EVA ASCARZA, ANJA LAMBRECHT, and NAUFEL VILCASSIM*

In many service industries, firms introduce three-part tariffs to replace or complement existing two-part tariffs. In contrast with two-part tariffs, three-part tariffs offer allowances, or "free" units of the service. Behavioral research suggests that the attributes of a pricing plan may affect behavior beyond their direct cost implications. Evidence suggests that customers value free units above and beyond what might be expected from the change in their budget constraint. Nonlinear pricing research, however, has not considered such an effect. The authors examine a market in which three-part tariffs were introduced for the first time. They analyze tariff choice and usage behavior for customers who switch from two-part to three-part tariffs. The findings show that switchers significantly "overuse" in comparison with their prior two-part tariff usage. That is, they attain a level of consumption that cannot be explained by a shift in the budget constraint. The authors estimate a discrete/continuous model of tariff choice and usage that accounts for the valuation of free units. The results show that the majority of three-part-tariff users value minutes under a three-part tariff more than they do under a two-part tariff. The authors derive recommendations for how the provider can exploit these insights to further increase revenues.

Keywords: pricing, nonlinear pricing, discrete/continuous choice model, three-part tariffs, uncertainty, learning, free products

When Talk Is "Free": The Effect of Tariff Structure on Usage Under Two- and Three-Part Tariffs

Firms in various sectors are complementing or replacing their two-part tariffs with three-part tariffs. For example, telecommunications providers offer plans with free minutes instead of charging for every minute, and banks offer a set number of free check-writing privileges instead of billing per check. Two-part-tariff customers pay both a regular, often monthly, access price and a usage price for every unit of consumption, whereas under a three-part tariff, a usage price applies only to consumption in excess of the usage allowance. Within the allowance, there is no usage charge; usage is "free."

Recent evidence indicates that consumers generally respond to free products or services differently than how they would respond to the same good if the firm charged for it. Specifically, when evaluating free products or services, consumers do not simply subtract costs from benefits but rather perceive the benefits associated with free products as higher than they would otherwise. This perception leads to increased demand (Shampanier, Mazar, and Ariely 2007), which has important implications for firms. For example, there is evidence that indicates that when AOL replaced its pay-per-use plans for dial-up Internet access with flat-rate plans, demand at a zero usage price was far greater than AOL had forecasted from the income effect of the price change alone (Cnet.com 1996).[1] Behavioral research suggests that this increased demand can be attributed to a positive affective response to the offer of a zero usage price—a "free" component of the tariff. This affective response

*Eva Ascarza is Assistant Professor of Marketing, Columbia Business School, Columbia University (e-mail: ascarza@columbia.edu). Anja Lambrecht is Assistant Professor in Marketing (e-mail: alambrecht@london.edu), and Naufel Vilcassim is Professor of Marketing (e-mail: nvilcassim@london.edu), London Business School. The authors thank the London Business School Centre for Marketing for financial support. They also thank Rajesh Chandy, Bruce Hardie, Kamel Jedidi, Carl Mela, Oded Netzer, Catarina Sismeiro, and Gonca Soysal, as well as participants at the UTD FORMS; the 2009 Marketing Science Conference; the 8th Triennial Invitational Choice Symposium; the 2011 Four School Colloquium; The Psychology and Economics of Scarce Attention conference; and seminars at INSEAD, Erasmus University, HKUST, Northwestern University, and University of Rochester for their helpful comments. Sunil Gupta served as associate editor for this article.

[1]As a result, AOL needed to manage significant congestion and dissatisfied customers.

ISSN: 0022-2437 (print), 1547-7193 (electronic)

Journal of Marketing Research
Vol. XLIX (December 2012), 882–899

increases usage beyond what might be predicted from the change of the budget constraint alone (Shampanier, Mazar, and Ariely 2007).

Importantly, research demonstrates that this positive affective response may likewise increase the valuation of other, even unrelated products. For example, Isen et al. (1978) show that people who obtain a small free gift (notepad or nail clipper) subsequently evaluate the performance of unrelated products (their own car and television) significantly higher than those who do not obtain a free gift. Similarly, field data show that coupons can increase purchases beyond the expected income effect for products other than those for which the coupon was issued (Heilman, Nakamoto, and Rao 2002).[2] These insights go against standard economic theory that assumes that a change in price will affect demand only through a change in the budget constraint.

In this research, we examine how consumer demand changes when consumers switch from a two-part to a three-part tariff. We build on Shampanier, Mazar, and Ariely's (2007) work and argue that the free component of a three-part tariff leads to a positive affective response. This positive affective response increases the valuation of the service as such. Building on research that illustrates persistence of demand effects from positive affect across products, we argue that the positive affective response to a tariff with free units persists even after consumers have exceeded their usage allowance. As a result, we propose that when consumers switch from two-part to three-part tariffs, demand should increase beyond what would be predicted from the change in the budget constraint alone. This demand effect should hold even for usage beyond the allowance.

