal. When a consumer is currently subscribed to a plan that is no longer offered (and is not dominated) there is option value to not switching, since switching plans will eliminate that plan from future choice sets. We ignore this issue for tractability.

where $\boldsymbol{\mu}_i$ is customer i's mean-type, φ is the common serial coefficient, and $\boldsymbol{\varepsilon}_{it} \sim N(0, \boldsymbol{\Sigma}_\varepsilon)$ is the normally-distributed mean-zero innovation with variance-covariance matrix

$$\boldsymbol{\Sigma}_\varepsilon = \begin{bmatrix} (\sigma_\varepsilon^{pk})^2 & \rho_\varepsilon \sigma_\varepsilon^{pk} \sigma_\varepsilon^{op} \\ \rho_\varepsilon \sigma_\varepsilon^{pk} \sigma_\varepsilon^{op} & (\sigma_\varepsilon^{op})^2 \end{bmatrix}.$$

(We assume AR(1) rather than AR(k) for simplicity.) Consumer types, $\boldsymbol{\mu}_i = (\mu_i^{pk}, \mu_i^{op})$, are normally distributed across the population as described below.

4.3.1 Near 9pm calling

Although prices in the model depend only on total peak and total off-peak calling, we additionally break out the share of calling demand for weekday outgoing-calls to landlines immediately before and after 9pm to help identify the calling price-coefficient. The shock $\mathbf{r}_{it}^{9pm} = (r_{it}^{9pk}, r_{it}^{9op}) \in [0,1]^2$ captures the share of peak and off-peak calling demand that is within 60 minutes of 9pm on a weekday and is for an outgoing call to a landline. The distribution of r_{it}^k for $k \in \{9pk, 9op\}$ is a censored normal,

$$\begin{aligned} \tilde{r}_{it}^k &= \alpha_i^k + e_{it}^{r,k} \\ r_{it}^k &= \begin{cases} 0 & \text{if} \quad \tilde{r}_{it}^k \leq 0 \\ \tilde{r}_{it}^k & \text{if} \quad 0 < \tilde{r}_{it}^k < 1 \\ 1 & \text{if} \quad \tilde{r}_{it}^k \geq 1 \end{cases}, \end{aligned}$$

where α_i^k is unobserved heterogeneity and $e_{it}^{r,k}$ is a mean-zero shock normally distributed with variance $(\sigma_e^k)^2$ independent across i, t, and k. We assume that α_i^{9pk} is normally distributed in the population with mean μ_α^{9pk} and variance $(\sigma_\alpha^{9pk})^2$.

Our identifying assumption for the calling price-coefficient is that consumer i's expected outgoing calling demand to landlines on weekdays is the same between 8:00pm and 9:00pm as it is between 9:00pm and 10:00pm:

$$E\left[r_{it}^{9pk}\right] E\left[\theta_{it}^{pk}\right] = E\left[r_{it}^{9op}\right] E\left[\theta_{it}^{op}\right]. \tag{7}$$

In other words, we assume that the increase in observed calling to landlines on weekdays immediately after off-peak begins at 9pm is a price effect rather than a discontinuous increase in demand at

9pm.[30] As a result, equation (7) implicitly defines α_i^{9op} as a function of α_i^{9pk} and other parameters.

4.4 Beliefs and Learning

Estimation of consumer beliefs and learning is focused on a single dimension of usage: total peak-calling. This is because plans 1-3 always offer free off-peak calling and hence the choice data are not rich enough to allow us to identify beliefs about off-peak calling. For simplicity, we assume that while consumers are learning about their peak type μ_i^{pk} over time, there is no learning about off-peak demand because consumers know their off-peak types μ_i^{op}.

We assume the serial-correlation coefficient φ is known by all consumers. While taste innovations $\boldsymbol{\varepsilon}_{it}$ have variance-covariance $\boldsymbol{\Sigma}_\varepsilon$, consumers believe the variance-covariance matrix is

$$\tilde{\boldsymbol{\Sigma}}_\varepsilon = \begin{bmatrix} (\tilde{\sigma}_\varepsilon^{pk})^2 & \rho_\varepsilon \tilde{\sigma}_\varepsilon^{pk} \sigma_\varepsilon^{op} \\ \rho_\varepsilon \tilde{\sigma}_\varepsilon^{pk} \sigma_\varepsilon^{op} & (\sigma_\varepsilon^{op})^2 \end{bmatrix},$$

where $\tilde{\sigma}_\varepsilon^{pk} = \delta_\varepsilon \sigma_\varepsilon^{pk}$ and $\delta_\varepsilon > 0$. If $\delta_\varepsilon = 1$, then consumers' perceptions match reality. If $\delta_\varepsilon < 1$, then consumers underestimate the volatility of their peak tastes from month-to-month and exhibit projection bias. If $\delta_\varepsilon < 1$, then consumers will predictably choose too risky plans and overreact to past usage when deciding whether or not to switch plans.[31] Consumer beliefs about the variance of off-peak tastes and the correlation between peak and off-peak tastes are both correct.

Consumers learn about their own peak-type μ_i^{pk} over time. At date t, consumer i believes that μ_i^{pk} is normally distributed with mean $\tilde{\mu}_{i,t}^{pk}$ and variance $\tilde{\sigma}_t^2$: $\mu_i^{pk}|\Im_{i,t} \sim N(\tilde{\mu}_{i,t}^{pk}, \tilde{\sigma}_t^2)$. At the end of each billing period, usage q_{it}^{pk} is realized and consumers can infer $\theta_{it}^{pk} = q_{it}^{pk}/\hat{q}(v_{itj}^{pk})$. When $q_{it}^{pk} = \theta_{it}^{pk} = 0$, we assume that consumers can observe the latent taste shock $\tilde{\theta}_{it}^{pk}$. The latent shock provides an unbiased normal signal about μ_i^{pk}.[32] In particular, at the end of the first billing period

[30] We focus on calls to landlines because the other party to the call pays nothing both before and after 9pm. The assumption would be unreasonable for calls to or from cellular numbers since such calling opportunities increase at 9pm when the calls become cheaper for the other party and the other party is more likely to call or answer.

[31] Our model assumes that projection bias does not disappear as consumers learn over time. This is consistent with evidence on projection bias (Loewenstein et al. 2003). For instance, as Loewenstein et al. (2003) note, "Several studies lend support to the folk wisdom that shopping on an empty stomach leads people to buy too much" (Nisbett and Kanouse 1968, Read and van Leeuwen 1998, Gilbert, Gill and Wilson 2002).

[32] In fact, given our assumption that consumers know μ_i^{op}, consumers can also infer ε_{it}^{op} from off peak usage which is informative about μ_i^{pk} because it is correlated with ε_{it}^{pk}. We assume consumers only update beliefs using θ_{it}^{pk} and not ε_{it}^{op}. This choice is conservative in the sense that our finding that consumers respond to data too little is biased downwards. It is also realistic for two reasons. First, consumers are unlikely to pay attention to off-peak usage when they are on contract with free off-peak calls. Second, we only assume consumers know μ_i^{op} for simplicity as we cannot identify off-peak beliefs. In reality, consumers are unlikely to know μ_i^{op} so cannot actually infer ε_{it}^{op}.

$t = 1$, consumer i learns

$$z_{i1} = (1 - \varphi)\,\tilde{\theta}_{i1}^{pk},$$

which she believes has distribution

$$N\left(\mu_i^{pk}, \frac{1-\varphi}{1+\varphi}(\tilde{\sigma}_{\varepsilon}^{pk})^2\right).$$

Then, in later periods $t > 1$, consumer i learns

$$z_{it} = \tilde{\theta}_{it}^{pk} - \varphi\tilde{\theta}_{i,t-1}^{pk},$$

which she believes has distribution

$$N\left(\mu_i^{pk}, (\delta_{\varepsilon}\sigma_{\varepsilon}^{pk})^2\right).$$

Define $\bar{z}_{it} = \frac{1}{t}\sum_{\tau=1}^{t} z_{i\tau}$. Then by Bayes rule (DeGroot 1970), updated time $t+1$ beliefs about μ_i^{pk} are $\mu_i^{pk}|\Im_{i,t+1} \sim N(\tilde{\mu}_{i,t+1}^{pk}, \tilde{\sigma}_{t+1}^2)$ where

$$\tilde{\mu}_{i,t+1}^{pk} = \frac{\tilde{\mu}_{i1}^{pk}\tilde{\sigma}_1^{-2} + \left(\frac{2\varphi}{1-\varphi}z_{i1} + t\bar{z}_{it}\right)(\tilde{\sigma}_{\varepsilon}^{pk})^{-2}}{\tilde{\sigma}_1^{-2} + \left(\frac{2\varphi}{1-\varphi} + t\right)(\tilde{\sigma}_{\varepsilon}^{pk})^{-2}}, \tag{8}$$

and

$$\tilde{\sigma}_{t+1}^2 = \left(\tilde{\sigma}_1^{-2} + \left(\frac{2\varphi}{1-\varphi} + t\right)(\tilde{\sigma}_{\varepsilon}^{pk})^{-2}\right)^{-1}.$$

Over time consumers learn their own types: $\tilde{\mu}_{i,t}^{pk}$ converges to μ_i^{pk} and $\tilde{\sigma}_t^2$ converges to zero.

Consumers' plan choices and threshold choices depend on beliefs about the distribution of tastes $\boldsymbol{\theta}_{it}$. When choosing a plan and a usage threshold for the first time, consumers believe:

$$\tilde{\theta}_{i1}^{pk} \sim N\left(\frac{\tilde{\mu}_{i1}^{pk}}{1-\varphi}, \tilde{\sigma}_{\theta 1}^2\right), \tag{9}$$

where

$$\tilde{\sigma}_{\theta 1}^2 = \frac{\tilde{\sigma}_1^2}{(1-\varphi)^2} + \frac{(\tilde{\sigma}_{\varepsilon}^{pk})^2}{1-\varphi^2}. \tag{10}$$

In all later periods $t > 1$, when consumers can condition on $\tilde{\theta}_{i,t-1}^{pk}$, beliefs are:

$$\tilde{\theta}_{it}^{pk} \mid \Im_{it} \sim N\left(\tilde{\mu}_{it}^{pk} + \varphi\tilde{\theta}_{it-1}^{pk}, \tilde{\sigma}_t^2 + (\delta_{\varepsilon}\sigma_{\varepsilon}^{pk})^2\right).$$

Following a month with surprisingly high usage, consumer i's beliefs about the distribution of demand in the following month increases for two reasons. First the consumer increases his estimate of his type ($\tilde{\mu}^{pk}_{i,t+1} > \tilde{\mu}^{pk}_{it}$), and second he knows that his demand is positively correlated over time. In the standard model the only behavior change that might result is a switch to a larger plan. In our model, a consumer might also switch to a larger plan but, conditional on not switching, would cut back on usage by choosing a higher calling threshold ($v^{pk}_{i,t+1} > v^{pk}_{i,t}$) and being more selective about calls.

4.5 Priors

Each customer is characterized by the individual specific triple $\{\mu^{pk}_i, \mu^{op}_i, \tilde{\mu}^{pk}_{i1}\}$. Together with the population parameter $\tilde{\sigma}^2_1$, this triple specifies each customer's true mean-type $\boldsymbol{\mu}_i$ and prior belief $\mu^{pk}_i \sim N(\tilde{\mu}^{pk}_{i1}, \tilde{\sigma}^2_1)$. (Consumers are assumed to know their own off-peak types.) The population is described by the joint distribution of $\{\mu^{pk}_i, \mu^{op}_i, \tilde{\mu}^{pk}_{i1}\}$. We assume that $\{\mu^{pk}_i, \mu^{op}_i, \tilde{\mu}^{pk}_{i1}\}$ has a trivariate normal distribution. Specifically, the marginal distribution of initial point estimates is

$$\tilde{\mu}^{pk}_{i1} \sim N(\tilde{\mu}^{pk}_0, \tilde{\sigma}^2_{\mu^{pk}}),$$

and the population distribution of true-types $\boldsymbol{\mu}_i$ conditional on the point estimate is

$$\boldsymbol{\mu}_i \mid \tilde{\mu}^{pk}_{i1} \sim N\left(\boldsymbol{\mu}_0 + \psi\left(\tilde{\mu}^{pk}_{i1} - \tilde{\mu}^{pk}_0\right), \boldsymbol{\Sigma}_\mu\right), \tag{11}$$

where $\boldsymbol{\mu}_0 = (\mu^{pk}_0, \mu^{op}_0)$, $\psi = (\psi^{pk}, \psi^{op})$, and

$$\boldsymbol{\Sigma}_\mu \equiv \begin{bmatrix} \sigma^2_{\mu^{pk}} & \rho_\mu \sigma_{\mu^{pk}} \sigma_{\mu^{op}} \\ \rho_\mu \sigma_{\mu^{pk}} \sigma_{\mu^{op}} & \sigma^2_{\mu^{op}} \end{bmatrix}. \tag{12}$$

Here μ^{pk}_0 is the average true peak-type μ^{pk}_i and $\tilde{\mu}^{pk}_0$ is the average prior $\tilde{\mu}^{pk}_{i1}$. Similarly, $\sigma^2_{\mu^{pk}}$ is the conditional variance of true peak-types μ^{pk}_i and $\tilde{\sigma}^{pk}_\mu$ is the variance of priors $\tilde{\mu}^{pk}_{i1}$.

