PRICING MISPERCEPTION:

EXPLAINING PRICING STRUCTURE IN THE CELLULAR SERVICE MARKET

Oren Bar-Gill and Rebecca Stone*

1. Introduction

The cellular service market is an economically significant market that has substantially increased consumer welfare. The U.S. market has grown from 5 million subscribers in 1990 to 263 million subscribers in 2008. Eighty-six percent of Americans have a cell phone, and an increasing number of households rely entirely on wireless communications, giving up landlines altogether. Annual revenues of the four national carriers, AT&T, Verizon, Sprint and T-Mobile, total over $150 billion. In this paper, we focus on the pricing of cellular service. The common pricing structure in the cellular service market is a three-part tariff comprising: (1) a monthly charge; (2) a number of allotted minutes – talking minutes that the monthly charge pays for; and (3) an overage charge – a per-minute price for minutes beyond the plan limit.

We argue that the three-part tariff is a rational response by sophisticated carriers to consumers' misperceptions about their cell phone usage. The consumer chooses a calling plan based on a forecast of future use patterns. The problem is that many

* NYU School of Law. We thank Jennifer Arlen, Adi Ayal, Lucian Bebchuk, Kevin Davis, Clay Gillette, Michael Grubb, Raghuram Iyengar, Florenica Marotta Wurgler, Rick Pildes, Martin Raphan, Howard Shelanski, Phil Weiser and workshop participants at NYU for helpful comments and suggestions. We gratefully acknowledge the financial support of the D'Agostino/Greenberg Fund at NYU School of Law." We thank the Center for Customer Relationship Management at Duke University for letting us use their Telecom Dataset. Michael Biondi, Osnat Dafna and Paul McLaughlin provided outstanding research assistance.

consumers do not have a very good sense of their future use patterns. The three-part tariff is advantageous to carriers, because it exacerbates the effects of consumer misperception, leading consumers to underestimate the cost of cellular service.

To be more specific, some consumers underestimate their future usage, while others overestimate their future usage. Crucially, consumers are not aware that their estimates are incorrect, which enables firms to exploit their misperceptions. The overage fee component of the three-part tariff targets the underestimators. These consumers underestimate the probability of exceeding the plan limit and incurring an overage fee, and as a result will underestimate the cost of cellular service. The other components of the three-part tariff – the monthly charge and the fixed number of allotted minutes that come with it – target the overestimators. These consumers think that they will use all, or most, of their allotted minutes, and so expect to pay a per-minute price equal to the monthly charge divided by the number of allotted minutes. In fact, the overestimators end up using far fewer minutes and paying a much higher per-minute price. Thus, the overestimators also underestimate the cost of cellular service.

This paper provides empirical support for this theoretical account of the three-part tariff structure. We show that consumers inaccurately perceive their future usage. Since carriers should be credited with knowledge of these consumer misperceptions,[1] we can infer that the three-part tariff is used to exacerbate the effects of consumer misperception and to minimize the cost of cellular service as (mis)perceived by consumers.

We analyze a unique dataset of subscriber-level, monthly billing and usage information for 3,730 subscribers at a single wireless provider. These data allow us to

[1] According to a pricing manager at a top US cellular phone service provider, "people absolutely think they know how much they will use and it's pretty surprising how wrong they are." *See* Grubb (2008), p. 1, n. 2.

calculate not only the total cost of wireless service under each consumer's chosen plan, but also the total amount that he would have paid had he chosen other available plans. Thus, we can determine the plan that best fits his actual cell phone usage. We show that over 65% of consumers chose the wrong plan – some chose plans with an insufficient number of allotted minutes, indicating underestimation of future use, while others chose plans with an excessive number of allotted minutes, indicating overestimation of future use. For a substantial minority of consumers, the cost of these mistakes in plan choice is substantial. While the average per-consumer cost is only $47.68 a year, 17% of consumers made mistakes that cost them at least 20% of their total wireless bill, or $146 annually. This suggests that many consumers do not have a very good sense of their future use patterns.

The identified mistakes have substantial implications for consumer welfare. Extrapolating from our data to the entire U.S. population of cell phone users, we estimate that 42.5 million consumers made mistakes that cost them at least $146 annually. We estimate the total annual reduction in consumer surplus from the three-part tariff structure to be $11.92 billion. We outline a series of policy proposals that could address this problem. In particular, since consumers have poor information about their own use-patterns, we propose that carriers be required to disclose use-pattern information to consumers.

The paper is organized as follows. Section 2 reviews related literature. Section 3 develops a simple behavioral economics model, in which the three-part tariff emerges as the carriers' optimal response to consumer misperception. Section 4 describes the empirical analysis that supports our theoretical account. Section 5 offers concluding remarks, focusing on the normative and prescriptive implications of our analysis.

2. The Literature

This paper adds to a growing literature that seeks to explain the design of products, contracts and pricing schemes, across different markets, as a response by sophisticated sellers to predictable mistakes made by less sophisticated consumers. DellaVigna and Malmendier (2004) study optimal contract design when consumers have imperfect self-control, and overestimate their self-control. They show that their theoretical results are consistent with pricing practices in the health club and credit card markets. They also present stylized evidence on the gambling, life insurance, mail order, time-share and cell phone markets, and argue that this evidence is consistent with their theory. DellaVigna and Malmendier (2006) provide a rigorous empirical account of pricing as a rational response to imperfect self-control in the health club market. Gabaix and Laibson (2006) develop a theoretical model of add-on pricing as an optimal response to consumer misperception. Finally, Bar-Gill argues that credit card pricing (Bar-Gill 2004) and subprime-mortgage pricing (Bar-Gill 2009) can be explained as rational responses to consumer misperception.

The paper which is most closely related to ours is Grubb (2008). Grubb also studies three-part tariffs in the cellular service market. He analyzes a data set that comprises billing records for 2,332 student accounts managed by a major US university for a national US cellular phone service provider. His dataset is slightly smaller than ours and the extent to which students are representative of the larger population of cell phone users is an open question. More importantly, his theoretical explanation of three-part tariffs differs from ours in a crucial respect. Like us, he argues that three-part tariffs are a rational response to consumer misperceptions, but in his model consumer misperceptions take a different form from ours. On his account, consumers are overconfident about their ability to predict their future demand for cellular services. This means that they underestimate the variance in their use patterns, which implies that the *same* consumers sometimes overestimate their usage and sometimes underestimate their usage. On our account by contrast, some consumers overestimate their usage and *others* underestimate their usage. While it is difficult to conclusively prove the dominance of one account over the other, the data appears to be more consistent with our account and less consistent with Grubb's (see Section 3.2 below).