We use data on tariff choice and usage of customers of a mobile phone company. An important feature of the data is the introduction of three-part tariffs in addition to the existing two-part tariffs during our observation period, which enables us to observe the same set of customers under different pricing regimes: (1) when only two-part tariffs were available and (2) when customers were able to switch to three-part tariffs. In this market, three-part tariffs were largely designed to increase customer acquisition. Thus, their addition to the choice set for existing customers is close to a natural experiment.

An initial exploration of the data shows that customers who switched to a three-part tariff significantly "overused" after switching: Their level of consumption cannot easily be explained by the change in the budget constraint only or by other plausible alternative explanations. To disentangle the effect of free minutes on customers' valuation of the service from the change in the budget constraint arising from the new pricing structure and from preferences at tariff choice, we jointly estimate each customer's tariff choice and usage decision, conditional on the chosen tariff. In the utility function, we explicitly allow for greater demand on three-part tariffs. Because three-part tariffs are new to this market, customers are likely to be initially unaware that free minutes might affect their demand beyond the budget constraint. Therefore, we allow for the possibility that customers learn about their three-part tariff usage, that is, that they learn about their valuation of free minutes.

Our results indicate that with three-part tariffs, 83.9% of customers use more than might be expected based on their previous usage. We interpret this to mean that these customers have a greater valuation of the service than under the two-part tariff. This effect increases the provider's revenue from three-part-tariff customers by 19.7%. We find that by reducing the fee charged for switching between tariffs, the provider could increase total revenues by 3.9%, and even more if it discontinued the option to switch to another two-part tariff. In both instances, customers' greater usage on three-part tariffs is crucial to any revenue increase.

Our findings are managerially relevant from three perspectives. First, projecting customer usage solely on the basis of observed usage under existing two-part tariffs could lead firms to incorrectly determine the optimal tariff structure and prices, with potentially serious consequences to their profits. Our analysis shows that firms may significantly underestimate the revenue effect from introducing three-part tariffs if they do not sufficiently account for the effect on preferences arising from the change in tariff structure. Second, our findings show that ignoring the effect of a greater valuation for free minutes underestimates three-part-tariff usage by 14.9%. This result implies that when changing their tariff structure, firms might also need to adjust their service capacity. Third, and more broadly, our results show that the attributes of a pricing plan not only change its monetary value but also affect the perceived characteristics of the service.

Our findings add to the nonlinear pricing literature that recognizes behavioral preferences at tariff choice: Customers choose flat-rate or three-part tariffs with large allowances even when these entail a greater bill than tariffs with lower allowances (DellaVigna and Malmendier 2006; Lambrecht, Seim, and Skiera 2007; Lambrecht and Skiera 2006; Nunes 2000). Our work confirms that such deviations from standard economic theory are not limited to the choice of a product or service but also affect its usage. Our work likewise complements recent research on choice and consumption under three-part tariffs (Bagh and Bhargava 2008; Grubb 2009; Grubb and Osborne 2011; Iyengar, Ansari, and Gupta 2007; Jensen 2006), which thus far has abstracted from potential effects of the tariff structure on usage. As an exception, Iyengar et al. (2011) explore how tariff structure affects usage on two-part versus pay-per-use tariffs. They find that customers' marginal utility of consumption is lower under a two-part tariff than under a pay-per-use tariff.

More broadly, our work contributes to research that explores behavioral effects of pricing. This includes the insight that attributes of a price or a pricing plan can affect behavior beyond their direct cost implications (Bertini and Wathieu 2008), systematic effects of price endings on consumers' purchase decisions (Anderson and Simester 2003; Thomas and Morwitz 2005), or valuation of discounts that goes beyond the change in prices (Darke and Dahl 2003).

[2]The authors attribute this to elevated mood and additionally propose that a psychological income effect from receiving an unexpected coupon could lead consumers to spend more than the windfall gain from the coupon. They do not test which of the two explanations is more relevant. Golden and Zimmer (1986), Sherman and Smith (1987), and Donovan et al. (1994) provide further evidence that positive affect increases product demand, even for products that do not directly stimulate the positive affect.

Next, we present the data, provide evidence that customers overuse under three-part tariffs, and discuss possible explanations. We then develop a joint model of tariff choice and usage that allows for a greater valuation of the service under a three-part tariff, resulting in greater usage. We present the results of the estimation and our counterfactual analyses and conclude with a summary of findings and implications of our work.

DATA

The data include a random sample of 5831 individual customers (i.e., noncorporate) of a South Asian mobile telephone provider. We observed customers for up to 12 months, beginning in May 2006. The focal firm had 16% of the installed base. Mobile phone service penetration was 35% in May 2007. On average, customers had been with the firm for 23.5 months before the start of our observation period. The data contain information on the tariff chosen and monthly usage of outgoing calls. During the observation period, 3.7% of customers left the firm. In total, the data include 69,878 monthly usage and tariff choice observations. During the first 3 months, customers were offered a choice of two-part tariffs (Tariff_2_1 to Tariff_2_4 in Table 1; we refer to currency as MU for “monetary units”). For each two-part tariff, the provider charges four different per-minute prices, all greater than zero, depending on the time of day and the call destination. The data include the total number of minutes used per month, but not by time of day or destination. In addition, the firm provided us with the number of minutes used across all customers in each tariff by time of day and call destination. As a result, we use the weighted average of usage prices per tariff as a measure of the usage price.