Let $b_1 = \tilde{\mu}^{pk}_0 - \mu^{pk}_0$ and $b_2 = 1 - \psi^{pk} + \psi^{op}\rho_\mu\left(\sigma_{\mu^{pk}}/\sigma_{\mu^{op}}\right)$. Then (as shown in Appendix C) taking expectations over the population distribution of tastes,

$$\tilde{\mu}^{pk}_{i1} - E\left[\mu^{pk}_i \mid \tilde{\mu}^{pk}_{i1}\right] = b_1 + b_2(\tilde{\mu}^{pk}_{i1} - \tilde{\mu}^{pk}_0). \tag{13}$$

A typical assumption (perhaps labeled rational expectations) is that $\tilde{\mu}^{pk}_{i1} - E[\mu^{pk}_i \mid \tilde{\mu}^{pk}_{i1}]$, or $b_1 = b_2 = 0$, which implies that individuals' initial point-estimates are unbiased estimates of their true types.[33] We do not impose this assumption. If $b_1 \neq 0$, then there is aggregate mean bias and consumers will predictably choose plans which are too small ($b_1 < 0$) or too large ($b_1 > 0$). If $b_2 \neq 0$, then there is conditional mean bias and consumers will predictably choose plans which are too moderate ($b_2 < 0$) or too extreme ($b_2 > 0$). (In the context of grocery home delivery service, Goettler and Clay (2010) find $b_2 > 0$ but do not reject $b_1 = 0$.)

Let $\delta_\mu = \tilde{\sigma}_1/(\sigma_{\mu^{pk}}\sqrt{1-\rho_\mu^2})$. Then (as shown in Appendix C), taking expectations over the population distribution of tastes:

$$\tilde{\sigma}_1 = \delta_\mu \sqrt{Var(\mu_i^{pk} \mid \tilde{\mu}_{i1}^{pk}, \mu_i^{op})}. \tag{15}$$

A typical assumption (perhaps labeled rational expectations) is that $\delta_\mu = 1$. We do not impose this assumption. If $\delta_\mu < 1$ then consumers exhibit overconfidence: they underestimate their own uncertainty about their type μ_i^{pk}.[34] Overconfident consumers, like those with projection bias, will predictably choose overly risky plans. However, in contrast to those with projection bias, they will under-react to past usage when making plan switching decisions. Grubb's (2009) analysis is static, so could not distinguish between overconfidence and projection bias, but found that customers do choose overly risky plans, so exhibit either overconfidence, projection bias, or both.

Note that the joint distribution of true types and priors described above can naturally be generated from the marginal distribution of true types, a normal common prior, and an unbiased normal signal with misperceived mean and variance. This is the presentation adopted by Goettler and Clay (2010).

---

[33]Strictly speaking, rational expectations would also impose the stronger condition $\tilde{\mu}_{i1}^{pk} - E\left[\mu_i^{pk} \mid \tilde{\mu}_{i1}^{pk}, \mu_i^{op}\right] = 0$ because we have assumed that consumers know μ_i^{op}. Let $b_3 = -\rho_\mu\left(\sigma_{\mu^{pk}}/\sigma_{\mu^{op}}\right)$. Then (as shown in Appendix C), taking expectations over the population distribution of tastes,

$$\tilde{\mu}_{i1}^{pk} - E\left[\mu_i^{pk} \mid \tilde{\mu}_{i1}^{pk}, \mu_i^{op}\right] = b_1 + b_2(\tilde{\mu}_{i1}^{pk} - \tilde{\mu}_0^{pk}) + b_3\left(\mu_i^{op} - \mu_0^{op}\right). \tag{14}$$

Therefore the additional restriction would be $b_3 = 0$. Our estimate of b_3 is negative, implying that consumers underreact to the information in their off-peak type when forming beliefs about their peak-type. However, we do not interpret this as an additional dimension of conditional mean bias, but rather the fact that consumers do not actually know their off-peak types. We assume that consumers do perfectly know their off-peak type only because variation in plan prices for off-peak usage is insufficiently rich to identify off-peak beliefs.

[34]When defining δ_μ, we calculate the population-variance of peak-types conditional both on a consumer's point estimate $\tilde{\mu}_{i1}^{pk}$ and her off-peak type μ_i^{op} because both are in a consumer's information set. While consumers may not actually know their off-peak types, this assumption is conservative in the sense that the alternative definition $\delta_\mu = \tilde{\sigma}_1/\sqrt{Var\left(\mu_i \mid \tilde{\mu}_{i1}\right)}$ would yield more overconfidence.

4.6 Comment about Risk Aversion

Our utility specification assumes that consumers are risk neutral. Our data do not allow us to separately identify preferences over risk from beliefs about risk. We assume risk neutrality and use plan choices to identify beliefs. In related work on health plan choice, Cullen, Einav, Finkelstein, Ryan and Schrimpf (2010) assume that subjective beliefs coincide with objective probabilities and use plan choices to identify risk preferences. Following our approach, we find that consumers are overconfident. As a result, if we followed Cullen et al.'s (2010) approach we would estimate that consumers are risk loving. We find it implausible that consumers take pleasure in the risk of accruing a high cell-phone bill. If one believes that consumers are in fact weakly risk-averse, then our estimates of overconfidence and projection bias are lower bounds on consumers' bias.

4.7 Identification

Parameters can be categorized into three groups: (1) price coefficients (α and β), (2) parameters governing beliefs ($\tilde{\mu}_0^{pk}$, $\tilde{\sigma}_{\mu^{pk}}$, $\tilde{\sigma}_1$, and $\tilde{\sigma}_{\varepsilon}^{pk}$), and (3) the true (conditional) distribution of tastes ($\boldsymbol{\mu}_0$, ψ, $\boldsymbol{\Sigma}_\mu$, $\boldsymbol{\Sigma}_\varepsilon$, and φ). Broadly speaking, plan choices identify beliefs, the distribution of actual usage identifies the distribution of true tastes, and changes in usage in response to the discontinuous change in marginal price between peak and off-peak hours identify the calling price coefficient β. Finally the rate of switching in the "wrong" direction identifies the contract price coefficient α.

4.7.1 Calling Price Coefficient

If consumers' chosen thresholds ($\mathbf{v}_{it}^*$) were known, the calling price-coefficient β could be inferred from marginal price variation and the induced variation in $\hat{q}(v_{it}^k)$. Unfortunately, we require β to calculate $\mathbf{v}_{it}^*$. We circumvent this problem by relying on a source of marginal price variation for which v_{it}^* is known. Prior to fall 2003, v_{it}^* is 11 cents during peak hours and 0 cents during off-peak hours for plan 0 subscribers.

Our identifying assumption in equation (7) is that underlying demand varies continuously over the hours of the day so that demand is the same on weekdays one hour before and one hour after 9pm. (Specifically we make this assumption only for calls to landlines since call demand will increase at 9pm for cellular calls because cellular subscribers at the other end of the phone line are more likely to place and answer calls when their calls become off-peak). Thus the discontinuous increase in calling at 9pm on weekdays is attributed entirely to the price response.

Given plan 0 pricing prior to fall 2003, $\theta_{it}^{op} = q_{it}^{op}$ and $\theta_{it}^{pk} = q_{it}^{pk}\,(1 + 0.11\beta)$. Moreover, the pre and post 9pm calling shares are always observed because calling thresholds are constant within peak and within off-peak hours: $r_{it}^{9op} = q_{it}^{9op}/q_{it}^{op}$ and $r_{it}^{9pk} = q_{it}^{9pk}/q_{it}^{pk}$. Thus equation (7) can be solved for β as a function of moments of the data:

$$\beta = \frac{100}{11}\left(\frac{E\left[q_{it}^{9op}/q_{it}^{op}\right]E\left[q_{it}^{op}\right]}{E\left[q_{it}^{9pk}/q_{it}^{pk}\right]E\left[q_{it}^{pk}\right]} - 1\right).$$

There is a second source of marginal price variation for which v_{it}^{*} is known. Off-peak v_{it}^{*} is either eleven cents for plan 0 in fall 2003 or zero cents for all other plans with free nights-and-weekends. Comparing usage within individuals who switch between these plans (97 switches) or across individuals on the different plans helps identify the price coefficient. This variation is less satisfactory, however, because the price coefficient will be confounded with selection effects and identification relies on our having correctly modeled and controlled for it. For instance, when consumers switch plans in response to a change in tastes, changes in usage after the switch result from the taste change as well as a price response. We incorporate and control for this in the model through the AR(1) persistence in tastes. Other price variation (Appendix A Figure 9).is less useful without knowing consumer thresholds. For instance, there is one clean experiment in the data in which existing plan 1 subscribers were automatically upgraded from 280 free minutes to 380 free minutes and increased their usage in response by an average of 53 minutes. (The 95% confidence interval on this increase is 26-81 minutes.) However without knowing how consumer thresholds were affected by the price change this does not identify β.

4.7.2 Serial Correlation

Data prior to fall 2003 identifies the AR1 coefficient φ. During this period, all plans offered free nights-and-weekends so that we observe

$$q_{it}^{op} = \theta_{it}^{op} = \mu_i^{op} + \varphi\theta_{it-1}^{op} + \epsilon_{it}^{op}. \tag{16}$$

The argument follows the identification argument for the parameters of a linear regression model with person level fixed effects and a lagged dependent variable. By taking the first difference of equation (16), we remove the impact of the fixed effect μ_i^{op}. Then φ can be estimated using past values of θ_{it}^{op} as instruments, as in Blundell and Bond (1998).

4.7.3 Beliefs

Next, consider identification of consumers' prior beliefs from plan choices. Choice data is quite informative about beliefs about peak usage, as illustrated by Figure 6, but relatively uninformative about beliefs about off-peak usage. Hence we assume consumers know their own off-peak taste distribution (including μ_i^{op} and σ_ε^{op}). Prior to fall 2003, when off-peak calling is free, an individual consumer's plan choice depends only on α, β and her beliefs about θ_{i1}^{pk} described by $\tilde{\mu}_{i1}^{pk} / (1-\varphi)$ and $\tilde{\sigma}_{\theta 1}$. Thus initial plan-choice shares depend only on α, β, φ, $\tilde{\sigma}_{\theta 1}$, and the population distribution of $\tilde{\mu}_{i1}^{pk}$, described by $\tilde{\mu}_0^{pk}$ and $\tilde{\sigma}^2_{\mu^{pk}}$. Parameters φ and β are already identified. For transparency of the argument, we begin by considering a restricted model that excludes logit errors ($1/\alpha = 0$). Initial plan choice shares identify the remaining parameters $\tilde{\mu}_0^{pk}$, $\tilde{\sigma}^2_{\mu^{pk}}$, and $\tilde{\sigma}_{\theta 1}$. Finally, the learning rate separately identifies δ_μ and δ_ε from $\tilde{\sigma}_{\theta 1}$. Initial choice shares in post fall 2003 data also aid identification, but require a more complicated argument involving beliefs about off-peak tastes.