3. Model

3.1 Framework of Analysis

Consider a market in which there are N consumers. Each consumer's utility is increasing linearly in minutes of cell phone use up to a saturation point. Specifically, he receives α units of utility for each minute of cell phone use consumed up to a saturation level, s, and zero units of utility thereafter. Thus, a consumer will never talk for more than s minutes. The consumer's per-minute utility function is therefore given by:

$$v(m) = \begin{cases} \alpha & \text{if } m \le s \\ 0 & \text{if } m > s \end{cases}$$

and his overall utility is given by:

$$V(m) = m \cdot v(m) = \alpha \cdot \min(m, s).$$

A carrier incurs a fixed cost of C per customer served and zero marginal costs. We assume that $s \cdot \alpha - C \geq 0$, which ensures that it will be optimal for consumers to purchase *s* minutes of use from a carrier.

We allow for misperception about future use. In particular, we let $\hat{s}$ denote the consumer's perceived saturation point, and allow for $\hat{s} \neq s$. Consumers are heterogeneous along two dimensions: First, consumers have different perceived saturation points. Second, for any given perceived saturation point $\hat{s}$, there are three groups of consumers: (1) *A* consumers – consumers with accurate perceptions, $\hat{s}_A = s_A$; (2) *U* consumers – consumers who underestimate their saturation point, $\hat{s}_U < s_U$; and (3) *O* consumers – consumers who overestimate their saturation point, $\hat{s}_O > s_O$. We let $N_T^{\hat{s}}$ denote the number of consumers with "perceived type" $\hat{s}$ and "behavioral type" $T \in \{A, U, O\}$, and $N^{\hat{s}}$ denote the number of consumers of "perceived type" $\hat{s}$ regardless of behavioral type. We let $n_T^{\hat{s}} \equiv N_T^{\hat{s}} / N^{\hat{s}}$ denote the proportion of consumers of "perceived type" $\hat{s}$ who have "behavioral type" T, and $n^{\hat{s}} \equiv N^{\hat{s}} / N$ denote the proportion of consumers with perceived type $\hat{s}$ regardless of behavioral type. An agent's perceived type is private information, but the firm knows the proportions of *A*, *U* and *O* consumers for any given perceived type.

Once locked in, all consumers learn their true saturation points and choose their minutes of use accordingly. We consider price-schedules of the form $\left(F^{\hat{s}}, \overline{p}_i^{\hat{s}}\right)$, where *F* is the monthly fixed fee and $\overline{p}_i \equiv (p_1, p_2, ...)$ is a vector of per-minute prices (p_1 is the price charged for the first minute, p_2 is the price charged for the first minute, etc.). The

$\hat{s}$ superscripts in $\left(F^{\hat{s}}, \overline{p}_i^{\hat{s}}\right)$ imply that different price-schedules can be designed for consumers of different perceived types.

Since $u(m) \le \alpha$, carriers will always set $p_i \le \alpha$ in equilibrium. We make the following assumption:

Assumption 1: The per-minute price vector $\overline{p}_i \equiv (p_1, p_2, ...)$ is weakly increasing in *i*.

Assumption 1 reflects the practical implausibility of a diminishing marginal price. With diminishing marginal prices, subscribers would have an incentive to pool into groups and share phones (and service). Furthermore, in a more realistic model, reducing the per-minute cost would lead to increased usage, and could put strain on the carrier's network.

We first find the equilibrium in a monopolistic market and then look for separating equilibria in a competitive market.

3.2 Monopoly

A monopolist maximizes its profits subject to (1) the consumers' participation constraints, and (2) the consumers' incentive constraints, which ensure that a consumer with a perceived saturation point $\hat{s}$ chooses the tariff designed for consumers with a perceived saturation point $\hat{s}$. The monopolist's maximization problem is:

$$\max\left\langle \sum_{\hat{s}} n^{\hat{s}} \left(F + \sum_{i=0}^{\hat{s}} p_i^{\hat{s}} + n_U^{\hat{s}} \sum_{\hat{s}}^{s_U} p_i^{\hat{s}} - n_O^{\hat{s}} \sum_{s_O}^{\hat{s}} p_i^{\hat{s}} \right) - C \right\rangle$$

such that

$$\forall \hat{s} : \hat{s}\alpha - F^{\hat{s}} - \sum_{i=0}^{\hat{s}} p_i^{\hat{s}} \geq 0$$

$$\forall \hat{s} > 1, \hat{t} < \hat{s} : \sum_{i=0}^{\hat{t}} p_i^{\hat{t}} - \sum_{i=0}^{\hat{t}} p_i^{\hat{s}} \leq F^{\hat{s}} - F^{\hat{t}} \leq \sum_{i=0}^{\hat{s}} p_i^{\hat{t}} - \sum_{i=0}^{\hat{s}} p_i^{\hat{s}}$$

The following proposition describes the equilibrium outcome.

Proposition 1: In equilibrium, the monopolist offers a range of three-part tariffs such that a consumer of perceived type $\hat{s}$ pays $p_i^{\hat{s}} = 0$ for $i \leq \hat{s}$, $p_i^{\hat{s}} = \alpha$ for $i > \hat{s}$ and $F^{\hat{s}} = \hat{s}\alpha$.

Proof: If for all $\hat{s}$, $p_i^{\hat{s}} = 0$ for $i \leq \hat{s}$ and $p_i^{\hat{s}} = \alpha$ for $i > \hat{s}$ then the incentive constraints become:

$$\forall \hat{s} > 1, \hat{t} < \hat{s} : \; 0 \leq F^{\hat{s}} - F^{\hat{t}} \leq (\hat{s} - \hat{t})\alpha,$$

which is satisfied when $F^{\hat{s}} = \hat{s}\alpha$ for all $\hat{s}$ since then $F^{\hat{s}} - F^{\hat{t}} = (\hat{s} - \hat{t})\alpha$. It is straightforward to verify that each perceived type's participation constraint will also bind. Finally, since the monopolist extracts every type's perceived surplus, the tariffs must be profit-maximizing. □

Remarks:

1. The three-part tariff structure maximizes perceived consumer surplus while ensuring that Assumption 1 is satisfied. Perceived consumer surplus of the underestimators is

maximized by setting the price of included minutes as low as possible, $p_i^{\hat{s}} = 0$ for $i \le \hat{s}$, while making the overage price as high as possible, $p_i^{\hat{s}} = \alpha$ for $i > \hat{s}$, since underestimators erroneously believe that they will use only $\hat{s}$ minutes of their plans and will therefore never pay the high overage charges. The overestimators correctly perceive that they will not pay the overage charges and so are correctly indifferent to a high overage, while they erroneously perceive that they will consume minutes s_O to $\hat{s}$ of their plan. Thus, their perceived consumer surplus is maximized by setting $p_i^{\hat{s}} = 0$ for $s_O < i \le \hat{s}$, which, in conjunction with Assumption 1, entails that $p_i^{\hat{s}} = 0$ for $i \le \hat{s}$.