On average, a customer uses 297 minutes a month and has a bill of MU 17.15 (Table 1). Customers can check their usage and bill by text message, by telephone, or through the Internet. They are free to leave the provider at any time (there are no contractual obligations) or to switch to another tariff of the same provider. Customers can switch tariffs by calling the firm’s customer service center, by visiting one of the firm’s retail outlets, or through an authorized agent. The provider charges MU 10 for switching to another tariff. This fee is higher than the two-part tariff access prices and represents a significant expenditure in this emerging market, in which customers are cash flow constrained.

We examine whether customers choose the *ex post* cost-minimizing two-part tariff on the basis of the average and standard deviation of their usage in the first three months of the data. We find that only 10.9% of customers would have paid less on a different two-part tariff (see the Web Appendix at www.marketingpower.com/jmr_webappendix).

Three months after the start of our observation period, the company added three three-part tariffs to the existing two-part tariffs (Tariff_3_1 to Tariff_3_3 in Table 1). Under a three-part tariff, the marginal price is zero for usage within the allowance. The provider charges a single price for usage above the allowance. The new tariffs were heavily advertised in print and on television. The provider introduced three-part tariffs to differentiate its offerings from those of its competitors and to increase customer acquisition, not as recognition of limitations of two-part tariffs in sorting customers (Jensen 2006; Wilson 1993). Thus, for existing customers, the introduction of three-part tariffs was close to a natural experiment. The firm was unaware that a change in the tariff structure could potentially change demand beyond what it would expect from the change in the budget constraint.

Our panel covers customers who were subscribers of the firm before the three-part tariffs were introduced but not newly acquired customers. Thus, we do not address market expansion effects. The two competitors offered similar two-part tariffs as those offered by the focal firm, but no three-part tariffs.

DESCRIPTIVE ANALYSES

Switching from Two- to Three-Part Tariffs

Although the focus of this research is on three-part tariff usage, we also provide an overview of tariff-switching behavior. In the data, 13.7% of customers switched between tariffs: 5.8% switched to a two-part tariff, and 7.9% switched to one of the three-part tariffs, resulting in 2357 three-part-tariff observations. A key strength of the data is that we observe customers under a two-part tariff before the introduction of the three-part tariffs and, subsequently, several months of usage behavior under a three-part tariff. Thus, the data provide us with a high number of observations per tariff, which enables us to identify the effect of tariff structure on three-part-tariff usage within individual customers.

As Table 2 illustrates, customers mostly switched from any of the two-part tariffs to Tariff_3_1. In determining whether customers’ decisions to switch to a three-part tariff could have been predicted from their past consumption, we find that of all customers who would

Table 1
TARIFF CHARACTERISTICS

Package	Access Price (MU)	Allowance (Minutes)	Usage Price (MU)	Average Usage (Minutes)	Average Bill (MU)	Observations (Number)	Customers April 2007 (Number)
Tariff_2_1	1	0	.079	157.17	12.09	9786	766
Tariff_2_2	2	0	.056	220.45	13.19	9512	752
Tariff_2_3	3	0	.055	267.97	15.93	32,166	2541
Tariff_2_4	5	0	.042	470.20	24.52	16,057	1346
Tariff_3_1	10	200	.050	324.97	16.95	1896	321
Tariff_3_2	20	500	.050	730.39	33.25	231	47
Tariff_3_3	30	900	.050	1142.86	48.75	230	58

Notes: MU refers to the local currency.

Table 2
TARIFF-SWITCHING MATRIX

	Switching to . . .						
	Tariff_2_1	*Tariff_2_2*	*Tariff_2_3*	*Tariff_2_4*	*Tariff_3_1*	*Tariff_3_2*	*Tariff_3_3*
Switching from . . .							
Tariff_2_1	87.3%	.0%	3.2%	2.8%	6.5%	.1%	.1%
Tariff_2_2	.0%	91.1%	3.2%	1.8%	3.8%	.1%	.0%
Tariff_2_3	.1%	.1%	85.6%	6.0%	7.1%	.6%	.6%
Tariff_2_4	.2%	.0%	1.3%	89.0%	4.4%	2.2%	2.8%

Notes: For customers who switched more than once, we only consider the first switch.

have benefited from switching, 9.0% did switch, whereas among customers who would not have benefited, only 4.7% switched (for details, see the Web Appendix at www.marketingpower.com/jmr_webappendix). This is consistent with the generally low switching rates in telecommunications services and likely a result of the high switching fee, which may have deterred customers from switching.[3]

To assess whether the high switching fee might have deterred customers from switching, we replicate the analysis presented previously, now considering a switch beneficial only when savings in the first month would compensate for the switching fee. In this case, a greater proportion of customers who would benefit from switching to a three-part tariff did switch (13.2% vs. 9.0%). These results provide some indication that whereas customers behave optimally when choosing tariffs, the switching fee could have prevented customers from switching to the three-part tariffs. Our econometric model takes this into account.