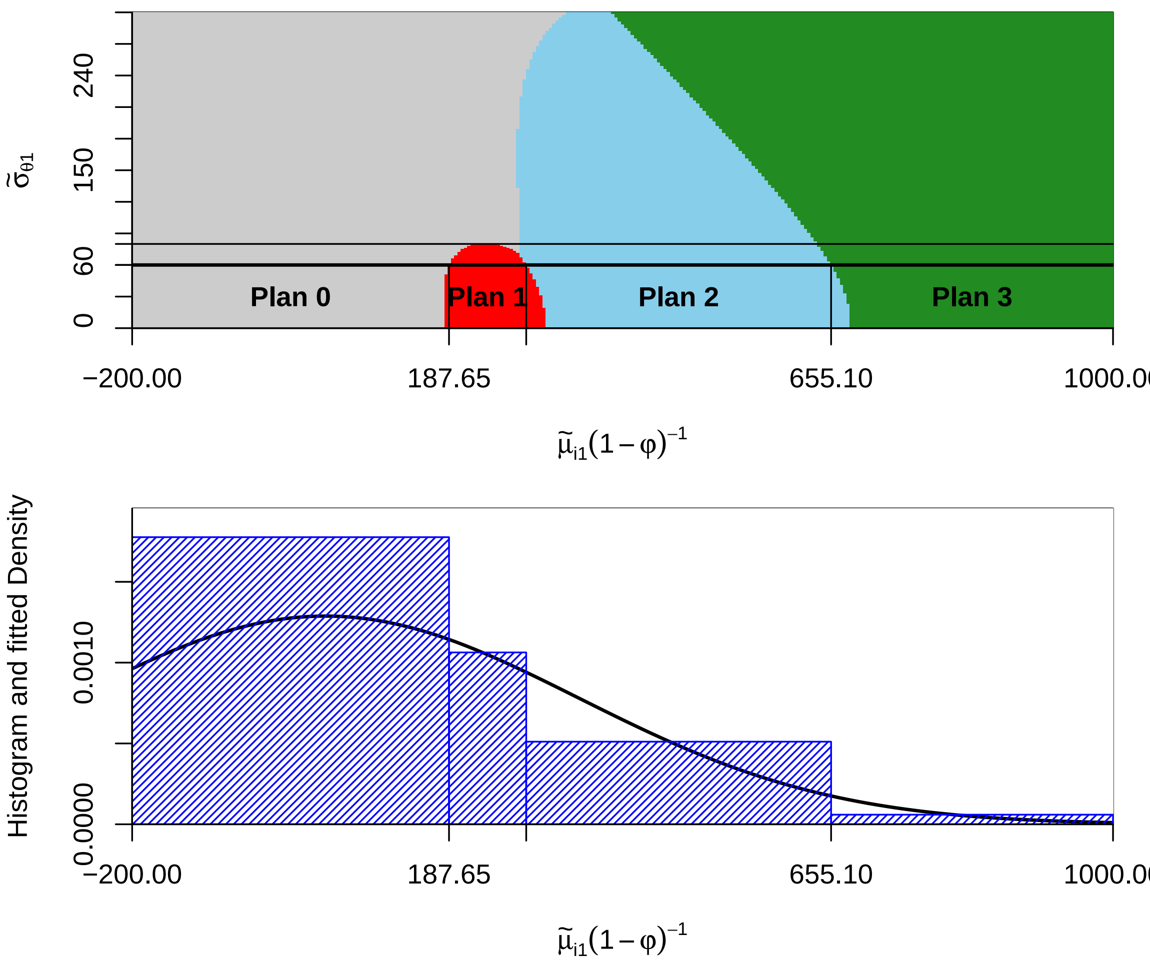


Figure 6: Top panel: Plan choice as a function of initial beliefs $\{\tilde{\mu}_{i1}(1-\phi)^{-1}, \tilde{\sigma}_{\theta 1}\}$ implied by the model evaluated at October-November 2002 prices given $\beta = 3.41$. Bottom panel: Histogram and fitted normal distribution over $\tilde{\mu}_{i1}(1-\phi)^{-1}$ implied by the assumption $\tilde{\sigma}_{\theta 1} = 80$ and October-November 2002 new subscriber plan choice shares of 69%, 10%, 19%, and 2% for plans 0 to 3 respectively.

Absent the logit-error, initial plan choices place bounds on each individual's prior beliefs about the mean ($\tilde{\mu}_{i1}^{pk}/(1-\varphi)$) and variance ($\tilde{\sigma}_{\theta 1}^2$) of their first taste shock, $\tilde{\theta}_{i1}^{pk}$. (Recall $\tilde{\sigma}_{\theta 1}^2$ is related to model parameters by equation (10).) Based on October-November 2002 pricing data (ignoring free in-network calling), Figure 6 (top panel) shows plan-choice as a function of prior beliefs $\{\tilde{\mu}_{i1}^{pk}/(1-\varphi),\, \tilde{\sigma}_{\theta 1}^2\}$ given $\beta = 3.41$.[35] Consumers joining in October-November 2002 with beliefs in the gray region choose plan 0, those with beliefs in the red region choose plan 1, those with beliefs in the blue region choose plan 2, and those with beliefs in the green region choose plan 3. This means that observing a new customer in October-November 2002 choose plan j will bound her beliefs to be within the relevant colored region.

Notice in Figure 6, that plan 0 is chosen both by individuals with low expectations of usage (low $\tilde{\mu}_{i1}^{pk}/(1-\varphi)$) since it has the lowest fixed fee, and by individuals with high uncertainty about usage (high $\tilde{\sigma}_{\theta 1}$) since it never charges more than 11 cents per minute and is therefore a safe option. Figure 6 shows that for any $\tilde{\sigma}_{\theta 1}$ larger than 114, plan 1 is never chosen. Thus the assumption that $\tilde{\sigma}_{\theta 1}$ is common across individuals and the fact that a sizable fraction of individuals chose plan 1 in October-November 2002 puts an upper bound on $\tilde{\sigma}_{\theta 1}$ of 114. (The implied upper bound in the structural model is lower since there we account for the fact that plan 0 was the only plan in fall 2002 to offer free in-network-calling.)

If we were to fix $\tilde{\sigma}_{\theta 1}$ at any level below 114, individual i's plan choice bounds $\tilde{\mu}_{i1}^{pk}/(1-\varphi)$ to an interval. For instance, if overconfidence and projection bias were complete ($\tilde{\sigma}_1 = \delta_\mu = \delta_\varepsilon = 0$) so that consumers believed they could predict their usage perfectly ($\tilde{\sigma}_{\theta 1} = 0$) and consumers were inelastic ($\beta = 0$), then consumers would choose from the lower envelope of the tariff menu, and initial choice of plan j would imply the following bounds on the prior point estimate $\tilde{\mu}_{i1}^{pk}/(1-\varphi)$:

$$(M_j - M_{j-1})/p_{j-1} + Q_{j-1} \leq \frac{\tilde{\mu}_{i1}^{pk}}{1-\varphi} \leq (M_{j+1} - M_j)/p_j + Q_j.$$

For $\tilde{\sigma}_{\theta 1}$ and β strictly positive, the bounds do not have an analytical solution but can be read from the corresponding horizontal slice of Figure 6. For example, the bounds are given for $\tilde{\sigma}_{\theta 1} = 80$ by the vertical lines in Figure 6. Combining plan share data from customers who join in October-November 2002 with these bounds generates of histogram over $\tilde{\mu}_{i1}^{pk}/(1-\varphi)$ with four bins, one for each of the four pricing plans. Since we assume that $\tilde{\mu}_{i1}^{pk}/(1-\varphi)$ is normally distributed with mean $\tilde{\mu}_0/(1-\varphi)$ and standard deviation $\tilde{\sigma}_\mu/(1-\varphi)$, this histogram would then (over) identify

[35] As $\tilde{\mu}_{i1}^{pk}/(1-\varphi)$ and $\tilde{\sigma}_{\theta 1}^2$ are mean and variance parameters of a censored-normal distribution, Appendix G Figure 10 depicts the same information as Figure 6 mapped onto the space $E[\theta_{i1}^{pk}] \times SD[\theta_{i1}^{pk}]$ which is measured in minutes and may be more readily interpretable.

the distribution. The resulting histogram and fitted normal distribution, are both shown in the lower panel of Figure 6 for the case $\tilde{\sigma}_{\theta 1} = 80$ and $\beta = 3.41$.

The model identifies $\tilde{\sigma}_{\theta 1}$ as the value between 0 and 114 that generates the best fit between the histogram and the fitted normal distribution. Choosing a larger value for $\tilde{\sigma}_{\theta 1}$ simply implies a higher mean, but lower variance for the distribution of $\tilde{\mu}_{i1}^{pk}/(1-\varphi)$.[36] Given $\beta = 3.41$, the overall best fit is at $\tilde{\sigma}_{\theta 1} = 83.5$.

The preceding argument for identifying $\tilde{\sigma}_{\theta 1}$, $\tilde{\mu}_0^{pk}$, and $\tilde{\sigma}_{\mu^{pk}}$ clearly bounds $\tilde{\sigma}_{\theta 1} \leq 114$ (given $\beta = 3.41$) but then relies heavily on the functional form assumption that $\tilde{\mu}_{i1}^{pk}$ is normally distributed for point identification. Nevertheless, there is additional information in the data which reduces reliance on the functional form assumption. First, subsequent choices, either to maintain an initial plan choice or to switch plans, refine the bounds on prior beliefs. Someone who chose plan j, for whom the initial choice implies an upper bound $\tilde{\mu}_{i1}^{pk}/(1-\varphi) \leq bound_1$ (for example), who does not upgrade to a larger plan after an overage, must in fact have had a strictly lower initial point estimate $\tilde{\mu}_{i1}^{pk}/(1-\varphi) < bound_1$. (An overage is a signal to upgrade, and if one does not upgrade plans following a signal to upgrade, then one must have had a strict preference for the smaller plan prior to the additional information.) Second, as prices change over time, the bounds depicted in Figure 6 change as well, so that plan share data from later dates provide additional restrictions on $\tilde{\sigma}_{\theta 1}$, $\tilde{\mu}_0^{pk}/(1-\varphi)$, and $\tilde{\sigma}_{\mu^{pk}}/(1-\varphi)$. Our structural model point identifies $\tilde{\sigma}_{\theta 1}$ as the value which is best able to fit all of this choice data (rather than just the October-November 2002 choice data).

The exercise described above identifies consumer uncertainty about initial tastes ($\tilde{\sigma}_{\theta 1}$) but it still remains to separate out uncertainty about own type ($\tilde{\sigma}_1$) from perceived taste volatility ($\tilde{\sigma}_{\varepsilon}^{pk}$) which in turn will distinguish overconfidence (δ_μ) from projection bias (δ_ε). By equation (10), $\tilde{\sigma}_{\theta 1}^2$ is a weighted sum of $\tilde{\sigma}_1^2$ and $(\tilde{\sigma}_{\varepsilon}^{pk})^2$. The two parameters are distinguished by the rate of learning and plan switching, which is decreasing in $\tilde{\sigma}_{\varepsilon}^{pk}/\tilde{\sigma}_1$. This is apparent by re-writing equation (8) to show that a consumer's updated beliefs are a weighted average of her prior and her signals, where the weight placed on her prior is proportional to $(\tilde{\sigma}_{\varepsilon}^{pk}/\tilde{\sigma}_1)^2$:

$$\tilde{\mu}_{i,t+1}^{pk} = \frac{(\tilde{\sigma}_{\varepsilon}^{pk}/\tilde{\sigma}_1)^2 \tilde{\mu}_{i1}^{pk} \tilde{\sigma}_1^{-2} + \left(\frac{2\varphi}{1-\varphi} z_{i1} + t\bar{z}_{it}\right)}{(\tilde{\sigma}_{\varepsilon}^{pk}/\tilde{\sigma}_1)^2 + \left(\frac{2\varphi}{1-\varphi} + t\right)}. \tag{17}$$

Note that we can back out signals z_{it} from observed usage, which helps us to see how much they

[36] This is because higher uncertainty (higher $\tilde{\sigma}_{\theta 1}$) leads individuals who choose plans 1-3 to insure themselves by choosing plans with more included minutes. They are willing to choose plan 2 over plan 1 and plan 3 over plan 2 at lower values of $\tilde{\mu}_{i1}^{pk}/(1-\varphi)$. However, they are only willing to choose plan 1 over plan 0 at higher values of $\tilde{\mu}_{i1}^{pk}/(1-\varphi)$.

impact switching and beliefs, and therefore to utilize equation (17) to separate $\tilde{\sigma}_1$ and $\tilde{\sigma}_{\varepsilon}^{pk}$. Recall from Section 4.2 that we identify the probability of an active choice based on the rate at which consumers fail to switch away from dominated plans. Thus we can distinguish slow learning from a failure to actively consider switching.

Contract Price Coefficient The preceding discussion ignores logit-errors, which the model does incorporate into plan choice. As a result, plan choices do not actually give sharp bounds on prior beliefs, but rather smooth likelihoods over priors, since beliefs outside the bounds described by Figure 6 can be explained by the logit error. Without logit-errors, all initial plan choices could be rationalized by prior beliefs. However the model requires logit-errors to rationalize switches that appear to be in the 'wrong' direction. For example, suppose a customer with high average usage chooses a small plan and subsequently experiences a string of overage charges. A low prior belief ($\tilde{\mu}_{i1}^{pk}$ small) could rationalize the initial choice of a small plan. However, given the assumption of Bayesian learning, no prior can simultaneously rationalize the initial choice and a subsequent switch to an even smaller plan. The degree to which switching is in the wrong direction identifies the contract price-coefficient α, which determines the importance of the logit-error.

4.7.4 Tastes

Having identified beliefs it is straightforward to identify taste process parameters. Given the AR1 coefficient φ, the calling price-coefficient β, and consumer beliefs, we can calculate v_{it}^{k} for $k \in \{$pk-in,pk-out,op-in,op-out$\}$ and infer taste-shocks $\boldsymbol{\theta}_{it}$ and $\mathbf{r}_{it}^{9pm}$ from usage. Observing r_{it}^{k} for $k \in \{$9pk,9op$\}$ (a censoring of $\tilde{r}_{it}^{k} = \alpha_i^k + e_{it}^{r,k}$) identifies $E\left[\alpha_i^k\right]$, $Var(\alpha_i^k)$, and $Var(e_{it}^{r,k})$.[37] Correlation between observed usage and initial plan choices identifies ψ, which determines the correlation between beliefs and true types. Given φ and $\boldsymbol{\theta}_{it}$, we can calculate the composite error $(\boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_{it}) = \boldsymbol{\theta}_{it} - \varphi\boldsymbol{\theta}_{i,t-1}$, which is joint-normally distributed conditional on $\tilde{\mu}_{i1}^{pk}$, so unconditionally is the mixture of joint normals. The argument for identifying this distribution is then similar to that for identifying the error structure in a random effects distribution. This delivers the parameters $\boldsymbol{\mu}_0$, $\boldsymbol{\Sigma}_{\mu}$, and $\boldsymbol{\Sigma}_{\varepsilon}$. Finally, bias measures δ_{μ}, δ_{ε}, b_1, b_2, and b_3 can be computed from their definitions.