2. A consumer who chooses the tariff designed for a lower perceived type pays a lower fixed charge, but this is exactly offset by the additional overages he expects to pay should he choose that tariff, rendering him indifferent between this tariff and the tariff designed for his perceived type. He strictly prefers the tariff designed for his perceived type to a tariff for a higher perceived type since, while he does not expect to pay overages under either, he must pay a higher fixed charge if he chooses the latter.

3. Consumers do not benefit from the heterogeneity in perceived types. No type makes informational rents in equilibrium and the carrier sets $F^{\hat{s}}$ in order to extract the entire perceived consumer surplus of each perceived type.

3.3 Perfect Competition

In a competitive market, the optimal pricing scheme maximizes each consumer's perceived utility subject to carriers breaking even (this ensures that entry into the market is not profitable). In addition, each consumer must have an incentive to choose the price schedule that has been designed for consumers of his perceived type.

The equilibrium pricing scheme must solve a set of maximization problems such that for each perceived saturation point $\hat{s}$:

$$\max\left\langle \hat{s}\cdot\alpha - F^{\hat{s}} - \sum_{i=0}^{\hat{s}} p_i^{\hat{s}} \right\rangle$$

such that

$$\left(F^{\hat{s}} + \sum_{i=0}^{\hat{s}} p_i^{\hat{s}}\right) + n_U^{\hat{s}} \sum_{\hat{s}}^{s_U} p_i^{\hat{s}} - n_O^{\hat{s}} \sum_{s_O}^{\hat{s}} p_i^{\hat{s}} \geq C$$

$$\forall \hat{t} \neq \hat{s}:\ \hat{s}\cdot\alpha - F^{\hat{s}} - \sum_{i=0}^{\hat{s}} p_i^{\hat{s}} \geq \hat{s}\cdot\alpha - F^{\hat{t}} - \sum_{i=0}^{\hat{s}} p_i^{\hat{t}}$$

Let $\tilde{U}(\hat{s}) = n_U^{\hat{s}}(s_U - \hat{s})$ denote the degree of underestimation – the number of underestimators multiplied by the extent of underestimation – for a perceived saturation point $\hat{s}$. We make the following assumptions about the relative degrees of underestimation in the market.

Assumption 2:

(i) The degree of underestimation is (weakly) increasing in the perceived saturation point, i.e., $\tilde{U}'(\hat{s}) \geq 0$.

(ii) The marginal degree of underestimation is lower than or equal to 1, i.e., $\tilde{U}'(\hat{s}) \leq 1$.

We view Assumption 2(i) as a weak assumption, formalizing the notion that for a higher perceived saturation point, the likelihood and magnitude of underestimation should be smaller. Assumption 2(ii) is less innocuous. It is, however, sufficient for the existence of a separating equilibrium. We note that Assumption 2(ii) is not necessary for the existence of a separating equilibrium. Appendix A describes the conditions for a separating equilibrium in a market with only two perceived saturation points, when Assumption 2(ii) is not satisfied.

The following proposition describes a separating equilibrium outcome.

Proposition 2: If Assumption 2 is satisfied, i.e., if $0 \leq \tilde{U}'(\hat{s}) \leq 1$, then there is a separating equilibrium in which carriers offer a consumer with perceived saturation point $\hat{s}$ a three-part tariff with $p_i^{\hat{s}} = 0$ for $i \leq \hat{s}$, $p_i^{\hat{s}} = \alpha$ for $i > \hat{s}$ and $F^{\hat{s}} = C - n_U^{\hat{S}} \cdot (s_U - \hat{s}) \cdot \alpha$.

Proof: Since no consumer believes he will consume more than his perceived saturation point, the utility of each perceived type is maximized if $p_i^{\hat{s}} = 0$ for $i \leq \hat{s}$, $p_i^{\hat{s}} = \alpha$ for $i > \hat{s}$ and $F^{\hat{s}} = C - n_U^{\hat{S}} \cdot (s_U - \hat{s}) \cdot \alpha$, which ensures that the carrier's participation constraint is binding and so $F^{\hat{s}}$ is as low as possible. With $p_i^{\hat{s}} = 0$ for $i \leq \hat{s}$ and $p_i^{\hat{s}} = \alpha$ for $i > \hat{s}$, the incentive constraints reduce to $\forall \hat{s} > 1,\ \hat{s} > \hat{t} : 0 \leq F^{\hat{s}} - F^{\hat{t}} \leq (\hat{s} - \hat{t}) \cdot \alpha$. Substituting $F^{\hat{s}} = C - n_U^{\hat{S}} \cdot (s_U - \hat{s}) \cdot \alpha$, the incentive constraints become:

$$\forall \hat{s} > 1,\ \hat{s} > \hat{t} : 0 \leq \left[n_U^{\hat{t}} \cdot (t_U - \hat{t}) - n_U^{\hat{S}} \cdot (s_U - \hat{s}) \right] \cdot \alpha \leq (\hat{s} - \hat{t}) \cdot \alpha$$
$$\Leftrightarrow \forall \hat{s} > 1,\ \hat{s} > \hat{t} : 0 \leq n_U^{\hat{t}} \cdot (t_U - \hat{t}) - n_U^{\hat{S}} \cdot (s_U - \hat{s}) \leq (\hat{s} - \hat{t})$$

which is satisfied when Assumption 2 holds.□

Remarks:

1. In equilibrium, U consumers are cross-subsidizing the A and O consumers. The U consumers end up paying overages on minutes $\hat{s}$ to s_U. Competition then pushes down the fixed charges by the amount that carriers' earn from this underestimation, thus passing on the windfall to the A and O consumers. This situation persists in equilibrium since no consumer believes that he will pay the overage fee when he signs up with the carrier. Thus, to attract U consumers of a particular perceived type away from a carrier setting the competitive tariff, a competing carrier would have to lower the monthly fixed fee, $F^{\hat{s}}$, which would entail a loss.

2. When assumption 2 holds, the heterogeneity of perceived types is non-distortionary as the carriers' participation constraints bind in equilibrium. However, this will not always be the case. When assumption 2 fails, it is not always possible for both the participation constraints to bind and the incentive compatability constraints to hold. Under these conditions, at least one of the participation constraints will end up slack to ensure incentive compatibility holds. But then it may be possible for a carrier to profitably steal customers from firms offering such a tariff by lowering the fixed cost offered to the relevant perceived types. Such a tariff will not satisfy incentive compatibility and so will attract consumers of more than one perceived type. But if the proportions of different perceived types in the population at large are right, the (partial) pooling tariff will be profitable thus undermining the existence of a separating equilibrium as in Rothschild and Stiglitz's (1976) competitive insurance markets.[2] We show how this can occur when Assumption 2 fails and there are only two perceived types in Appendix A.