Finally, we examine subsequent tariff choices among customers who switched to a three-part tariff. We find that 11.7% later switched to another tariff: 8.3% switched back to a two-part tariff, and 3.4% switched to a different three-part tariff. In contrast, none of the customers who switched to a two-part tariff in the first place later switched again. The difference in subsequent switching behavior between customers who switched to two-part tariffs and those who switched to three-part tariffs suggests that they had difficulty predicting their own usage under three-part tariffs but not under two-part tariffs. Furthermore, customers typically do not switch immediately after their first three-part tariff choice. Rather, they spend, on average, 2.7 months under a three-part tariff before switching again. This indicates that customers learn about their usage under three-part tariffs before adjusting their tariff choice. Our econometric model incorporates this learning process.

[3]The share of customers who switch to a three-part tariff is significantly lower than the share of customers who initially choose a three-part tariff in the market that Grubb (2009) studies. Multiple factors likely contribute to this difference. First, our customers are required to *switch* to a three-part tariff, so switching costs may deter customers from switching. Second, our customers had a long-time experience with the service, so overconfidence, which Grubb identifies as an important factor in choice, may have been less important. Third, in our market, three-part tariffs are new, and customers might be reluctant to try new plans.

Usage Under Three-Part Tariffs

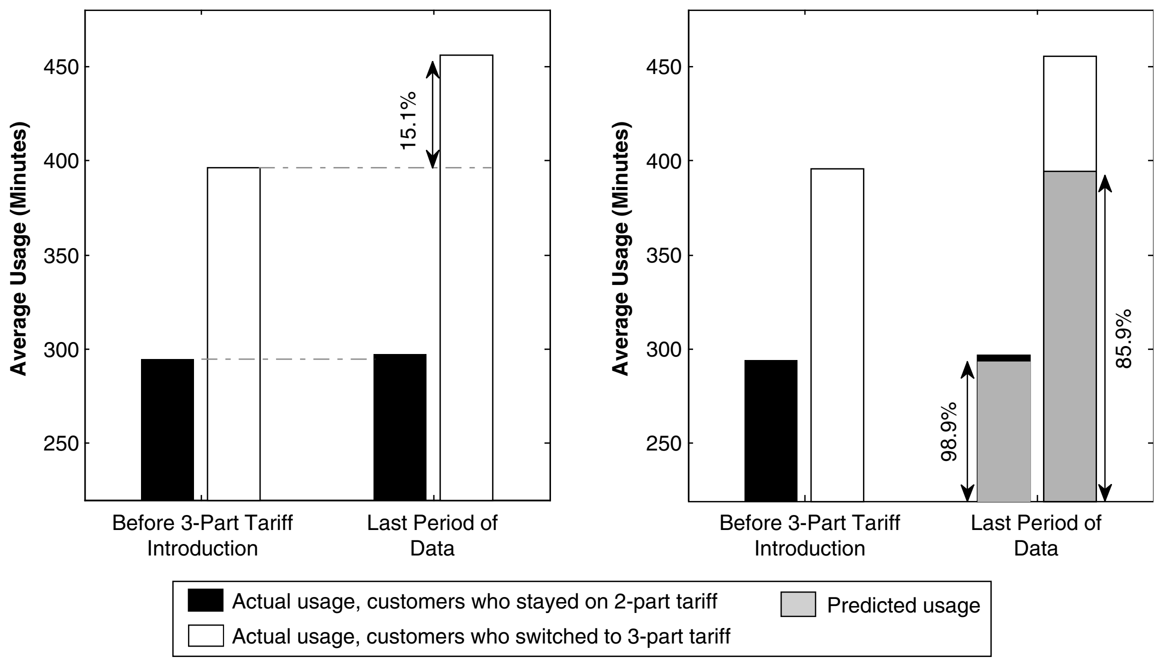
We turn to customers' usage behavior after the introduction of the three-part tariffs. A unique aspect of the data is that we observe all customers under two-part tariffs before the three-part tariffs were introduced. Moreover, not all customers switched to three-part tariffs after the introduction. Therefore, we can analyze whether switching to a tariff with free minutes affects consumption.

We compare the average monthly usage before the three-part tariff introduction to usage in the last month of the data. Customers who switched to a three-part tariff increased their usage by 15.1%, while customers who remained under a two-part tariff increased their usage by .9% (see the left-hand side of Figure 1). These results indicate a change in usage of three-part-tariff customers that goes beyond a general time trend. This pattern is persistent over time (see the Web Appendix at www.marketingpower.com/jmr_webappendix) and is consistent across all three-part tariffs: Under Tariff_3_1, usage increased by an average of 15.5%; under Tariff_3_2, usage increased by 16.2%; and under Tariff_3_3, usage increased by 19.2%. Furthermore, actual usage often significantly exceeds the allowance: 72% of three-part tariff observations exceeded the usage allowance by, on average, 88.4%.