[37]Without censoring, these would simply be $E\left[\alpha_i^k\right] = E\left[r_{it}^k\right]$, $Var\left(\alpha_i^k\right) = Cov(r_{it}^k, r_{it-1}^k)$, and $Var(e_{it}^{r,k}) = Var(r_{it}^k) - Var(\alpha_i^k)$.

5 Estimation Procedure

Before describing our estimation procedure, we outline the model parameters to be estimated. First are parameters associated with beliefs: the parameters governing the distribution of consumer beliefs, $\tilde{\mu}_0^{pk}$ and $\tilde{\sigma}_{\mu^{pk}}$, consumers' initial uncertainty about their peak type $\tilde{\sigma}_1$, and the projection bias δ_ε. The parameters associated with actual tastes for usage are the means of the μ_{it}^k's, μ_0^{pk}, μ_0^{op}, their variances and correlation $\sigma^2_{\mu^{pk}}$, $\sigma^2_{\mu^{op}}$, ρ_μ, and the variances and correlation of the idiosyncratic errors $(\sigma_\varepsilon^{pk})^2$, $(\sigma_\varepsilon^{op})^2$, ρ_ε, as well as ψ^{pk} and ψ^{op} which capture correlation between beliefs and actual usage. There are four parameters which govern the shares of outgoing landline calls occurring between 8:00 pm and 10:00 pm: the average peak share μ_α^{9pk}, the individual specific variance $(\sigma_\alpha^{9pk})^2$, and the two idiosyncratic variances $(\sigma_e^{9k})^2$ for $k \in \{pk, op\}$. Recall that we do not need to estimate a mean or individual specific variance for off peak 9:00 pm to 10:00 pm usage since we have restricted average peak and off peak tastes for 8:00 pm to 10:00 pm usage to be equal in equation (7). The final set of parameters that are discussed in the text include the calling price-coefficient β, the contract price-coefficient α, the probability a consumer makes an active choice P_C, and the utility of not using a cellphone anymore, O. Finally, we estimate an additional six parameters that govern the share of in-network usage and a parameter that reflects consumer beliefs about the share of in-network usage. We discuss these parameters further in Appendix D. We denote the vector of all the model parameters as Θ, which is 30 dimensional.

We begin this section by describing the structure of the likelihood function which arises from our model. As we will show below, the likelihood function for our model does not have a closed form expression due to the presence of unobserved heterogeneity. We therefore turn to Simulated Maximum Likelihood (SML) (see Gourieroux and Monfort (1993)) to approximate the likelihood function. We then conclude the section by describing some of the computational difficulties that arise during the estimation.

An observation in our model is a usage plan-choice pair for a consumer at a given date. At each observation, we must evaluate the joint likelihood of observed usage and plan choice conditional on observed prices and the consumer's usage and choice history. The likelihood for an observation will arise naturally from our model due to the distributional assumptions we have put on the model's unobservables. To facilitate the exposition, we are going to divide the unobservables into two groups. The first group consists of random variables that are independent across individuals, but are not independent across time within an individual. This consists of the unobservables $\tilde{\mu}_{i1}^{pk}$, μ_i^{pk}, μ_i^{op}, α_i^{9pk}, two normally distributed individual specific effects which govern the share of in network usage for peak and off peak, α_i^{pk} and α_i^{op}, and latent $\tilde{\theta}_{it}^k$ when $\theta_{it}^k = 0$ for $k \in \{pk, op\}$. As we will elaborate below, $\tilde{\theta}_{it}^k$ is unobserved when observed usage is zero for category k, which happens when

$\theta_{it}^k = 0$. We group these random variables together into a vector denoted $\boldsymbol{u}_i$. The other group of error terms consist of structural shocks that are independent across time and individuals: the logit plan choice error η_{itj}, the errors in the stochastic process of $\tilde{\theta}_{it}^k$ when $\tilde{\theta}_{it}^k > 0$, ε_{it}^k, idiosyncratic errors for the 8:00 pm to 10:00 pm shares, r_{it}^{9k}, e_{it}^k, for $k \in \{9pk, 9op\}$, as well as two normally distributed idiosyncratic errors governing in network usage, which we denote e_{it}^{pk} and e_{it}^{op}, respectively.

For individual i at time period t, we observe a plan choice j as well as a vector of usage, $\boldsymbol{q}_{it}$, where $\boldsymbol{q}_{it} = \{q_{it}^{pk,in}, q_{it}^{pk,out}, q_{it}^{op,in}, q_{it}^{op,out}, q_{it}^{9pk}, q_{it}^{9op}\}$, and the *in* and *out* superscripts refer to in and out of network usage. We begin by describing the joint likelihood of usage and plan choice in period 1, because that is simplest, and then describe the likelihood for $t > 1$. Conditional on making an active choice, an individual will choose plan j when that plan has the highest utility according to equation (2). We will denote the set of η_{itj}'s where plan j's utility is highest as $\boldsymbol{S}_{j,t}^{\eta}(\boldsymbol{u}_i, \boldsymbol{q}_{i1}, ..., \boldsymbol{q}_{i,t-1}, \Theta)$. Because we assume that the η_{itj}'s are independent type 1 Extreme value errors, the probability that plan j is chosen in period 1 can be written as

$$P_{i1}(\text{choose } j|\boldsymbol{u}_i) = \int_{\boldsymbol{S}_{j,t}^{\eta}(\boldsymbol{u}_i)} f_{\boldsymbol{\eta}}(\boldsymbol{\eta}_{i1}) d\boldsymbol{\eta}_{i1} = \frac{\exp(U_{ij1}(\Im_{i1}, \boldsymbol{u}_i))}{\sum_{k=1,..,J} \exp(U_{ik1}(\Im_{i1}, \boldsymbol{u}_i))}, \tag{18}$$

where $f_{\boldsymbol{\eta}}(\boldsymbol{\eta}_{it})$ is the density of $\boldsymbol{\eta}_{it} = \{\eta_{it1}, ..., \eta_{it,J_{it}}\}$. $\Im_{i1}$ is the consumer's information set at time 1, which will contain elements of $\boldsymbol{u}$ that are known to the consumer, as well as plan characteristics.

A consumer's observed usage, $\boldsymbol{q}_{it}$, will be a function of $\boldsymbol{u}_i$, the idiosyncratic errors ε_{it}^k and e_{it}^k, and past values of $\boldsymbol{q}_{it}$ for $t > 1$. Conditional on $\boldsymbol{u}_i$ and $\boldsymbol{q}_{i1}, ..., \boldsymbol{q}_{i,t-1}$, the distributions of ε_{it}^k and e_{it}^k will generate a distribution for $\boldsymbol{q}_{it}$. We denote this density function as $f_{\boldsymbol{q}}(\boldsymbol{q}_{it}|\boldsymbol{u}_i, \boldsymbol{q}_{i1}, ..., \boldsymbol{q}_{i,t-1}, \Theta)$. We describe the exact form of $f_{\boldsymbol{q}}$ in Appendix E; given distributional assumptions about ε_{it}^k and e_{it}^k, we derive the distribution of $\boldsymbol{q}_{it}$ using a change of variables. We note that one complication arises when we observe usage of zero in one or both categories. In our model this occurs when $\tilde{\theta}_{i1}^k < 0$. In this case, $\tilde{\theta}_{i1}^k < 0$ is unobserved and must be integrated out. This is why we stipulated that latent $\tilde{\theta}_{i1}^k$'s were included in the vector of $\boldsymbol{u}_i$'s. If usage is zero in a category, the functional form of $f_{\boldsymbol{q}}(\boldsymbol{q}_{it}|\boldsymbol{u}_i)$ also changes. If usage is zero in both categories, then $f_{\boldsymbol{q}}(\boldsymbol{q}_{it}|\boldsymbol{u}_i)$ becomes an indicator for both $\tilde{\theta}_{i1}^{pk} < 0$ and $\tilde{\theta}_{i1}^{op} < 0$. If only peak usage is zero, then $f_{\boldsymbol{q}}$ becomes an indicator for $\tilde{\theta}_{i1}^{pk} < 0$ multiplied by the density of q_{it}^{op} conditional on $\tilde{\theta}_{i1}^{pk}$, and vice-versa if off peak usage is zero.

We can now derive the likelihood of period 1 choice and usage. Marginalizing out over the distribution of $\boldsymbol{u}_i$, this likelihood can be written as

$$L_{i1}(\Theta) = Pr(\text{choose j}|t = 1) f_{\boldsymbol{q}}(\boldsymbol{q}_{i1}|\text{choose j}). \tag{19}$$

Using our definitions above and the fact that $\boldsymbol{\eta}_{it}$ is independent of $\boldsymbol{q}_{it}$, we can write the second term of (19) as

$$f_{\boldsymbol{q}}(\boldsymbol{q}_{i1}|\text{choose j}) = \int_{\boldsymbol{u}_i}\int_{\boldsymbol{\eta}_{i1}} \frac{f_{\boldsymbol{q}}(\boldsymbol{q}_{it}|\boldsymbol{u}_i)f_{\boldsymbol{u}}(\boldsymbol{u}_i)\mathbf{1}\{\boldsymbol{\eta}_{i1}\in \boldsymbol{S}^{\eta}_{j,t}(\boldsymbol{u}_i)\}f_{\boldsymbol{\eta}}(\boldsymbol{\eta}_{i1})\}}{Pr(\text{choose j}|t=1)} d\boldsymbol{\eta}_{i1}d\boldsymbol{u}_i.$$

Rearranging the above equation slightly and substituting it into (19) gives us a straightforward formula for L_{i1}:

$$L_{i1}(\Theta) = \int_{\boldsymbol{u}_i}\left(\int_{\boldsymbol{S}^{\eta}_{j,t}(\boldsymbol{u}_i)} f_{\boldsymbol{\eta}}(\boldsymbol{\eta}_{i1})d\boldsymbol{\eta}_{i1}\right) f_{\boldsymbol{q}}(\boldsymbol{q}_{it}|\boldsymbol{u}_i)f_{\boldsymbol{u}}(\boldsymbol{u}_i)d\boldsymbol{u}_i. \tag{20}$$

Note that the integral over the $\boldsymbol{\eta}_{i1}$ which is in large brackets is the probability that plan j is chosen in period 1, as derived in equation (18).

Writing the likelihood of a sequence of observed usage and choice decisions is straightforward and only requires a little more notation. Conditional on a vector of random effects, the choice probability in period t depends on previous usage, but also on previous choices due to the consideration probability. Recall that in each period, the consumer looks at prices and makes an active plan choice with probability P_C. Conditional on making an active choice[38] in period t and information $\Im_{it}$, the probability of a customer choosing plan $j \in J_{it}$ is

$$P_{it}(j|C;\Im_{it},\boldsymbol{u}_i,J_{it}) = \frac{\exp(U_{ijt}(\Im_{it},\boldsymbol{u}_i))}{\sum_{k\in J_{it}}\exp(U_{ikt}(\Im_{it},\boldsymbol{u}_i))}.$$

The consumer's information set in period t will contain some of the random draws, as well as past $\boldsymbol{q}_{it}$'s which impact the Bayesian updating process. The consumer's choice set J_{it} depends on the plan choices drawn from the non-university plans, and the consumer's past plan choices. For a new customer, the initial choice set J_{i1} includes plans currently offered through the university but does not include the outside option or any other plans, and does not vary with the simulation draw. Other options are not included for new customers because we only observe consumers who sign up; hence the probability of plan choice for these customers is the probability of choosing plan j conditional on signing up. For existing customers, the choice set J_{it} also includes the customer's existing plan and those currently offered by the other provider considered. We assume that the consumer considers only one outside provider (AT&T, Cingular, or Verizon), in addition to the possibility of quitting each month. The option considered is drawn from a discrete distribution which assigns probability 1/3 to each of the three providers.Since there are four possible choice sets, we index each choice set by J^k_{it}, $k = 1, ..., 4$.