[2] See also Grubb's (2008) discussion of perfectly competitive multi-tariff menus when overconfident consumers differ in their beliefs and in the true distributions of demand.

Proposition 2 states conditions for the existence of a separating equilibrium with a three-part tariff structure. The flipside, which is implied by Proposition 2 and the accompanying discussion, is that there are conditions, under which a separating equilibrium with a three-part tariff does not exist. It is of interest to note, however, that *whenever* a separating equilibrium exists in which each perceived type strictly prefers the tariff designed for his perceived type, it will have a three-part tariff structure. This result is stated in the following proposition.

Proposition 3: Whenever a separating equilibrium exists in which each perceived type strictly prefers the tariff designed for his perceived type to the tariff designed for any other perceived type, the equilibrium tariff offered to consumers with perceived saturation point $\hat{s}$ takes the form of a three-part tariff with $p_i^{\hat{s}} = 0$ for $i \leq \hat{s}$ and $p_i^{\hat{s}} = \alpha$ for $i > \hat{s}$.

Proof: See Appendix B.

4. Empirical Analysis: Testing the Model

4.1 The Data

We obtained a unique dataset of subscriber-level, monthly billing and usage information for 3,730 subscribers at a single wireless provider. These data provide information on which of four calling plans a subscriber has chosen and his monthly consumption of peak minutes for the period of September 2001 to May 2003. Each of the four calling plans offer a standard three-part tariff with a fixed allocation of peak minutes and steep overages for additional peak minutes consumed (off-peak and weekend minutes are free) as described in the table 1 below.[3]

	Plan 1	Plan 2	Plan 3	Plan 4
Market share (%)	47.36	9.92	32.1	10.62
Monthly fixed charge ($)	30	35	40	50
Number of included minutes	200	300	400	500
Overage rate ($)	0.40	0.40	0.40	0.40

Table 1: Menu of Three-Part Tariffs

The data reveal substantial variance in usage. Summary statistics are provided in tables 2a-2d below. Tables 2a-2c present the overall mean and standard deviation of minutes used for plans 1, 3 and 4, respectively. We omit information on plan 2 since no plan 2 subscriber remained with the plan for more than 10 months, which, as we explain below, precludes a meaningful analysis of mistakes in plan choice. To gain an initial sense of underestimation versus overestimation of usage, we also present, for each plan, average use figures for subscribers who use less than their included minutes (the average "underage") and for subscribers who use more than their included minutes (the average "overage"). We then aggregate this information across all three plans in table 2d.

	Plan 1		
	Share	Usage/Allowance	
		Mean	Std. Dev.
Under Allowance	0.815	0.45	0.294

[3] The database was provided by the Center for Customer Relationship Management at Duke University. The description of the data in the text is based on the decription provided by the Center for Customer Relationship Management, FUQUA School of Business, Duke University (http://www.fuqua-europe.duke.edu/centers/ccrm/index.html#data). *See also* Iyengar, Ansari and Gupta (2007). The four plans are offered with many different optional features that consumers can choose from, including messaging, long-distance and roaming. Iyengar, Ansari and Gupta determined that actual use of these features was negligible in the data set and thus ignored the added variation in contractual design. We do the same. Furthermore, it is not entirely clear from the data that all four plans were offered at all dates in all markets. We acknowledge this limitation of the data and qualify our results accordingly.

Over Allowance	0.178	1.46	0.624
All Subscribers	1	0.633	0.538

Table 2a: Summary Statistics – Plan 1

	Plan 3		
	Share	Usage/Allowance	
		Mean	Std. Dev.
Under Allowance	0.836	0.466	0.297
Over Allowance	0.16	1.284	0.343
All Subscribers	1	0.599	0.428

Table 2b: Summary Statistics – Plan 3

	Plan 4		
	Share	Usage/Allowance	
		Mean	Std. Dev.
Under Allowance	0.717	0.573	0.296
Over Allowance	0.278	1.259	0.29
All Subscribers	1	0.766	0.424

Table 2c: Summary Statistics – Plan 4

	All Plans		
	Share	Usage/Allowance	
		Mean	Std. Dev.
Under Allowance	0.813	0.466	0.297
Over Allowance	0.165	1.326	0.433
All Subscribers	1	0.612	0.456

Table 2d: Summary Statistics – Aggregate

To summarize, in aggregate, subscribers exceed their minute allowance 16.5% of the time, by an average of 32.6%. And during the 81.3% of the time when the allowance is not exceeded, subscribers use only 46.6% of their minute allowance on average. It should be emphasized that underages and overages do not, in and of themselves, imply overestimation and underestimation of use, respectively. A perfectly rational subscriber, with variable use will experience underages and overages. We must do more to prove our claim that the three-part tariff structure responds to subscribers' misperception of use.

4.2 Overestimation by Some and Underestimation by Others

In Section 2, we described the key difference between our behavioral account and Grubb’s behavioral account. While Grubb claims that the same consumers sometimes overestimate their usage and sometimes underestimate their usage, we argued that some consumers overestimate their usage and *others* underestimate their usage. In this section, we argue that the data appear to be more consistent with our account.

We calculate, for each subscriber, the percentage of overages (bills in which the subscriber exceeded the plan limit) and the percentage of underages (bills in which the subscriber did not exhaust his allotted minutes). The correlation between the percentages of overages and the percentages of underages across subscribers is -0.998. This strong negative correlation suggests that some subscribers experience overages while others experience underages – a result which is more consistent with our theory and less consistent with Grubb's.

To bolster this finding, we redo the analysis focusing on subscribers with substantial overages and underages. We define a substantial overage as an overage that is greater than the average overage in our data, 32.6% above the plan limit (Table 2d). Similarly, we define a substantial underage as an underage that is greater than the average underage in our data, 53.4% below the plan limit (Table 2d). We then calculate, for each subscriber, the percentage of substantial overages and the percentage of substantial underages. The correlation between the percentages of substantial overages and the percentages of substantial underages across subscribers is -0.363. Again, the negative correlation supports our behavioral account and is less consistent with Grubb's.