At first glance, this increase in usage could be due to the change in tariffs' marginal prices; a utility-maximizing customer may use more under a three-part tariff simply because of the change in the budget constraint. To explore whether the change in the budget constraint can explain the increased three-part tariff usage, we estimate a linear demand model of monthly usage. Using the observations before the three-part tariff introduction, we estimate an individual-level demand intercept and a price coefficient for monthly consumption. We then predict usage conditional on the chosen tariff in the last month in the data and compare these predictions with the actual behavior. This approach accounts for changes in the budget constraint because the prediction is based on prices and allowances of the tariff that each customer faces in the last month (for details of the analysis and results, see the Web Appendix at www.marketingpower.com/jmr_webappendix).

The right-hand side of Figure 1 shows the actual and predicted average usage in the last period. For customers who stayed on a two-part tariff, predicted usage is 98.9% of observed usage. However, for customers who switched to a three-part tariff, predicted usage is only 85.9% of actual usage. In other words, the demand model, which accounts for the change in the budget constraint, predicts two-part tariff usage accurately but underestimates three-part tariff

Figure 1
USAGE BEFORE AND AFTER THE INTRODUCTION OF THE THREE-PART TARIFFS



usage by almost 15%. This result is consistent across all three-part tariffs and is independent of the month in which customers switch to a three-part tariff. This indicates that the effect persists over time (see the Web Appendix at www.marketingpower.com/jmr_webappendix).

Moreover, this result still holds when we relax model assumptions that could lead us to systematically overpredict usage. First, if demand was not linear in price but rather convex, imposing a linear demand specification could lead to an underestimation of usage in regions in which price is very low (or zero), thus systematically underestimating three-part tariff usage. We relax the linearity assumption and find that nonlinear utility specifications lead to the same results qualitatively (see the Web Appendix at www.marketingpower.com/jmr_webappendix). Second, it might be that customers who switched to a three-part tariff had different usage price sensitivities than customers who did not switch to a three-part tariff. Thus, the assumption of homogeneous usage price sensitivity could lead us to under- or overestimate their price response and to underpredict three-part-tariff usage. Similar to existing research on multipart tariffs (Iyengar, Ansari, and Gupta 2007; Lambrecht, Seim, and Skiera 2007; Narayanan, Chintagunta, and Miravete 2007), we cannot estimate an individual-level price coefficient because of a lack of individual-level price variation in the data, though we can check whether the homogeneity assumption is driving our results. We estimate a demand specification in which customers who switched to a three-part tariff have a different set of parameters than customers who did not switch to a three-part tariff. Again, we obtain qualitatively similar results (see the Web Appendix at www.marketingpower.com/jmr_webappendix).

In summary, the large increase in usage after switching to a tariff with free minutes cannot be explained by the difference in the budget constraint. It is consistent, however, with previous research that shows that free goods lead to a positive affective response, increasing the valuation of the product or service and, thus, its demand (Shampanier, Mazar, and Ariely 2007). Furthermore, this positive response to free minutes can increase customers' valuation of other goods (Heilman, Nakamoto, and Rao 2002; Isen et al. 1978), thus affecting the entire consumption experience. This finding is supported by the data, in that 72% of three-part tariff customers consume, on average, 88.4% above the tariff allowance. We therefore conclude that customers who switch to a three-part tariff may assign greater value to minutes on that tariff, including those above the allowance.

Alternative Explanations

We acknowledge that other explanations could plausibly lead to a similar pattern of usage. We examine these in detail in the following subsections.

Risk aversion. Risk aversion might lead customers to choose tariffs with large allowances—that is, tariffs for which greater usage may be optimal—which in turn may result in greater consumption. However, risk-averse customers would also be more likely to keep their usage at or slightly below the allowance, which is not consistent with the data: 72% of three-part-tariff observations exceed the usage allowance by, on average, 88.4%, and 92% of three-part-tariff customers exceed their usage allowance at least once. Furthermore, we observe that for 61% of the three-part-tariff customers, usage is above their two-part-tariff satiation level. Risk aversion cannot explain such a change to the satiation level.

Regret avoidance. Under a three-part tariff, customers might want to "get their money's worth"—that is, entirely use their allowance even if this exceeds the optimal usage of a rational utility-maximizing consumer. However, this explanation is not consistent with the data, because most three-part-tariff customers exceed the allowance by a large amount (on average, 88.4%).

Within-day price variation. Under the two-part tariffs, usage prices vary with the time of day (peak and off-peak rates) and call destination (within and out of network). We do not have access to usage data by call type or records of individual calls, and so our previous analysis relied on the average usage price provided by the provider. However, in theory, all observed two-part-tariff calls could have been made during peak hours when a greater usage price was charged, in which case the previous analysis may have overestimated the usage price sensitivity and thus underpredicted three-part-tariff usage.

As a robustness check, we reestimate the usage model discussed in the section "Usage Under Three-Part Tarifs," this time assuming that all two-part-tariff calls were made at the highest marginal price (i.e., during peak hours out of network). In doing so, we possibly underestimate customers' price sensitivity and possibly overpredict, but not underpredict, three-part-tariff usage. Again, only 85.8% of usage of three-part-tariff customers can be explained by the shift in the budget constraint.