The probability that an existing customer switches to plan j' (where j' could imply stopping

[38] Notation: conditioning on C means conditioning on an active choice.

use of a cellular phone) in period t or keeps the existing plan j are $P_C P_{it}(j'|C; \Im_{it}, \boldsymbol{u}_i, J_{it})$ and $P_C P_{it}(j|C; \Im_{it}, \boldsymbol{u}_i, J_{it}) + (1 - P_C)$ respectively:

$$P(\text{Choose } j'|\Im_{it}, \boldsymbol{u}_i, J_{it}) = \begin{cases} P_C P_{it}(j'|C; \Im_{it}, \boldsymbol{u}_i, J_{it}) & \text{if } j \neq j' \\ P_C P_{it}(j'|C; \Im_{it}, \boldsymbol{u}_i, J_{it}) + (1 - P_C) & \text{if } j = j' \end{cases}. \tag{21}$$

Turning to $\boldsymbol{q}_{it}$, similar to period 1, the density $f_{\boldsymbol{q}}(\boldsymbol{q}_{it}|\boldsymbol{u}_i, \boldsymbol{q}_{i1}, ..., \boldsymbol{q}_{i,t-1})$ can be derived from the density of the error terms for $\tilde{\theta}^k_{it}$ and r^k_{it}. Note that $\boldsymbol{q}_{it}$ is a function of previous $\boldsymbol{q}_{it}$'s due to its dependence on $\tilde{\theta}^k_{it}$, which is a function of $\tilde{\theta}^k_{i,t-1}$ through the AR1 process, as well as its dependence on v^*, which is a function of all previous $\boldsymbol{q}_{it}$'s through the Bayesian updating process. Again, if $\boldsymbol{q}_{it}$ is zero in some category, we include $\tilde{\theta}^k_{it}$ as a random effect and $f_{\boldsymbol{q}}(\boldsymbol{q}_{it}|\boldsymbol{u}_i, \boldsymbol{q}_{i1}, ..., \boldsymbol{q}_{i,t-1})$ will include an indicator for $\tilde{\theta}^k_{it} < 0$. Full derivations of $f_{\boldsymbol{q}}$ are left to Appendix E.

We can now write out the likelihood of a sequence of choice and usage decisions:

$$L_i(\Theta) = \int_{\boldsymbol{u}_i} \prod_{t=1}^{T_i} \left[\left(\sum_{k=1}^{3} \frac{1}{3} P(\text{Choose } j'|\Im_{it}, \boldsymbol{u}_i, J^k_{it}) \right) f_{\boldsymbol{q}}(\boldsymbol{q}_{it}|\boldsymbol{u}_i, \boldsymbol{q}_{i1}, ..., \boldsymbol{q}_{i,t-1}) \right] f_{\boldsymbol{u}}(\boldsymbol{u}_i) d\boldsymbol{u}_i. \tag{22}$$

As we noted above, the individual specific likelihood has no closed form solution due to integration over $\boldsymbol{u}_i$. We approximate this integral using Monte Carlo Simulation. For each individual, we take S draws on the random effects from $f_{\boldsymbol{u}}(\boldsymbol{u}_i)$ and the choice sets J^k_{it}, and approximate the likelihood using

$$\hat{L}_i(\Theta) = \frac{1}{S} \sum_{s=1}^{S} \left[\left(P(\text{Choose } j'|\Im_{it,s}, \boldsymbol{u}_{is}, J^k_{it,s}) \right) f_{\boldsymbol{q}}(\boldsymbol{q}_{it}|\boldsymbol{u}_{is}, \boldsymbol{q}_{i1}, ..., \boldsymbol{q}_{i,t-1}) \right].$$

The model log-likelihood is the sum of the logarithms of the individual likelihoods:

$$\hat{LL}(\Theta) = \sum_{i=1}^{I} \log(\hat{L}_i(\Theta)). \tag{23}$$

It is well-known that the value of Θ which maximizes $\hat{LL}$ is inconsistent for fixed S due to the logarithmic transformation in equation (23). However, it is consistent if $S \to \infty$ as $I \to \infty$, as discussed in Hajivassiliou and Ruud (1994). We chose $S = 300$; to arrive at this value we conducted some simple artificial data experiments where we simulated our model and attempted to recover the parameters, finding that 300 draws was sufficient to recover the true parameter draws to roughly 5% accuracy. We also found in our experiments that we were able to reduce simulation bias significantly by using a deterministic Sobol sequence generator to create the random draws, rather than canonical random number generators. Goettler and Shachar (2001) describe some of the advantages of this technique in detail. We use the algorithm provided in the R package randtoolbox to create the draws (Dutang and Savicky 2010).

A second issue that arises in the formulation of equation (23) is that with a finite number of draws, the inclusion of an indicator function in $f_{\boldsymbol{q}}$ when usage is zero produces a discontinuous likelihood, which is difficult to maximize. If only one period were observed for each consumer, then we could solve this problem by simply substituting the probability that $\tilde{\theta}_{it}^{k}$ is censored for $f_{\boldsymbol{q}}$ and not including $\tilde{\theta}_{it}^{k}$ in $\boldsymbol{u}_i$, rather than using an indicator function and integrating out $\tilde{\theta}_{it}^{k}$. However, because we observe multiple periods per individual and we have assumed that $\tilde{\theta}$ follows an AR1 process, if q_{it}^{k} is zero in period t we need a value of $\tilde{\theta}_{it}^{k}$ to compute the likelihood of period $t+1$'s $\boldsymbol{q}$. Hence, we need to have a draw of $\tilde{\theta}_{it}^{k}$ available when censoring occurs. This type of problem also arises in the estimation of dynamic Tobit models. In that literature, Lee (1999) proposes integrating out serially correlated latent unobservables using an importance-sampling procedure that results in a smooth likelihood function. We adapt that procedure to our problem. In Appendix E, we derive the densities used to draw out the latent $\tilde{\theta}_{it}^{k}$, and we describe the changes to the likelihood that make it smooth.

We wrote the program to evaluate the likelihood in R and Fortran. The evaluation of this likelihood is computationally intensive for two reasons: first, it must be evaluated at many simulation draws; second, for each choice a consumer could make, at each time period and each draw, we often must solve for $\mathbf{v}_{it}^{*}$ and $\alpha_i^{9,op}$ using a nonlinear equation solver. Our estimation method therefore falls into an inner-loop outer-loop framework, where the inner loop is the solution of the $\mathbf{v}_{it}^{*}$'s and $\alpha_i^{9,op}$'s, and the outer loop maximizes the likelihood.

We summarize the algorithm for computing these variables in four steps. Step 1 is to compute $\alpha_{i,s}^{9op}$ conditional on the simulated draws and the other model parameters. Recall that we assume that a consumer's average taste for weekday-evening landline-usage is the same thirty minutes before and after 9pm. For each consumer i and each simulation draw s, we compute $\alpha_{i,s}^{9op}$ as the solution to equation (29) in Appendix D, which extends equation (7) to account for in-network calling. As this equation does not have an analytic solution, we compute $\alpha_{i,s}^{9op}$ with a nonlinear equation solver. The result of this step is used to compute the structural error for r_{it}^{9op}.

The next three steps compute the calling threshold vector $\mathbf{v}_{it,s}^{*}$ and $\tilde{\boldsymbol{\theta}}_{it,s}$ period-by-period. Because the $\mathbf{v}_{it,s}^{*}$ is a function of past values of $\tilde{\boldsymbol{\theta}}_{it,s}$ through the Bayesian learning and the AR1 process, these three steps are iterated across both individuals i, and time periods t. Step 2 calculates consumer beliefs about $\tilde{\boldsymbol{\theta}}_{it,s}$ in two parts following Section 4.4. First, consumer beliefs about μ_i^{pk}, $(\tilde{\mu}_{it,s}^{pk}, \tilde{\boldsymbol{\sigma}}_{it}^{-2})$ are updated via Bayes rule by conditioning on the lagged value $\tilde{\theta}_{i,t-1,s}^{pk}$. Second, beliefs about $\tilde{\boldsymbol{\theta}}_{it,s}$ are computed from $(\tilde{\mu}^{pk}_{it,s}, \tilde{\boldsymbol{\sigma}}^{-2}_{it})$, $\mu^{op}_{i,s}$ and the the lagged value $\tilde{\boldsymbol{\theta}}_{i,t-1,s}$ which enters through the AR1 process. (No updating is required for $t = 1$.) In step 3 we calculate $\mathbf{v}^*_{it,s}$ following it's characterization in Appendix D, which depends on the beliefs calculated in step 2. Recall that components of $\mathbf{v}^*_{it}$ are either known to be 0 cents or 11 cents or must be calculated by numerically solving a first-order condition (either equation (27) or (28) which are the extensions to equation (5) that account for in-network calling given in Appendix D). In step 4, we calculate $\tilde{\boldsymbol{\theta}}_{it,s}$. When $\boldsymbol{\theta}_{it,s}$ is not censored, we can compute $\tilde{\boldsymbol{\theta}}_{it,s}$ from observed usage conditional on β and $\mathbf{v}^*_{it,s}$ using equations (30)-(31) in Appendix D. When censoring occurs, we use the simulated value for $\tilde{\boldsymbol{\theta}}_{it,s}$.

With $\alpha^{9op}_{i,s}$, $\tilde{\boldsymbol{\theta}}_{it,s}$ and $\mathbf{v}^*_{it,s}$ in hand we can compute the choice probabilities and the density of observed usage in equation (22). We optimize our likelihood in two steps. The first step uses a Nelder-Mead optimizer to get close to the optimum. From there we use a Newton-Raphson optimizer to reach the optimum within a tighter tolerance. Because the optimization algorithms will stop at local optima, it is important to have good starting points. To arrive at starting points for the model, we choose the usage parameters (the means and variances of the μ's, α's, and ε's) and the β to match observed usage.[39] Conditional on these choices of usage parameters, we choose initial belief parameters to match the observed plan shares. To do this, we use our model to simulate plan shares for the 2002 to 2003 school year and the 2003 to 2004 school year, and match those simulated shares to the observed shares during these two years. We chose to split the data in that way to exploit the fact that plan 0 stopped offering free off peak minutes at the beginning of the 2003 to 2004 school year.

6 Results

Our parameter estimates are shown in Table 7. The first three columns show the coefficients, estimates, and standard errors for the first 15 parameters, while the fourth through sixth columns show the same for the next 15 parameters. The calling price coefficient β is 3.41, which indicates that a price increase from 0 cents to 11 cents per minute decreases usage by about 27%. The next two parameters relate to overconfidence and projection bias. The standard deviation of consumer uncertainty about mean type, $\tilde{\sigma}_1$, is 13.6 minutes. In contrast, further down the column the variance of μ^{pk}_i conditional on $\tilde{\mu}^{pk}_{i1}$ is much higher at $\sigma^{pk}_m = 77$ minutes. Thus consumers are overconfident,

[39]We assume that v^* is equal to 3 cents for plan 3, 5 cents for plan 2, and 8 cents for plan 1, and maximize the likelihood of usage conditional on those guesses at v^*. We chose those values of v^* because they matched the average values of v^* that were produced by simulating the model at parameters which were in the neighborhood of the estimates. We stress that we only use the guesses at v^* to arrive at starting points; we solve for the endogenous v^*'s when running the full simulated maximum likelihood.

underestimating uncertainty about mean type by about 82%. Moreover, the estimate of δ_ε shows that consumers exhibit projection bias, believing the variance of ε_{it}^{pk} to be 54% of its true value.

Table 7: Parameter Estimates

Coefficient	Estimate	Std. Err	Coefficient	Estimate	Std. Err
β	3.41	(0.047)	μ_α^{9pk}	-0.004	(0.001)
$\tilde{\sigma}_1$	13.553	(0.973)	$(\sigma_\alpha^{9pk})^2$	0.059	(0.001)
δ_ϵ	0.537	(0.007)	$(\sigma_e^{9pk})^2$	0.105	(0.001)
$\tilde{\mu}_0^{pk}$	-24.966	(6.698)	$(\sigma_e^{9op})^2$	0.116	(0.001)
μ_0^{pk}	107.469	(1.792)	φ	0.518	(0.019)
μ_0^{op}	109.64	(3.354)	α	0.309	(0.135)
$\tilde{\sigma}_\mu^{pk}$	111.794	(0.025)	Price Consideration	0.061	(0.046)
$\sigma_{\mu^{pk}}$	77.038	(1.304)	Outside Good Utility	-65.773	(28.631)
$\sigma_{\mu^{op}}$	171.326	(2.096)	δ_r	0.003	(0.047)
ψ^{pk}	-0.212	(0.019)	μ_α^{pk}	0.349	(0.002)
ψ^{op}	0.188	(0.022)	μ_α^{op}	0.402	(0.002)
ρ_μ	0.981	(0.002)	$(\sigma_\alpha^{pk})^2$	0.034	(0.001)
σ_ε^{pk}	169.803	(0.539)	$(\sigma_\alpha^{op})^2$	0.04	(0.001)
ρ_ε	0.397	(0.004)	$(\sigma_e^{pk})^2$	0.03	(0)
σ_ε^{op}	305.534	(0.667)	$(\sigma_e^{op})^2$	0.026	(0)
Log-likelihood	-265155.8				

The next 9 parameters characterize the distribution of $\tilde{\mu}_{i1}^{pk}$ and the distribution of μ_i^{pk} and μ_i^{op} conditional on $\tilde{\mu}_{i1}^{pk}$. On average, consumers believe their initial draw on $\tilde{\theta}_{it}^{pk}$ to be negative, while the actual mean of $\tilde{\theta}_{it}^{pk}$ is 107 minutes. (Note that even though the average user believes her initial $\tilde{\theta}_{it}^{pk}$ is negative on average, she believes that her initial θ_{it}^{pk} to be positive on average because θ_{it}^{pk} is censored at zero.) The average off-peak draw $\tilde{\theta}_{it}^{op}$ is a little bit higher than the peak value at 109 minutes. The standard deviation in consumers' initial belief $\tilde{\mu}_{i1}^{pk}$ is 112 minutes. Conditional on $\tilde{\mu}_{i1}^{pk}$, the standard deviations of μ_i^{pk} and μ_i^{op} are 77 and 171 minutes respectively. The estimates of ψ indicate that initial beliefs are negatively correlated with μ_i^{pk}, and positively correlated with μ_i^{op}. Finally, conditional on $\tilde{\mu}_{i1}^{pk}$ the correlation between peak and off-peak μ_i^k is high at 98%. The unconditional correlation between peak and off-peak μ_i^k is somewhat lower at 89%; the unconditional standard deviations of the peak and off-peak μ_i^k are slightly higher than their conditional values at 80 minutes and 172 minutes respectively. The last three rows of column 1 describe the distribution of the error term $\boldsymbol{\varepsilon}$. The variances of peak and off peak errors are higher than the unconditional variances of μ_i^{pk} and μ_i^{op}, indicating that more of the variation in usage can be attributed to monthly volatility than the consumer-level fixed effect; additionally, their correlation is much lower.