Next, we count the number of subscribers who experienced at least two substantial overages, 686, the number of subscribers who experienced at least two substantial underages, 2827, and the number of subscribers who experienced at least two substantial overages and at least two substantial underages, 449. (We focus on subscribers with two substantial overages and/or underages, since a single substantial overage/underage may not be representative.) This analysis lends further support to our theory that most subscribers either overestimate their use or underestimate their use, and is less consistent with Grubb's overconfidence theory. The vast majority, 78% (= (2827-449)/(2827+686-449)), of subscribers with substantial biases are overestimators. A small minority, 8% (= (686-449)/(2827+686-449)), are underestimators. And only a relatively modest 15% (= 449/(2827+686-449)) experience both substantial overages and substantial underages, the likely use pattern if subscribers suffer from an overconfidence bias as Grubb suggests.[4]

[4] These results are not strictly inconsistent with Grubb's overconfidence theory. Variance in usage can be the product of a single large random variable that affects usage across multiple months, e.g., will the subscriber get a new job that requires much travel. Underestimating the probability of such an occurrence can lead to overconfidence about usage, and the resulting use patterns would be consistent with the data described above: Some subscribers would not get the new job and would experience only underages, while others would get the job and experience only overages. The problem with this story is that variance in usage usually is not the product of a single large random variable that affects usage across multiple months,

4.3 Choosing the Wrong Plan

We next estimate the percentage of consumers who arguably chose the wrong plan, and the cost of the mistake to these consumers. We consider a plan choice to be a mistake when, given the consumer's usage, a different plan would have cost the consumer less. We limit our analysis to the 3,456 consumers who stayed with a plan for at least 10 months, and take, as our unit of analysis, the consumer's tenure with a plan. Given the natural variance in usage from month to month, identifying mistakes over shorter time horizons is less reliable. For each of the 3,456 consumers, we calculate the total cost of wireless service under the consumer's chosen plan and compare it to the total amount that this consumer would have paid had she chosen each of the other three plans. We measure the magnitude of the mistake by the difference, in both percentage and dollar terms, between the consumer's actual costs and the lowest possible cost – the cost that the consumer would have paid if she could have predicted her usage with certainty.[5]

The results are collected in tables 3a and 3b. In these tables, each row represents the group of subscribers who chose a certain plan. This group is then divided into four sub-groups according to the plan that these subscribers *should* have chosen. For instance, the cell located at the intersection of the Plan 3 row and the Plan 1 column represents the sub-group of subscribers who chose plan 3 but should have chosen plan 1. Table 3a presents the size, in percentage terms, of these sub-groups. Table 3b presents the

---

but rather the product of multiple small random variables that affect usage on a month-to-month basis (e.g., will I have to work much overtime in a specific month, will I go on an out-of-town vacation in a specific month, will my friend in Chicago fall ill in a specific month, etc.) And in this case, overconfidence would predict use patterns that are inconsistent with the data described above.

[5] This analysis assumes risk neutrality. The sums involved are small enough that this assumption seems reasonable.

magnitude of the mistakes (or cost-savings), both in percentage terms and in annual dollar terms, for each sub-group.

		Optimal Plan			
		Plan 1	Plan 2	Plan 3	Plan 4
Chosen Plan	Plan 1	74.09%	21.79%	1.49%	2.49%
	Plan 3	27.20%	35.61%	21.19%	16%
	Plan 4	9.00%	10.66%	8.00%	73.33%

Table 3a: The likelihood of mistakes

		Optimal Plan			
		Plan 1	Plan 2	Plan 3	Plan 4
Chosen Plan	Plan 1	0% $0	9.56% $54.16	26.97% $203.58	28.22% $341.71
	Plan 3	21.09% $101.58	6.55% $32.59	0% $0	11.34% $102.98
	Plan 4	36.71% $220.27	12.38% $75.31	7.00% $39.90	0% $0

Table 3b: The magnitude of mistakes

We present the results for one group of subscribers, those who chose plan 3, in figure 1. We focus on this group of subscribers, since it includes significant numbers of both underestimators, who should have chosen plan 4, and overestimators, who should have chosen either plan 2 or plan 1. Figure 1 displays the share of consumers, among those who chose plan 3, who should have chosen each of the four plans (the dark gray bars), and for those who should not have chosen plan 3 the amount of money they would have saved, both in percentage terms (the light gray bars) and in dollar figures.

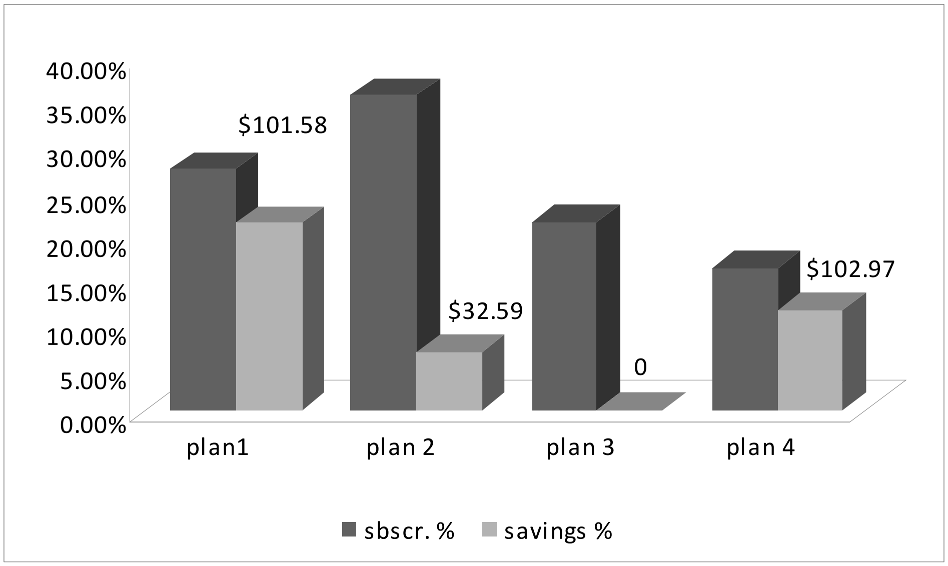

Figure 1: Plan 3 subscribers – the likelihood and magnitude of mistakes

These figures underestimate the number and cost of mistakes, especially for plans with a lower allocation of minutes. For example, for subscribers who chose plan 1 our data reveal only mistakes arising from underestimation of use when the subscriber should have chosen plan 2, plan 3 or even plan 4. In our data, which includes only the four postpaid plans, a subscriber who chose plan 1 could not have made a mistake that arises from overestimation of use. But, in fact, many plan 1 subscribers, who overestimated their use, probably could have done better by choosing a prepaid plan. We offer a conservative estimate of the number and magnitude of such mistakes by adding a hypothetical prepaid plan with a high per-minute charge of 40 cents (equal to the overage charges in our data). 24.4% of plan 1 subscribers would have saved money, $149 annually on average, had they chosen the prepaid plan.[6]

[6] These conclusions are tentative, since pre-paid plans may defer from post-paid plans on other dimensions. In particular, while the service quality offered by pre-paid plans is improving, in the period when the data were collected there was still a non-negligible difference in quality between pre-paid and post-paid plans.

We acknowledge that, in theory, the use patterns revealed in our data are consistent with the behavior of perfectly rational, but risk averse subscribers. These subscribers would choose plans with more allotted minutes to reduce the risk of paying substantial overage fees. And most of these subscribers will end up using much less than their allotted minutes. This explanation fails for two reasons. First, given the small sums of money involved the observed plan choices are not consistent with risk aversion under Expected Utility Theory (Rabin 2000). (However, they may be consistent with certain behavioral accounts of risk aversion.) Second, even if risk aversion could explain the patterns of overage and underage given the three-part tariff structure, it cannot explain the emergence of the three-part tariff as the equilibrium pricing structure. In fact, with rational, risk-averse subscribers, we would expect to see two-part tariffs.