Self-selection. Customers may have switched to a three-part tariff because they anticipated greater usage in future periods. Econometrically, self-selection should lead to a high correlation between factors that affect three-part-tariff choice, beyond expected savings, and factors that affect three-part-tariff usage, beyond what would be predicted from a consumer's demand parameters from two-part-tariff usage. Investigating whether such a correlation exists is not feasible in a purely descriptive way or by examining demand alone. This can only be tested in a joint model of usage and tariff choice. We therefore check for such a correlation when estimating our full discrete/continuous choice model in the "Model Development and Estimation" section.

Alternatively, self-selection could arise if usage followed an autoregressive process. If this were the case, customers who switched to a three-part tariff because they had a positive usage shock in the last period should also be more likely to increase their usage in future periods. We rule out this possibility in two ways. First, we check for serial correlation among monthly usage shocks. Second, we investigate whether customers who have (higher) positive usage shocks are indeed more likely to switch to a three-part tariff. Neither of the analyses supports the notion that an autoregressive process leads to self-selection in the data (for details, see the Web Appendix at www.marketingpower.com/jmr_webappendix).

Tariff-specific services or marketing activities. Usage behavior could change if the three-part tariff offered other services unavailable on the two-part tariff. For example, if text messages were offered for free on the two-part but not on the three-part tariff, three-part-tariff customers could substitute calls for text messages. In our context, text messages are rarely used because the language of conversation is different from the language script on the handset and because neither calling plan is connected with other services that would lead to such an increase in usage. Similarly, marketing activities that aim to switch customers to three-part tariffs and simultaneously increase usage on these tariffs could explain high three-part-tariff usage (e.g., advertising or the introduction of handsets with new features that stimulate usage and were limited to three-part tariff users). We know from the provider that no such activities were undertaken in this market.

Awareness of usage. Customers might overuse because they are unaware of their usage level. Several indications, however, suggest that overusage is not due to a lack of awareness. First, customers had used the service for, on average, 23.5 months before the start of our data, and the firm provides many possibilities for monitoring usage. Thus, it seems unlikely that three-part-tariff switchers were completely unaware of their usage level. Even if the availability of free minutes initially resulted in unintended overusage, it seems reasonable to assume that customers would adjust their usage within the next months, a pattern not observed in the data (see the Web Appendix at www.marketingpower.com/jmr_webappendix). Second, the data show a mass point of usage observations at the allowance, providing further support that customers track their usage (see the Web Appendix). Third, if overusage were largely due to a lack of awareness, customers should have stopped using at their satiation point. However, 61.0% of customers used more than the satiation level predicted from their two-part tariff usage. These indications provide further support that a lack of awareness is likely not the main reason behind the usage increase but rather that the satiation level changed. Still, we acknowledge that because the data do not include information on when or how often customers check their usage levels, we cannot fully rule out that a lack of awareness contributed to the observed increase in usage.

Intramonth usage uncertainty. Under a three-part tariff, customers' decisions to make a call depend on their expected valuation of future calls during that month. For example, customers should prefer to use the entire allowance today even if such calls are of low value to them, as long as the expected value of tomorrow's calls is of even lower value. However, if tomorrow's calls are unexpectedly of greater value than the three-part tariff's usage price, customers will still make these calls, possibly resulting in overusage. As a consequence, intramonth usage uncertainty could lead to the overusage we observe in the data. To investigate this possibility, we analyze whether customers with greater intramonth usage uncertainty also have greater overusage after they switch to a three-part tariff.

Because the data are limited to aggregate monthly usage and do not contain information on individual calls, we cannot measure the level of uncertainty in an individual month. However, the level of usage variation across months provides a measure of the degree of uncertainty customers face in their overall usage. We check whether the degree of usage variation before the three-part tariffs were introduced is correlated with overusage after switching. We measure overusage as the ratio of actual to predicted three-part-tariff usage. We find no correlation between the two measures (correlation = .080, p = .098). Alternatively, we measure overusage as the difference between actual and predicted three-part-tariff usage and correlate this with usage variation. Again, we find no correlation with the degree of usage variation (correlation = .038, $p = .432$). We conclude that the degree of intramonth uncertainty is unlikely to explain the increase in usage we observe in the data.

In conclusion, even if we cannot simultaneously rule out all alternative accounts, these accounts do not explain the full extent of the increase in usage we observe. Instead, the increase in usage seems consistent with the interpretation that customers have a greater valuation when a tariff offers free minutes. There are, however, difficulties in precisely pinning down the effect of free minutes in the descriptive model we have presented so far. First, we cannot identify a systematic increase in usage separately from random usage shocks and time-varying demand shifters. Thus, we are unable to precisely estimate by how much customers' demand changes as a result of the change in tariff structure. This is important for firms that want to accurately forecast their revenues and capacity needs. Second, modeling usage independently of tariff choice could lead to selection bias (Dubin and McFadden 1984) and thus could provide biased estimates of demand parameters. Because the same set of parameters affects choice and usage, jointly modeling both should yield unbiased and more reliable parameter estimates. Third, to conduct policy simulations, we need a consistent set of parameters that fully describes customers' usage and tariff choice decisions and thus need to incorporate factors that affect tariff choice alone (e.g., switching costs, tariff preferences), which have been overlooked so far. Finally, jointly modeling choice and usage allows us to conclusively rule out self-selection as an alternative explanation. Thus, we next build a joint model of usage and tariff choice that enables us to estimate the effect of free minutes and to make inferences about customers' behavior under three-part tariffs.