The first four parameters of column 2 describe the consumer's taste for 8:00 pm to 10:00 pm usage. The low value of μ_α^{9pk} indicates that outgoing 8:00 to 9:00 pm landline usage is small as a fraction of total peak usage, which is consistent with the data. The φ value of 0.52 indicates a significant amount of serial correlation in tastes from month to month. The contract price coefficient, α, is estimated to be 0.31. The price consideration parameter is 0.061, indicating that consumers seldom look at prices. However, the parameter is not precisely estimated; this is consistent with our artificial data experiments, where we found that this parameter was difficult to identify. The outside good utility is estimated to be -66. Compared to average utilities of about -16, this implies that consumers prefer the inside goods to the outside good by a large margin.

The last seven parameters relate to in-network usage. We describe the modifications to the model needed to distinguish in and out-of-network usage in Appendix D. Loosely, the parameter δ_r is the consumer's belief about what fraction of her usage is in-network. Since our estimate of this parameter is close to zero, consumers believe that almost all usage is out of network.[40] The next two parameters govern the shares of θ_{it} which can be apportioned to peak and off-peak in-network usage, respectively, while the final four govern the variances of in-network usage.

Table 8: Estimates of Consumer Beliefs

Coefficient	Estimate	Std. Err
δ_μ	0.176	(0.013)
δ_ε	0.537	(0.007)
$\tilde{\sigma}_1$	13.553	(0.973)
$\tilde{\sigma}_\varepsilon$	91.166	(1.077)
$\tilde{\sigma}_{\theta 1}$	110.212	(1.304)
b_1	-132.435	(6.976)
b_2	1.295	(0.019)
b_3	-0.441	(0.005)

Turning back to consumer beliefs, some of the parameters which characterize consumer beliefs, such as δ_μ and b_1, are functions of our estimated parameters. We display estimates of these parameters in Table 8. Our estimate of δ_μ indicates strong overconfidence: consumers underestimate the true standard deviation of μ_i^{pk} by about 82% believing it to be 13 minutes (compare to a true con-

[40]Plan 0 always offered free in-network usage and plan 2 did so as well near the end of our sample period. We incorporated this parameter to help explain the high share of Plan 1 relative to Plan 0, as plan 0 dominates plan 1 for anyone with a median in-network usage share.

ditional standard deviation of 77 minutes). The impact of the overconfidence on consumer beliefs can be seen in Figure 7. The solid black line shows a consumer's perceived distribution of μ_i^{pk} when $\tilde{\mu}_{i1}^{pk} = \tilde{\mu}_0^{pk}$ and $\mu_i^{op} = \mu_0^{op}$, while the dotted red line shows the true distribution of μ_i^{pk} conditional on $\tilde{\mu}_{i1}^{pk} = \tilde{\mu}_0^{pk}$ and $\mu_i^{op} = \mu_0^{op}$. The mean of the perceived distribution is lower than the true distribution as a result of the aggregate mean bias, and the variance of the perceived distribution is considerably lower than the variance of the true distribution due to the overconfidence. Our estimate of δ_ε indicates strong (although milder) projection bias: consumers underestimate the standard deviation of the monthly innovation in tastes by about 54%. Consumers believe the standard deviation of the peak error is 91 minutes (compared to a true standard deviation of 170 minutes). Aggregate mean bias is negative, indicating that the average consumer underestimates her initial $\tilde{\theta}_{it}$ draw by 132 minutes. Accounting for censoring of the latent shock, the average consumer believes the mean of θ_{it} is 32 minutes. Consumers initial uncertainty about $\tilde{\theta}_{it}$, $\tilde{\sigma}_{\theta 1}$, is 110 minutes. In contrast, an average unbiased consumer would believe the mean of θ_{it} is 140 minutes and her initial uncertainty about $\tilde{\theta}_{it}$, $\sigma_{\theta 1}$, would be 250 minutes. Finally, the positive estimate of b_2 reflects strong positive conditional mean bias. (See footnote 33 for our interpretation of the negative estimate of b_3.)

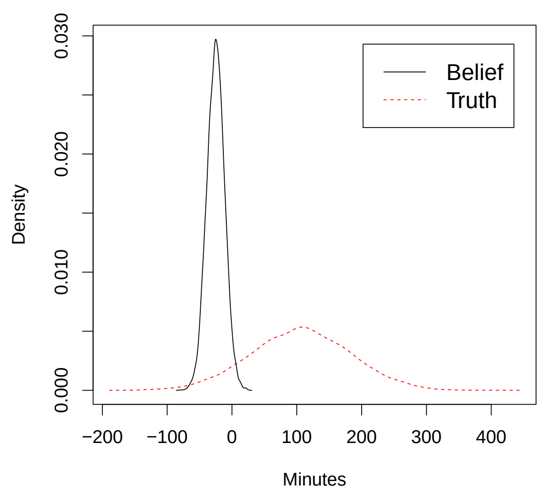


Figure 7: Perceived and true distributions of μ_i^{pk} conditional on $\tilde{\mu}_{i1}^{pk} = \tilde{\mu}_0^{pk}$ and $\mu_i^{op} = \mu_0^{op}$

The fact that overconfidence is stronger than projection bias ($\delta_\mu < \delta_\varepsilon$) implies that consumers overweight their priors relative to new experience and hence learn slowly. This is illustrated in Figure 8, which plots an average of consumers' evolving point-estimates $\tilde{\mu}_{it}^{pk}$ for consumers whose true value is $\mu_i^{pk} = \mu_0^{pk} \approx 107$. A consumer's time t point-estimate $\tilde{\mu}_{it}^{pk}$ is a function of her initial belief $\tilde{\mu}_{i1}^{pk}$ and her signals from past usage z_{it}, which we integrate out using simulation. The dotted lines in the figure show the average of $\tilde{\mu}_{it}^{pk}$ for 1000 simulated consumers, where each consumer's $\tilde{\mu}_{i1}^{pk}$ and z_{it} are drawn from their estimated distributions. The red thick line shows how consumers' beliefs evolve given estimated overconfidence and projection bias. A consumer whose true μ_i^{pk} is roughly 107 minutes and who enters the sample believing her $\tilde{\mu}_{i1} = \tilde{\mu}_0 = -24$ increases her belief to $\tilde{\mu}_{i,13} = 2$ after one year. The blue dashed line shows how beliefs evolve when overconfidence and projection bias are removed. Debiasing consumers speeds up learning and after 1 year a consumer's belief about μ_i^{pk} will be $\tilde{\mu}_{i,13} = 68$ minutes, only 39 minutes below its actual value of 107 minutes.

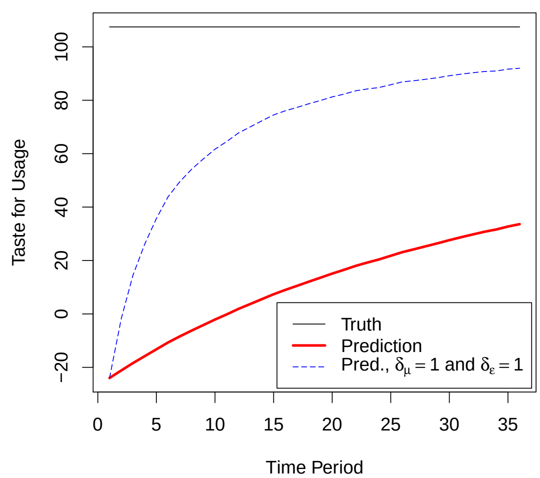


Figure 8: Posterior Estimate of μ_i vs Actual μ_i for $\mu_i = \mu_0$ (Peak calls)

6.1 Fixed-Price Counterfactual: Impact of Biased Beliefs

While firms will naturally alter prices in response to a change in consumer beliefs or regulation, we defer modeling endogenous price changes in our counterfactual simulations until Section 6.3. Here, we begin by considering a set of prices-fixed counterfactual simulations. In Tables 9 and 10 we show the results of fixed-price counterfactual experiments in which we remove one or more consumer biases. We construct these counterfactual simulations at our data in the sense that we hold fixed the number of consumers, and when consumers enter and exit the data set.

Table 9 shows counterfactual plan choice shares. The first three rows show simulated shares of university plans under estimated beliefs, overconfidence and projection bias free beliefs, and unbiased beliefs respectively. Moving down the first three rows, consumers switch away from plan 1 as consumers are progressively debiased, which is to be expected. Biased consumers who choose plan 1 incur more overages than they expect; once bias is removed, consumers will choose alternate

Table 9: Counterfactual: Impact of beliefs on plan choice shares (fixed prices)

		Share of Plan				
Offered Plans	Beliefs	0	1	2	3	4
University	Estimated	42.8	28.7	13.7	5.5	0
University	$\delta_\mu = 1$ and $\delta_\varepsilon = 1$	39.8	25.7	17.8	7.1	0
University	No Biases	46.6	17.5	18.4	8.5	0
Public	Estimated	N/A	23.2	11.7	4.9	0.9
Public	$\delta_\mu = 1$ and $\delta_\varepsilon = 1$	N/A	20.9	13.9	5.6	1.0
Public	No Biases	N/A	15.5	16.6	8.9	1.4

We omitted a column for the share of the outside good and other carriers. Plan 4 refers to a $59.99 plan which offered 650 to 1150 peak minutes. This plan was available to the general public, and a similar plan was available to students but was not chosen by any of them.

plans. Holding prices fixed, consumer welfare rises from removing this bias by about $24 per student over the two year period where they are observed. The last three rows show simulated shares if university plans are removed from the choice set and consumers choose between the university carrier's public plans as well as plans from all other carriers when they enter the data. This change greatly increases the share choosing alternate carriers. Moreover, because the public plans did not include a flat rate plan, 1 is relatively more popular among biased consumers and its fall in share is correspondingly larger when consumers are debiased.

Table 10: Counterfactual: Per student change in surpluses from bias elimination (fixed prices)

	University Plans			Public Plans		
Beliefs	Profits	Cons. Welf.	Total Welf.	Profits	Cons. Welf.	Total Welf.
$\delta_\mu = 1$	−32.56	24.26	−8.29	−71.62	53.36	−18.26
$\delta_\varepsilon = 1$	−22.32	18.69	−3.64	−49.13	39.18	−9.96
$\delta_\mu = 1$ and $\delta_\varepsilon = 1$	−39.25	28.69	−10.56	−86.24	62.35	−23.89
No Biases	−58.40	49.28	−9.12	−98.22	83.35	−14.87

Changes in surpluses (profits, consumer welfare, and total welfare) are measured in dollars per student over the 2 year sample period. Changes are relative to surpluses at estimates.

Table 10 shows the change in firm profits, consumer welfare, and total welfare that results from removing one or more consumer biases. Surplus changes are measured in dollars per student over the two year period that they are observed. The first three columns show the welfare effects when students face university prices, while the last three columns show the welfare effects when consumers fact publicly available prices. The consequences of debiasing are larger in the later case because the flat rate university plan, which is absent from the public menu, tended to protect biased consumers. The first row of Table 10 shows the impact of removing overconfidence, which raises consumer surplus at the expense of lowers profits and lower total welfare. Row two shows similar effects of removing projection bias and row three shows the combined effects of removing both biases. Finally row four shows the total effect of removing all biases, including aggregate mean bias, conditional mean bias, and underestimation of in-network calling (discussed in Appendix D). On average, debiased consumers make fewer calls because they are more aware of overage risk and this reduces total welfare because marginal costs are approximately zero. Thus gains in consumer surplus are overshadowed by profit losses. Interestingly, while consumers naturally benefit most (about $49 per student) when completely debiased, the total welfare loss is largest (about $11 per student) when only overconfidence and projection bias are removed.

6.2 Fixed Price Counterfactual: Impact of Bill-Shock Regulation

In this section we evaluate the welfare impact of a counterfactual experiment where we implement bill-shock regulation similar to recent voluntary agreement between cellular carriers and the FCC. We hold prices fixed at the levels we observe in the university billing data, and the outside plan price data from EconOne. In the next section we will allow firms to vary prices and recompute the equilibrium. In this counterfactual, consumers are informed when their usage reaches Q, their allotment of free minutes.[41] In response to this new policy, a consumer's usage rule changes: A consumer will accept all calls valued above v^* until she exhausts her included minutes. After that point, she only accepts calls valued above p. Because the consumer adjusts her calling threshold upon making Q calls, the optimal initial threshold v^* differs from that characterized by equation (5). (Appendix B describes expected utility and characterizes v^* under bill-shock regulation.) In our counterfactual experiment, we first simulate consumer usage and choices under the standard regime, where consumers are not informed about when they use their free minutes. Then, we solve for new v^*'s and re-simulate choices and usage under the bill-shock regime.