To sum up, many consumers fail to accurately anticipate their use patterns, and the three-part tariff design can be explained as a market response to such use-pattern mistakes. Telling evidence that providers are not troubled, to say the least, by consumers' use-pattern mistakes is that they foster these mistakes by requiring, as a condition for network access, that handset manufacturers disable the call-timer feature (Wu, 2007; Nachman, 2002). Consumers are becoming more aware of their use-pattern mistakes and more frustrated with carriers who take advantage of these mistakes. A 2007 survey found that complaints about unexpected overage fees are the most common complaints voiced by subscribers (PersonalMaestro, 2007). This survey also identified a high demand for an external “minutes alert” service. Consumers are learning. And, the market is responding to the demand generated by these more sophisticated consumers through, e.g., unlimited calling plans and prepaid plans.

5. Conclusion

We have attempted to provide a descriptive account of consumer mistakes in the cellular service market – mistakes that are exacerbated by the three-part tariff structure prevalent in this market. In this Conclusion, we briefly explore the normative and prescriptive implications of our descriptive account.

5.1 Welfare Implications

The average consumer in our data made a mistake that cost him 8% of his total wireless bill, or $47.68 annually. Extrapolating from our data onto the entire U.S. population of cell phone users, numbering 250 million, we obtain a $11.92 billion annual reduction in consumer surplus. While the $11.92 billion figure is substantial, the average per-consumer harm, $47.68, is small. But averages hide potentially important distributional implications. The $11.92 billion are not evenly divided among the $250 million U.S. subscribers. In our data, 35% of subscribers chose the right plan. And even among subscribers who chose the wrong plan, the magnitude of the mistake, i.e., the extra payment as compared to the right plan, varies substantially. In our data, 34% of consumers made mistakes that cost them at least 10% of their total wireless bill, or $113 annually, and 17% of consumers made mistakes that cost them at least 20% of their total wireless bill, or $146 annually. (10% of consumers made mistakes that cost them at least 25% of their total wireless bill, or $60 annually; this implies that the really large mistakes, in percentage terms, had smaller stakes, in dollar terms.)

While harm to consumers is important, it should be emphasized that a reduction in the consumer surplus is not a welfare cost in and of itself. Yet the identified consumer mistakes do generate welfare costs: First, consumer mistakes imply allocative inefficiency, since consumers are not buying the right product for them. Second, social welfare is reduced by regressive redistribution. Such redistribution occurs when carriers profit from consumer mistakes. But regressive redistribution occurs even if these excess profits are competed away. The distribution of mistakes implies that revenues from consumers who make mistakes keep prices low for consumers who do not make mistakes.

5.2 Policy Implications

Consumers are making costly mistakes, but consumers are also learning from their mistakes and the market is responding to the demand generated by these increasingly sophisticated consumers. We are, therefore, skeptical about hard-handed regulatory intervention in the cellular service market. Still, we believe that smart regulation that facilitates rather than inhibits market forces can reduce the adverse consequences of consumer misperception.

We focus on disclosure regulation. Our proposals, however, deviate from existing disclosure regulation and from other proposals for heightened disclosure regulation in an important way. Current disclosure regulation and other proposals focus on the disclosure of product attribute information, i.e., information on the different features and price dimensions of cellular service.[7] Our proposal, on the other hand, emphasizes the disclosure of use-pattern information, i.e., information on how the consumer will use the

[7] See, e.g., the proposed Cell Phone User Bill of Rights (Schumer, 2003).

product. The emphasis on use-pattern information follows directly from our descriptive analysis, which revealed systemic misperceptions about usage levels.

Conventional wisdom assumes that sellers have better information on product attributes while buyers have better information about use patterns. If a buyer has better information about how he will use the product, then it makes no sense to require sellers to disclose use-pattern information. While in many markets the conventional wisdom is correct, it is not true of the cellular service market. Carriers have valuable statistical use-pattern information that is not available to subscribers. More importantly, they have individualized use-pattern data, collected over the course of their relationships with their subscribers. Disclosure of this information can empower consumers and facilitate the efficient functioning of the cellular service market.

At the average-use level, carriers could be required to disclose the average overage charges that consumers pay. Carriers could also be required to disclose the percentage of consumers who use, say, 50% or less of the allotted minutes or the percentage of consumers who would save money if they switched to a lower fixed-fee, lower limit plan.

At the individual-use level, carriers already provide consumers with information on overage charges on their monthly bills. Arguably, this disclosure has reduced consumers' underestimation of use and created the impetus for the demand to eliminate overage fees—a demand that is now partially met by unlimited calling plans for high-demand users. We propose a parallel disclosure that would help reduce the costs to consumers of overestimation of use. Carriers should be required to disclose the number of minutes used. (Some carriers already do so voluntarily.) Moreover, they should be required to disclose the actual per-minute price, calculated as the monthly fixed-fee divided by the number of minutes used.[8]

We have focused on the disclosure of use-pattern information, but, in fact, the more appealing proposals argue for total cost disclosures, which combine both product attribute and use-pattern information. For example, the disclosure of actual per-minute prices combines product attribute information, the monthly fixed-fee, and use-pattern information, the number of minutes used. Taking total cost disclosure one step further, carriers can be required to disclose a comprehensive total cost of ownership (TCO) figure for their calling plans—the total amount paid, or to be paid, by a consumer, over the duration of a plan (or on a yearly basis). For new subscribers, this TCO figure can be based on average-use information. For existing subscribers, who are considering whether to renew their plan, switch plans or even switch carriers, the TCO figure can be based on individual-use information.

TCO information for a single plan, specifically for the subscriber's current plan, may be insufficient. To effectively compare different plans, the subscriber needs TCO information on all plans. Carriers can be required to provide TCO information for their entire menu of plans or, at least, for several main offerings. Perhaps a better solution would be to require carriers to disclose only the plan with the lowest TCO for the prospective subscriber and for the existing subscriber whose use patterns have changed. An even better solution would utilize the emerging market for "comparison-shopping services." Companies like BillShrink.com and FixMyCellBill.com promise to find the

[8] A related disclosure is voluntarily implemented by the carrier Orange in France. In many plans, Orange offers to monitors customers' consumption with a report of minutes not consumed ("le report des minute non consommees"). Orange also offers to send a text message notification when plan limits are exceeded ("l'alerte par SMS en cas de depassement du forfait"). *See* http://www.orange.fr.

right plan for you. But they currently do this based on minimal, usually self-reported, use-pattern information. If carriers are required to provide comprehensive use-pattern information in electronic form, the subscriber could then take this information to BillShrink or FixMyCellBill and they will find the carrier and the plan with the lowest TCO for that subscriber.