MODEL DEVELOPMENT AND ESTIMATION

Model Setup

At the beginning of each month, customers choose one of the available tariffs or leave the provider depending on their expected usage for that month. Conditional on their tariff choice, customers next determine their monthly usage from the utility they derive from the service. We capture this behavior with a discrete/continuous choice model (Dubin and McFadden 1984; Hanemann 1984).

Building on our descriptive analyses, we incorporate a factor in the utility function to capture the possibility of greater utility if the tariff provides free minutes. This valuation of free minutes affects three-part-tariff usage directly, and choice indirectly, through expected usage. Given that three-part tariffs were completely new to this market, we assume that customers are initially unaware of the possible effect of free minutes on their consumption, beyond what the change in their budget constraint would account for.[4] Only when customers first experience a three-part tariff do they become aware of their valuation of free minutes. Consistent with tariff-switching patterns in the data, we assume that customers learn about their three-part tariff usage, and thus their value of consuming under a three-part tariff, over time.

Mobile phone penetration is increasing and the firm's customers have used the service for a long time, so we assume that customers who leave the provider switch to a competitor rather than disconnecting the service. The competitors' tariff offerings did not change during the observation period, so explicitly accounting for their tariffs in estimation would not differ greatly from normalizing the price of the outside option to one, which we do for simplicity.

We model customers' tariff choice on the basis of the expected utility in the next period only (Iyengar, Ansari, and Gupta 2007; Lambrecht, Seim, and Skiera 2007; Narayanan, Chintagunta, and Miravete 2007). Alternatively, we could assume that customers trade off current-period switching costs against all future utility gains (Goettler and Clay 2011). However, accounting for future periods would require assumptions about the discount rate and complicate the estimation. Note that if consumers were forward looking, a static model might potentially overestimate their switching costs but would not bias our main parameter of interest, which captures the additional value of a three-part tariff.

Utility Function

Customers choose among a set of J tariffs. Each tariff j includes a monthly access price, denoted by F_j; an allowance measured in minutes of usage, $\tilde{q}_j$; and a marginal price, p_j, charged for each minute that exceeds the tariff's monthly allowance. A higher access price is associated with a higher usage allowance, so that $F_j < F_{j'}$ if $\tilde{q}_j < \tilde{q}_{j'}$. A two-part tariff is similar to a three-part tariff, but its allowance, $\tilde{q}_j$, is, by definition, set to zero. For each two-part tariff, a higher access price is associated with a lower usage price, so that $p_j > p_{j'}$ if $F_j < F_{j'}$.

We assume that customer i at time t chooses the optimal tariff j and consumption levels for minutes of calls, q^*_{ijt}, and the outside good, q^*_{i0t}, to maximize his or her utility subject to the budget constraint. We choose a quadratic utility function to allow for satiation (Iyengar, Ansari, and Gupta 2007; Lambrecht, Seim, and Skiera 2007). This is important because it reflects the behavior in the data in which some customers use less than the allowance, the maximum possible usage at a zero usage price. It also assumes that customers are risk averse. Because the utility function is linear in q_{i0t}, it does not capture nonlinearities in the outside good. Utility on tariff j is represented by

$$(1)\quad U_{ijt}(q_{ijt}, q_{i0t}) = c_i\left\{\frac{1}{b_i}\left[d_{ijt}q_{ijt} - \frac{(q^2_{ijt} + d^2_{ijt})}{2}\right] + q_{i0t}\right\} + \varsigma_{ijt},$$
$$b_i, c_i, d_{ijt} > 0,$$

where c_i represents the marginal utility of income and d_{ijt} is the satiation level (i.e., demand at a zero usage price). The demand slope, b_i, measures sensitivity to the usage price. The term ς_{ijt} captures observable and unobservable characteristics that affect tariff choice but not consumption. Customer i's budget constraint is given by

$$(2)\quad y_{it} = q_{i0t} + F_j + (q_{ijt} - \tilde{q}_j)I_{q_{ijt} \geq \tilde{q}_j}p_j,$$

[4]The firm did not expect any change in the usage from the introduction of three-part tariffs beyond what the change to the cost structure would entail.

where we normalize the price of the outside good to one. Under a three-part tariff, the usage price, p_j, is strictly positive only for $q_{ijt} > \tilde{q}_j$—instances that we capture with the indicator variable $I_{q_{ijt} \geq \tilde{q}_j}$ set to one if $q_{ijt} \geq \tilde{q}_j$ and zero if otherwise. Under two-part tariffs, $\tilde{q}_j$ is, by definition, set to zero; thus, $I_{q_{ijt} \geq \tilde{q}_j}$ is always one.