We find that the bill-shock regulation reduces usage. For three-part tariffs, overall usage drops by 13 minutes per bill. This usage drop is primarily driven by consumers cutting back usage after an overage: for consumers who make overages pre-regulation, the bill shock regulation reduces usage by about 66 minutes (consumers who make underages increase their usage by a little more than one minute). The regulation has very little impact on plan choice.[42] Table 11 shows the impact of the counterfactual exercise on profits, consumer welfare, and total welfare. As in the case of debiasing

[41]This counterfactual experiment will not impact the behavior of consumers who stay on the flat rate plan.

[42]The most noticeable increase is in the share of Plan 1, which rises from 28.7% to 28.9% when the university plans are offered, and from 23.2% to 23.3% when the public plans are offered.

Table 11: Bill-Shock Counterfactual (Fixed Prices): Welfare Impact

	University Plans			Public Plans		
Beliefs	Profits	Cons. Welf.	Total Welf.	Profits	Cons. Welf.	Total Welf.
Estimated	−38.83	21.55	−17.29	−78.28	41.87	−36.41
$\delta_\mu = 1$ and $\delta_\varepsilon = 1$	−7.79	3.49	−4.30	−15.90	7.55	−8.35
No Biases	−1.39	1.87	0.48	−5.90	4.89	−1.01

Per student change in surpluses (profits, consumer welfare, and total welfare) due to bill-shock regulation are measured in dollars per student over the 2 year sample period.

shown in Table 10, the consequences of bill-shock regulation are larger when consumers face public prices because they lack the protection of the flat rate university plan. Row one of Table 11 shows the effect of regulation at the estimated parameters. Consumer welfare increases by bit less than \$22 per student, which is less than half the amount that complete debiasing raises consumer welfare. At the same time, the total welfare loss of \$17 per student is much larger than the corresponding loss from debiasing. Rows two and three show that when consumers are partially or fully debiased the regulation has a much smaller effect, increasing consumer welfare by less than two dollars per student when consumers are unbiased. This is to be expected because the primary impact of the bill-shock regulation is to reduce the size of overages. When consumers are debiased they make better plan choices, moving away from plan 1, and raise their calling thresholds v_{it}^{pk}, consuming less and paying less overages. As unbiased consumers do a better job of avoiding overages, Bill-shock regulation has less scope to improve their welfare. Overall, this suggests that the main value of the bill-shock regulation arises from the presence of consumer biases. In the absence of a way to remove biases, bill-shock regulation provides a socially costly way to mitigate their effects.

6.3 Endogenous Price Counterfactual: Impact of Bill-Shock Regulation

6.3.1 Nested Logit Specification

To predict the effect of bill-shock regulation on equilibrium prices it is important to correctly capture the degree of competition between carriers. Hence, we modify the error structure of the demand model to be a two level nested logit, rather than logit. In our nested logit specification, we assume that each inside nest contains the plans offered by a carrier (the option of shutting off cell-phone service is also put in its own nest). The outside nest consists of all the carriers (including no service) in a particular consumer's consideration set.[43] We assume that the inclusive value parameter, denote by λ, is the same for each option.

The more restrictive logit specification implies that if consumers choose plans within carrier primarily based on price then carriers are close substitutes. Thus the logit specification leads to unrealistically high competition and low prices in counter factual simulations. We choose the more flexible nested logit specification because it allows consumers to have strong idiosyncratic carrier preferences (due to network coverage or phone availability) that create market power, while at the same time making within carrier plan choices primarily based on price.

Ideally, we would like to estimate λ jointly with the other demand parameters using demand side choice data. Unfortunately, we observe neither carrier market shares on campus nor the alternate carriers chosen by students quitting university plans. Hence only the quitting rate is available to identify utility of the outside good, average utility of university plans relative to other carriers, and λ. In our demand estimates we assume $\lambda = 1$ (logit specification) and carrier symmetry to identify the outside good utility.[44] To address this identification problem, we calibrate λ using supply-side price data: We select the value of λ that best rationalizes observed prices conditional on our demand estimates. Our algorithm, which is described in Appendix F, calibrates λ to be 0.2.

Before proceeding, we make two comments on our calibration approach. First, one potential problem is that our demand estimates were made conditional on $\lambda = 1$ (which generates the logit model), but different values of λ might produce different demand estimates. Fortunately, our demand estimates are relatively insensitive to λ, which we show in Appendix F. Second, in principle we could have estimated λ and the other parameters jointly by using constrained maximum likelihood and constraining observed prices to be optimal at the estimated parameters. We avoided this approach because we prefer only to impose our supply-side structural assumptions (that competition is symmetric static Nash in prices and that our student population is representative.[45]) only when they are necessary in the endogenous-price counterfactual simulations.

---

[43] A new consumer chooses among three carriers and no-serivce whereas an existing consumer who considers switching chooses among her current carrier, a randomly chosen outside carrier, and no service.

[44] Outside price variation is too limited to separately identify λ. Moroever, in an unreported specification, we instead chose a natural normalization for the outside good and estimated λ. We rejected this alternative, however, because the resulting estimate of λ was zero, an implausible number that implies carriers have monopoly power. This may have been due to the fact that our normalization of the outside good was too low, that carrier symmetry is a bad assumption when including the university plans, or the fact that forced quits due to graduations (outside the model) biased the estimate downwards.

[45] In reality, university plans are not symmetric to other carrier offerings and our population of students is likely overweighted towards new and low volume users relative to the overall population.

6.3.2 Counterfactual Simulation

Table 12 shows the results of our endogenous-price counterfactual experiments. We assume that carriers are symmetric, equilibrium is symmetric static Nash in prices, overage rates are at most fifty cents,[46] and each carrier offers a menu of two plans. Column 1 shows predicted plan prices and welfare outcomes under our estimated demand parameters. The model predicts that the firm offers a two-part tariff at fifty cents per minute for $28.46 per month and a three-part tariff with 295 included minutes for $61.28 per month. Relative to observed prices, these predictions have too few included minutes for the given price points. This is likely due to the fact that our calibration ensured we did a reasonable job of matching price points but our student population comprises relatively low usage customers compared to the total population.

Table 12: The Impact of Bill Shock Regulation and Removing Biases on Equilibrium Prices

		Est (1)	Est, Bill Shock (fixed prices) (2)	Est, Bill Shock (3)	$\delta_\mu = 1$ and $\delta_\varepsilon = 1$ (4)	No Biases (5)
Plan 1	M	28.46	28.46	28.36	29.26	78.07
	Q	0	0	0	0	∞
	p	50	50	50	50	N/A
	Share	63	63	72	64	49
	Margin	76	76	76	76	78
Plan 2	M	61.28	61.28	73.99	78.95	78.07
	Q	295	295	374	∞	∞
	p	50	50	50	N/A	N/A
	Share	37	37	28	36	51
	Margin	77	69	79	79	78
Outside Good Share		0	0	0	0	0
Profit		915	882	919	925	937
Cons Welfare		5497	5515	5465	5501	5695
Total Welfare		6413	6396	6384	6425	6632

All welfare and profit numbers are expressed in thousands of dollars. Because the counterfactuals in columns (4) and (5) produced two part tariffs, under bill shock regulation equilibrium prices are unchanged. We simulate 1000 consumers for 12 months.

Column 2 of Table 12 holds fixed the predicted prices from column 1 but imposes bill-shock regulation. Holding prices fixed, bill-shock regulation transfers money from firms to consumers by helping (plan 2) consumers avoid overage fees. Thus annual profits fall by $33.84 per consumer

[46] Otherwise the combination of biased beliefs and inattention lead to implausibly high overage rate predictions.

while consumer surplus increases by \$17.37 per consumer. At the same time, the reduction in calling leads to an annual total welfare loss of \$16.47 per consumer as marginal cost is assumed to be zero. Column 3 imposes bill-shock regulation but allows firm prices to adjust. In equilibrium, markups are determined primarily by the calibrated inclusive value parameter, which at $\lambda = 0.2$ implies markups of about \$78 per month. As a result, the loss of overage revenue leads firms to raise the fixed fee on plan 2 by \$12.71. Hence annual profits are stable (rising by only \$3.80 per consumer) and consumers lose because they essentially become residual claimants of total welfare. Total welfare falls by \$28.67 per consumer annually, a larger reduction than seen in column 2. One might have expected a smaller welfare loss because firms respond to the inefficiency of consumers reducing calling at the included allowance by increasing included minutes from 295 to 374. However, biased consumers view the net price change (of higher fixed fee, more included minutes, and fewer accidental overages) as a price increase and hence switch away towards plan 1. As plan 1 has no included minutes, the net effect is a further reduction in usage in column 3 compared to column 2 and hence even lower total welfare.

Column 5 of Table 12 shows the effect of eliminating all biases. In this scenario, firms offer unlimited talk plans for \$78. This follows because absent biases the only reason not to set the per minute price equal to marginal cost would be for purposes of price discrimination. However, in equilibrium it is optimal to charge everyone the same \$78 markup because the calibrated inclusive value is assumed constant across individuals. Thus a single unlimited talk plan is sufficient. (In fact two identical plans are better than one because of the red bus/blue bus problem). By returning marginal price to marginal cost, eliminating biases increases total welfare by \$220 per consumer. As noted before, consumers are essentially residual claimants of total welfare, so consumer welfare increases by almost the same amount while profits are stable. Finally, because eliminating bias also eliminates three-part tariff pricing, bill-shock regulation has no effect without bias.

The fact that eliminating biases leads to a single unlimited talk plan points out that the pricing in columns 1 through 3 is driven by biased beliefs. To understand why biases lead to this pricing prediction it is worth starting by considering column 4. Column 4 shows predicted prices when overconfidence and projection bias are removed but consumers still suffer estimated mean biases. The estimated positive conditional mean bias implies that consumer beliefs are too extreme: consumers with low expectations of demand underestimate their demand while consumers with high expectations of demand overestimate their demand.[47] Grubb (2009) shows that firms should raise marginal prices when consumers underestimate demand but lower them when consumers overesti-

[47] The latter effect is sufficiently strong that it is not overturned by the estimated negative aggregate mean bias.

mate demand. This prediction is illustrated by plan 1, fifty cents a minute for \$29.26 per month, and plan 2, unlimited talk for \$78.95. Plan 1 charges the maximum marginal price for every minute because its customers underestimate their demand. Plan 2 charges the minimum marginal price for every minute because its customers overestimate their demand.

Returning to columns 1 through 3, we see that adding consumer overconfidence and projection bias does not affect plan 1. Overconfident and projection biased consumers underestimate both left and right tails of their demand shock distribution. Plan 1 consumers, however, have sufficiently low expectations of demand that the left tail is censored at zero. As a result, overconfidence and projection bias only aggravate demand underestimation for plan 1 consumers. In contrast, overconfidence and projection bias are not effected by censoring for plan 2 customers and lead to three-part tariff pricing as predicted by Grubb (2009).

7 Conclusion

We specify and estimate a model of consumer cellular-phone plan and usage choices. We identify the distribution of consumer tastes from observed usage and consumers' beliefs about their future usage from observed plan choices. Comparing the two we find that consumers underestimate their average taste for calling, underestimate their own uncertainty about their average tastes, and underestimate the volatility of their tastes from month-to-month. Because the magnitude of overconfidence is substantially larger than that of projection bias, consumers correct initial plan choice mistakes more slowly than would unbiased consumers.

We then conduct counterfactual experiments where we (a) eliminate biases and (b) quantify the welfare impact of bill-shock regulation. We find that eliminating biases significantly increases consumer welfare, by about \$49 per consumer over one year. If prices do not respond to bill-shock regulation, consumers will save about \$39 per year, but at the cost of forgone phone calls, so average consumer welfare increases are a more moderate \$22 over one year. This finding is reversed when firms optimally respond to bill-shock regulation. Although firms increase the number of included minutes in three-part tariff offerings in response to bill-shock regulation, they also raise fixed fees, resulting in a net reduction in consumer welfare.

Although our results have clear implications for the wireless telephone industry, which is of growing importance in the world economy, they should have implications for many other product categories. For example, consumers face multipart tariffs when choosing and using many utilities, such as electricity and water. Our model could be used to inform policy makers about how to price these utilities in a manner that increases consumer welfare. Additionally, our evaluation of bill-shock regulation could be insightful in other relevant contexts as well. For instance, in 2009 US checking overdraft fees totalled more than $38 billion and have been the subject of new Federal Reserve Board regulation (Martin 2010, Federal Reserve Board 2009). Convincing evidence of consumer inattention (Stango and Zinman 2009, Stango and Zinman 2010) suggests that this fee revenue would be dramatically curtailed if the Fed imposed its own bill-shock regulation by requiring debit card processing terminals to ask users "$35 overdraft fee applies, continue Yes/No?" before charging fees. Our counterfactual shows that in the cellular context consumers are nevertheless made worse off after accounting for endogenously higher fixed fees. Interestingly, Grubb (2011) suggests that this might not be case for overdraft fees because "free checking" with zero monthly fees on accounts is the current industry norm. Although bill-shock regulation applied to overdrafts might mean an end to free checking, that in itself could intensify competition between banks due to the high salience of monthly account fees.