References

Bar-Gill, Oren (2004), Seduction By Plastic, *Northwestern University Law Review*, 98, 1373-1434.

Bar-Gill, Oren (2009), The Law, Economics and Psychology of Subprime Mortgage Contracts, *Cornell Law Review*, 94, [TBA].

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

DellaVigna, Stefano, and Malmendier, Ulrike (2006), Paying Not to Go to the Gym, *American Economic Review*, 96, 694-719.

Gabaix, Xavier, and Laibson, David (2006), Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets, *Quarterly Journal of Economics*, 121, 505-540.

Grubb, Michael (2008), Selling to Overconfident Consumers, MIT Sloan Working Paper.

Iyengar, Raghuram, Ansari, Asim, and Gupta, Suni L. (2007), A Model of Consumer Learning of Consumer Service Quality and Usage, *Journal of Marketing Research*, 44, 529-544.

Nachman, Sherrie (2002), Cranky Consumer: How to Check Up on Your Cell Phone Minutes, *Wall Street Journal*, June 18, 2002.

PersonalMaestro.com (2007), Press Release: New Survey - Cell Phone Overage Fees Number One Complaint Against Cellular Providers, Feb. 19, 2007 (available at http://www.1888pressrelease.com/new-survey-cell-phone-overage-fees-number-one-complaint-ag-pr-8ps51je411.html).

Rabin, Matthew (2000), Risk Aversion and Expected-Utility Theory: A Calibration Theorem, *Econometrica*, 68, 1281-1292.

Rothschild and Stiglitz (1976), Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information, *Quarterly Journal of Economics*, 90, 629-49.

Schumer, Charles (2003), Press Release: Schumer Unveils First Comprehensive Cell Phone User Bill of Rights, Feb. 25, 2003 (available at http://www.senate.gov/~schumer/SchumerWebsite/pressroom/press_releases/PR01504.html).

Wu, Tim (2007), Wireless Net Neutrality: Cellular Caterfone on Mobile Networks, New America Foundation, Wireless Future Program Working Paper No. 17 (ver. 2.1).

Appendix A

This Appendix extends the analysis in Section 3.3, deriving the conditions for the existence of a separating equilibrium in a market with only two perceived saturation points, when Assumption 2(ii) is not satisfied. Note that when there are only two perceived saturation points, $\hat{s}$ and $\hat{t}$, and $\hat{s} > \hat{t}$, Assumption 2 becomes: $\frac{n_U^{\hat{t}}\left(t_U - \hat{t}\right) - n_U^{\hat{s}}\left(s_U - \hat{s}\right)}{\hat{s} - \hat{t}} \leq 1$.

The results are summarized in the following proposition.

<u>Proposition 1A</u>: In a market with only two perceived saturation points, $\hat{s}$ and $\hat{t}$ where $\hat{s} > \hat{t}$, when Assumption 2(ii) is not satisfied, i.e., when $\frac{n_U^{\hat{t}}\left(t_U - \hat{t}\right) - n_U^{\hat{s}}\left(s_U - \hat{s}\right)}{\hat{s} - \hat{t}} > 1$, then if $\frac{n_U^{\hat{t}}\left(t_U - \hat{t}\right) - n_U^{\hat{s}}\left(s_U - \hat{s}\right)}{\hat{s} - \hat{t}} \leq 1 + \frac{n^{\hat{s}}}{n^{\hat{t}}} n_O^{\hat{s}}$ there exists a separating equilibrium in which carriers offer two three part tariffs, one designed for each perceived type—a tariff consumed by $\hat{s}$ perceived types where $p_i^{\hat{s}} = 0$ for $i \leq \hat{s}$, $p_i^{\hat{s}} = \alpha$ for $i > \hat{s}$ and $F^{\hat{s}} = C - n_U^{\hat{S}} \cdot \left(s_U - \hat{s}\right) \cdot \alpha$ and a tariff consumed by $\hat{t}$ perceived types where $p_i^{\hat{t}} = 0$ for $i \leq \hat{t}$, $p_i^{\hat{t}} = \alpha$ for $i > \hat{t}$ and $F^{\hat{t}} = C - n_U^{\hat{s}}\left(s_U - \hat{s}\right)\alpha - \left(\hat{s} - \hat{t}\right) \cdot \alpha$.

<u>Proof</u>:

The utility of each perceived type is maximized by setting $p_i^{\hat{s}} = 0$ for $i \le \hat{s}$, $p_i^{\hat{s}} = \alpha$ for $i > \hat{s}$ and $p_i^{\hat{t}} = 0$ for $i \le \hat{t}$, $p_i^{\hat{t}} = \alpha$ for $i > \hat{t}$ and $F^{\hat{s}}$ and $F^{\hat{t}}$ is as low as possible. Therefore, to ensure that the participation constraints are satisfied, we require that

$$F^{\hat{s}} + n_U^{\hat{s}}(s_U - \hat{s})\alpha \ge C$$
$$F^{\hat{t}} + n_U^{\hat{t}}(t_U - \hat{t})\alpha \ge C$$

and to ensure that the incentive compatibility constraints are satisfied we require that

$$0 \le F^{\hat{s}} - F^{\hat{t}} \le (\hat{s} - \hat{t}) \cdot \alpha \Leftrightarrow F^{\hat{t}} \ge F^{\hat{s}} - (\hat{s} - \hat{t}) \cdot \alpha$$

The failure of Assumption 2(ii), $(\hat{s} - \hat{t}) < n_U^{\hat{t}}(t_U - \hat{t}) - n_U^{\hat{s}}(s_U - \hat{s})$, means that only the participation constraint for the perceived $\hat{s}$ types binds in equilibrium. If the participation constraint for the perceived $\hat{t}$ types binds, $F^{\hat{t}} = C - n_U^{\hat{t}}(t_U - \hat{t})\alpha$, then incentive compatibility implies that $F^{\hat{s}} \le (\hat{s} - \hat{t})\alpha + F^{\hat{t}} = C + [(\hat{s} - \hat{t}) - n_U^{\hat{t}}(t_U - \hat{t})]\alpha < C - n_U^{\hat{s}}(s_U - \hat{s})\alpha$, which means that the participation constraint for the perceived $\hat{s}$ types is not satisfied. If, on the other hand, the participation constraint for the perceived $\hat{s}$ types binds $F^{\hat{s}} = C - n_U^{\hat{s}}(s_U - \hat{s})\alpha$, then incentive compatibility implies that $F^{\hat{t}} \ge C - n_U^{\hat{s}}(s_U - \hat{s})\alpha - (\hat{s} - \hat{t}) \cdot \alpha > C - n_U^{\hat{t}}(t_U - \hat{t})\alpha$, which means that the participation constraint for the perceived $\hat{t}$ types is strictly satisfied. $F^{\hat{t}}$ is as small as possible when the incentive constraint binds, and therefore fixed charges in a separating equilibrium are given by $F^{\hat{s}} = C - n_U^{\hat{s}}(s_U - \hat{s})\alpha$ and $F^{\hat{t}} = C - n_U^{\hat{s}}(s_U - \hat{s})\alpha - (\hat{s} - \hat{t}) \cdot \alpha$.