Finally, we comment on some future directions for this research. One possible avenue would be to relax the assumption of Bayesian learning. Work in experimental economics has suggested that consumer learning may not proceed according to Bayesian updating (Tversky and Kahneman 1974, Camerer 1995, Rabin 1998). It would be interesting to know how our findings would change under non-Bayesian learning. Another possible direction for future work would be to analyze a market in which consumers' decisions to experiment with new experience goods are important. Biases, such as negative mean bias or overconfidence, would tend to drive down the value of experimentation in the absence of switching costs. However, overconfidence and projection bias could have the reverse effect when switching costs are important. Switching costs should make a consumer less likely to experiment with a new product when she is uncertain of its quality because she would like to avoid being locked-in with a bad product. However, an overconfident consumer would be more sure of her prediction of the new product's quality prior to experimenting with it and hence under appreciate the risk of unwanted lock-in.

References

Ackerberg, Daniel A., "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination," *International Economic Review*, 2003, *44* (3), 1007–1040.

Armstrong, Mark and John Vickers, "Competitive Price Discrimination," *RAND Journal of Economics*, 2001, *32* (4), 579–605.

Bar-Gill, Oren and Rebecca Stone, "Pricing Misperception: Explaining Pricing Structure in the Cellular Service Market," *SSRN eLibrary*, 2009. `http://ssrn.com/paper=1425046`.

Baron, David P. and David Besanko, "Regulation and Information in a Continuing Relationship," *Information Economics and Policy*, 1984, *1* (3), 267–302.

Blundell, Richard and Stephen Bond, "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 1998, *87* (1), 115–143.

Borenstein, Severin, "To What Electricity Price Do Consumers Respond? Residential Demand Elasticity Under Increasing-Block Pricing," Preliminary Draft April 30 2009.

Busse, M. R., "Multimarket Contact and Price Coordination in the Cellular Telephone Industry," *Journal of Economics and Management Strategy*, 2000, *9* (3), 287–320.

Camacho, Nuno, Bas Donkers, and Stefan Stremersch, "Predictably Non-Bayesian: Quantifying Salience Effects in Physician Learning about Drug Quality," Working Paper June 2010.

Camerer, Colin, *Individual decision-making*, Princeton University Press, 1995.

Cardon, James H. and Igal Hendel, "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey," *The RAND Journal of Economics*, 2001, *32* (3), 408–427.

Ching, Andrew, Tülin Erdem, and Michael P. Keane, "The Price Consideration Model of Brand Choice," *Journal of Applied Econometrics*, 2009, *24* (3), 393–420.

Chintagunta, Pradeep, Puneet Manchanda, and S. Sriram, "Empirical Investigation of Consumer Adoption, Consumption, and Termination of a Video on Demand Service," Work in Progress 2009.

Conlin, Michael, Ted O'Donoghue, and Timothy J. Vogelsang, "Projection Bias in Catalog Orders," *American Economic Review*, 2007, *97* (4), 1217–1249.

Courty, Pascal and Hao Li, "Sequential Screening," *The Review of Economic Studies*, 2000, *67* (4), 697–717.

Crawford, Gregory S. and Matthew Shum, "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 2005, *73* (4), 1137–1173.

CTIA - The Wireless Association, "CTIA-The Wireless Association, Federal Communications Commission and Consumers Union Announce Free Alerts to Help Consumers Avoid Unexpected Overage Charges," Technical Report October 17 2011. `http://www.ctia.org/media/press/body.cfm/prid/2137`.

——, "Year-End 2010 Top-Line Survey Results," Technical Report 2011. `http://files.ctia.org/pdf/CTIA_Survey_Year_End_2010_Graphics.pdf`.

Cullen, Mark, Liran Einav, Amy Finkelstein, Stephen Ryan, and Paul Schrimpf, "Selection on Moral Hazard in Health Insurance," Work in Progress 2010.

DeGroot, Morris H., *Optimal Statistical Decisions*, New York: McGraw-Hill, 1970.

DellaVigna, Stefano and Ulrike Malmendier, "Contract Design and Self-Control: Theory and Evidence," *The Quarterly Journal of Economics*, 2004, *119* (2), 353–402.

Dutang, Christophe and Petr Savicky, "randtoolbox: Generating and Testing Random Numbers.," 2010. R package version 1.10 `http://cran.r-project.org/web/packages/randtoolbox/`.

Eliaz, Kfir and Ran Spiegler, "Contracting with Diversely Naive Agents," *The Review of Economic Studies*, 2006, *73* (3), 689–714.

—— **and** ——, "Consumer Optimism and Price Discrimination," *Theoretical Economics*, 2008, *3* (4), 459–497.

Erdem, Tülin and Michael P. Keane, "Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 1996, *15* (1), 1–20.

Federal Reserve Board, "Federal Reserve announces final rules prohibiting institutions from charging fees for overdrafts on ATM and one-time debit card transactions," Press Release November 12 2009.

fei Lee, Lung, "Estimation of dynamic and ARCH Tobit models," *Journal of Econometrics*, 1999, *92* (2), 355–390.

Gaynor, Martin S., Yunfeng Shi, Rahul Telang, and William B. Vogt, "Cell Phone Demand and Consumer Learning - An Empirical Analysis," *SSRN eLibrary*, 2005.

Gilbert, Daniel T., Michael J. Gill, and Timothy D. Wilson, "The Future Is Now: Temporal Correction in Affective Forecasting," *Organizational Behavior and Human Decision Processes*, 2002, *88* (1), 430–444.

Goettler, Ronald L. and Karen B. Clay, "Tariff Choice with Consumer Learning and Switching Costs," Working Paper 2010.

—— **and Ron Shachar**, "Spatial Competition in the Network Television Industry," *The RAND Journal of Economics*, 2001, *32* (4), 624–656.

Gourieroux, Christian and Alain Monfort, "Simulation-based inference : A survey with special reference to panel data models," *Journal of Econometrics*, 1993, *59* (1-2), 5–33.

Grubb, Michael D., "Selling to Overconfident Consumers," *American Economic Review*, 2009, *99* (5), 1770–1807.

——, "Bill Shock: Inattention and Price-Posting Regulation," mimeo 2011.

Hajivassiliou, Vassilis A. and Paul A. Ruud, *Classical Estimation Methods for LDV Models Using Simulation*, Amsterdam: North-Holland, 1994.

Herweg, Fabian and Konrad Mierendorff, "Uncertain Demand, Consumer Loss Aversion, and Flat-Rate Tariffs," *Journal of the European Economic Association*, Forthcoming.

Huang, Ching-I., "Estimating Demand for Cellular Phone Service Under Nonlinear Pricing," *Quantitative Marketing and Economics*, 2008, *6* (4), 371–413.

Ito, Koichiro, "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing," 2010.

Iyengar, Raghuram, Asim Ansari, and Sunil Gupta, "A Model of Consumer Learning for Service Quality and Usage," *Journal of Marketing Research*, 2007, *44* (4), 529–544.

——, Kamel Jedidi, and Rajeev Kohli, "A Conjoint Approach to Multipart Pricing," *Journal of Marketing Research*, 2008, *45* (2), 195–210.

Jiang, Lai, "The Welfare Effects of ''Bill Shock" Regulation in Mobile Telecommunication Markets," Working Paper July 2011.

Kim, Jiyoung, "Consumers' Dynamic Switching Decisions in the Cellular Service Industry," Working Paper, SSRN November 2006.

Lambrecht, Anja and Bernd Skiera, "Paying Too Much and Being Happy About It: Existence, Causes, and Consequences of Tariff-Choice Biases," *Journal of Marketing Research*, 2006, *43* (2), 212–223.

Leider, Steve and Özge Şahin, "Contracts, Biases and Consumption of Access Services," Ross School of Business Paper July 14 2011. `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1895468`.

Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence D. Phillips, *Calibration of Probabilities: The State of the Art to 1980*, Cambridge ; New York: Cambridge University Press,

Liebman, Jeffrey B. and Richard Zeckhauser, "Schmeduling," Working Paper October 2004.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin, "Projection Bias in Predicting Future Utility," *The Quarterly Journal of Economics*, 2003, *118* (4), 1209–1248.

Martin, Andrew, "Bank of America to End Debit Overdraft Fees," March 10 2010.

Miravete, Eugenio J., "Screening Consumers Through Alternative Pricing Mechanisms," *Journal of Regulatory Economics*, 1996, *9* (2), 111–132.

——, "Estimating Demand for Local Telephone Service with Asymmetric Information and Optional Calling Plans," *The Review of Economic Studies*, 2002, *69* (4), 943–971.

——, "Choosing the wrong calling plan? Ignorance and learning," *American Economic Review*, 2003, *93* (1), 297–310.

———, "The Welfare Performance of Sequential Pricing Mechanisms," *International Economic Review*, 2005, *46* (4), 1321–1360.

——— **and Ignacio Palacios-Huerta**, "Inattention, Choice Dependence, and Learning from Experience in a Repeated Decision Problem," mimeo 2011.

——— **and Lars-Hendrik Röller**, "Estimating Price - Cost Markups Under Nonlinear Pricing Competition," *Journal of the European Economic Association*, 2004, *2* (2-3), 526–535.

Narayanan, Sridhar, Pradeep K. Chintagunta, and Eugenio J. Miravete, "The role of self selection, usage uncertainty and learning in the demand for local telephone service," *Quantitative Marketing and Economics*, 2007, *5* (1), 1–34.

Nisbett, Richard E. and David E. Kanouse, "Obesity, hunger, and supermarket shopping behavior," 1968.

Osborne, Matthew, "Consumer Learning, Switching Costs and Heterogeneity: A Structural Examination," *Quantitative Marketing and Economics*, 2011, *9* (1), 25–70.

Park, Minjung, "The Economic Impact of Wireless Number Portability," 2009.

Rabin, Matthew, "Psychology and Economics," *Journal of Economic Literature*, 1998, *36* (1), 11–46.

Read, Daniel and Barbara van Leeuwen, "Predicting Hunger: The Effects of Appetite and Delay on Choice," *Organizational Behavior and Human Decision Processes*, 1998, *76* (2), 189–205.

Reiss, Peter C. and Matthew W. White, "Household Electricity Demand, Revisited," *The Review of Economic Studies*, 2005, *72* (3), 853–883.

Riordan, Michael H. and David E. M. Sappington, "Awarding Monopoly Franchises," *American Economic Review*, 1987, *77* (3), 375–387.

Rochet, Jean-Charles and Lars A. Stole, "Nonlinear Pricing with Random Participation," *The Review of Economic Studies*, 2002, *69* (1), 277–311.

——— **and Lars Stole**, *The Economics of Multidimensional Screening*, Vol. 36 of Econometric Society Monographs, New York: Cambridge University Press, 2003.

Roodman, David, "How to do xtabond2: An introduction to difference and system GMM in Stata," *Stata Journal*, 2009, *9* (1), 86–136.

Saez, Emmanuel, "Do Taxpayers Bunch at Kink Points?," Working Paper June 2002.

Seim, Katja and V. Brian Viard, "The Effect of Market Structure on Cellular Technology Adoption and Pricing," 2010.

Spiegler, Ran, *Bounded Rationality and Industrial Organization*, Oxford University Press, 2011.

Stango, Victor and Jonathan Zinman, "What do Consumers Really Pay on Their Checking and Credit Card Accounts? Explicit, Implicit, and Avoidable Costs," *American Economic Review Papers and Proceedings*, 2009, *99* (2).

——— **and** ——— , “Limited and Varying Consumer Attention: Evidence from Shocks to the Salience of Overdraft Fees,” 2010.

Stole, Lars A., *Price Discrimination and Competition*, Vol. Volume 3, Elsevier,

Train, Kenneth, *Discrete Choice Methods with Simulation*, 2nd ed., Cambridge University Press, 2009.

Tversky, Amos and Daniel Kahneman, “Judgment under Uncertainty: Heuristics and Biases,” *Science*, 1974, *185* (4157), 1124–1131.

U.S. Census Bureau, “Table 1. Preliminary Annual Estimates of the Resident Population for the United States, Regions, States, and Puerto Rico: April 1, 2000 to July 1, 2010,” Technical Report NST-PEST2010-01 February 2011. `http://www.census.gov/popest/states/tables/NST-PEST2010-01.xls`.

Uthemann, Andreas, “Competitive Screening of Customers with Non-Common Priors,” 2005. `http://www.ucl.ac.uk/~uctpaut/research/uthemann_ncp.pdf`.

Yao, Song, Carl F. Mela, Jeongwen Chiang, and Yuxin Chen, “Determining Consumers’ Discount Rates with Field Studies,” 2011.