But since the participation constraint for the perceived $\hat{t}$ types is not binding, an entrant may have an incentive to steal away customers by offering $F^{\hat{t}} = C - n_U^{\hat{s}}(s_U - \hat{s})\alpha - (\hat{s} - \hat{t})\cdot\alpha - \varepsilon$ where $\varepsilon > 0$. Since incentive compatibility will no longer hold, both perceived types will choose that tariff, and the following condition must therefore be held for the tariff to be profitable:

$$n^{\hat{s}}\left[F^{\hat{t}} + (\hat{s} - \hat{t})\alpha + n_U^{\hat{s}}(s_U - \hat{s})\alpha - n_O^{\hat{s}}(\hat{s} - \hat{t})\alpha - C\right] + n^{\hat{t}}\left[F^{\hat{t}} + n_U^{\hat{t}}(t_U - \hat{t})\alpha - C\right] \geq 0$$

Substituting $F^{\hat{t}} = C - n_U^{\hat{s}}(s_U - \hat{s})\alpha - (\hat{s} - \hat{t})\cdot\alpha - \varepsilon$ for the limiting case where $\varepsilon = 0$ is strictly satisfied, profitability is ensured if:

$$n^{\hat{s}}\left[-n_O^{\hat{s}}(\hat{s} - \hat{t})\right]\alpha + n^{\hat{t}}\left[n_U^{\hat{t}}(t_U - \hat{t}) - n_U^{\hat{s}}(s_U - \hat{s}) - (\hat{s} - \hat{t})\right]\alpha > 0$$

$$\Leftrightarrow \frac{n^{\hat{s}}}{n^{\hat{t}}} < \frac{n_U^{\hat{t}}(t_U - \hat{t}) - n_U^{\hat{s}}(s_U - \hat{s}) - (\hat{s} - \hat{t})}{n_O^{\hat{s}}(\hat{s} - \hat{t})}$$

$$\Leftrightarrow \frac{n_U^{\hat{t}}(t_U - \hat{t}) - n_U^{\hat{s}}(s_U - \hat{s})}{\hat{s} - \hat{t}} > 1 + \frac{n^{\hat{s}}}{n^{\hat{t}}} n_O^{\hat{s}}.$$

Therefore, a separating equilibrium only exists when $\frac{n_U^{\hat{t}}(t_U - \hat{t}) - n_U^{\hat{s}}(s_U - \hat{s})}{\hat{s} - \hat{t}} \leq 1 + \frac{n^{\hat{s}}}{n^{\hat{t}}} n_O^{\hat{s}}$, i.e. when the proportion of perceived $\hat{s}$ types (relative to the proportion of perceived $\hat{t}$ types) is sufficiently small, and when the proportion of perceived $\hat{s}$ types who are also overestimators is sufficiently large. When this condition fails, a pooling response is profitable and a separating equilibrium does not exist.□

Appendix B

This Appendix contains the proof of Proposition 3.

Proof of Proposition 3: Towards a contradiction, suppose that there is such an equilibrium in which at least one of the equilibrium tariffs doesn't take this form. Specifically, suppose that (i) for some $i \leq \hat{s}$, $p_i^{\hat{s}} > 0$ or (ii) for some $i > \hat{s}$, $p_i^{\hat{s}} < \alpha$. By assumption, the participation constraints hold and the incentive constraints strictly hold. We show that entry to steal away perceived $\hat{s}$ types from incumbent carriers is profitable under these conditions, which contradicts the assumption that we have an equilibrium.

First suppose that (i) holds. Notice that Assumption 1 then implies that we must have $p_i^{\hat{s}} > 0$ for some $s_O \leq i \leq \hat{s}$. But then an entrant could steal perceived $\hat{s}$ types from incumbent carriers by lowering all the positive $p_i^{\hat{s}}$s in the range $i \leq \hat{s}$ by $\Delta p_i^{\hat{s}}$ and increasing $F^{\hat{s}}$ by $\sum_{i=0}^{\hat{s}} \Delta p_i^{\hat{s}} - \varepsilon$ so that $\hat{s}$ types are (nearly) indifferent. Since ε can be set to be arbitrarily small, the entrant ends up earning $\left| n_O^{\hat{s}} \sum_{s_O}^{\hat{s}} \Delta p_i^{\hat{s}} \right|$ more profits than the incumbent carrier, an amount equal to the additional amount the overestimators end up paying in fixed fees, while falsely anticipating that they will benefit from the lower prices on minutes in the range $s_O \leq i \leq \hat{s}$. Since, by assumption, the incumbent earned non-negative profits (as the participation constraint must hold given our assumption of equilibrium), the entrant's profits from the perceived $\hat{s}$ types will be positive. Further, since, by assumption, the old tariffs satisfy the incentive compatibility constraints strictly (so that other perceived types strictly prefer tariffs designed for them to the incumbent's $\hat{s}$ tariff), we can set ε small enough to ensure that no other perceived type will want to choose the tariff of the entrant since they view the entrant's tariff as yielding at most a price reduction of ε relative to the incumbent's $\hat{s}$ tariff. Therefore, entry is profitable, which contradicts the assumption that (i) can hold in a separating equilibrium.

Likewise, if (ii) holds, there cannot be a separating equilibrium since an entrant could attract away perceived $\hat{s}$ types from incumbent providers by increasing $p_i^{\hat{s}}$s in the range $i > \hat{s}$ by $\Delta p_i^{\hat{s}}$ and lowering $F^{\hat{s}}$ by an arbitrarily small increment ε. Since perceived $\hat{s}$ types are insensitive to prices in the range $i > \hat{s}$, they will strictly prefer the new tariff, regarding it as increasing their utility by ε. Lower perceived types will perceive that the new tariff yields at most a price reduction of ε relative to the incumbent's $\hat{s}$ tariff, while higher perceived types will regard it as more expensive, and so, once again, with ε sufficiently small, no other perceived types will find it optimal to switch to the new tariff. Thus, the entrant's profits will exceed those of the incumbent provider by $n_U^{\hat{s}} \sum_{\hat{s}}^{s_U} \Delta p_i^{\hat{s}}$, the additional amount the underestimators end up paying. Since, by assumption, the incumbent provider earned non-negative profits, the entrant's tariff must be profitable. □