From Equations 1 and 2, we derive the customer's optimal two-part tariff usage as

$$(3) \qquad q^*_{ijt} = \begin{cases} 0 & \text{if } d_{ijt} \leq b_i p_j \\ d_{ijt} - b_i p_j & \text{if } d_{ijt} > b_i p_j \end{cases}$$

and under a three-part tariff as

$$(4) \qquad q^*_{ijt} = \begin{cases} d_{ijt} & \text{if } d_{ijt} < \tilde{q}_j \\ d_{ijt} - b_i p_j & \text{if } d_{ijt} - b_i p_j > \tilde{q}_j \\ \tilde{q}_j & \text{if } d_{ijt} - b_i p_j \leq \tilde{q}_j \leq d_{ijt}. \end{cases}$$

The first part of Equation 4 reflects consumption when usage is below the allowance. The second part determines usage when consumption exceeds the allowance and a strictly positive usage price, p_j, applies. The last part accounts for situations when the optimal usage would exceed the allowance of $\tilde{q}_j$ at a marginal price of zero but falls short of the allowance at the positive marginal price. Because the incremental value of usage beyond the allowance is not justified by the additional usage charges that accrue abruptly, q^*_{ijt} must be equal to $\tilde{q}_j$.

Substituting the optimal demand for the outside good and usage into the utility function yields the conditional indirect utility function under a two-part tariff

$$(5) \qquad V_{ijt}(y_{it}, p_j, F_j) = \begin{cases} c_i(y_{it} - F_j) + \varsigma_{ijt} & \text{if } q^*_{ijt} = 0 \\ c_i\left[y_{it} - F_j - \left(d_{ijt} - \frac{b_i p_j}{2}\right)p_j\right] + \varsigma_{ijt} & \text{if } q^*_{ijt} > 0 \end{cases}$$

and under a three-part tariff

$$(6) \qquad V_{ijt}(y_{it}, p_j, F_j) = \begin{cases} c_i(y_{it} - F_j) + \varsigma_{ijt} & \text{if } q^*_{ijt} \leq \tilde{q}_j \\ c_i\left[y_{it} - F_j + p_j\tilde{q}_j - \left(d_{ijt} - \frac{b_i p_j}{2}\right)p_j\right] + \varsigma_{ijt} & \text{if } q^*_{ijt} > \tilde{q}_j. \end{cases}$$

We decompose ς_{ijt} into three observed tariff preference shifters: (1) the cost of switching to another tariff of the same provider, (2) the cost of switching to the outside option (i.e., churn), and (3) the preference for choosing a three-part over a two-part tariff:

$$(7) \qquad \varsigma_{ijt} = \rho_1 SC^T \times I^T_j + \rho_2 I^P_j + \lambda_i I^{3\ pt}_j + \varepsilon_{ijt}.$$

The term SC^T reflects the provider's fee for switching to one of its own tariffs, ρ_1 reflects the sensitivity to this switching cost, and I^T_j is an indicator for switching to another tariff of the same provider. The term I^P_j is an indicator for switching to another provider, and ρ_2 captures the nonmonetary costs of switching to a provider other than the focal firm. The indicator $I^{3\ pt}_j$ is one only under a three-part tariff, so λ_i captures unobserved factors that affect three-part-tariff choice independent of usage expectations. It determines the individual-specific propensity to choose a three-part tariff. For customers who do not switch to three-part tariffs, it may, for example, capture a preference for a familiar tariff structure or aversion to a high access price. For switchers to a three-part tariff, it may capture a willingness to experiment with new tariffs that overrides the cost of the switching fee. Thus, it is a factor that is known to customers, even though it is unobserved to the researcher. Importantly, this preference is reflected in choice and explains switching to three-part tariffs, but not the changes in usage. We assume that λ_i is normally distributed across the population with mean and variance $(\mu_\lambda, \sigma^2_\lambda)$. The term ε_{ijt} is an unobserved preference shifter that the customer knows at the time of tariff choice. It is assumed to follow a Type 1 extreme value distribution.

Last, we specify the factors that determine the satiation level d_{ijt}. Our challenge is in modeling increased usage under a three-part versus a two-part tariff beyond what is due to the change in the budget constraint. Empirically, only a change to the demand intercept, d_{ijt}, can explain the observed increase in usage at any level of consumption. (A change in the demand slope cannot explain a change in usage when usage is below the allowance.) We specify δ_i as the additional value from the service under a three-part tariff. It is independent of the change of the budget constraint and measures a change in behavior conditional on tariff choice. Note that while the parameter λ_i captures unobserved factors that lead to the *choice* of a three-part tariff, δ_i captures how the valuation of free minutes changes customers' *usage* behavior.

Because three-part tariffs are new to this market, customers are not yet aware of their positive affective response to the free allowance and its subsequent effect on usage at their initial three-part tariff choice. This means that they do not yet know that access to free minutes may change their usage behavior. Therefore, δ_i does not affect the initial three-part-tariff choice and enters the demand intercept, d_{ijt}, only after a customer initially chooses a three-part tariff. Note that we would be unable to identify both δ_i and λ_i if the consumer was aware of δ_i at his or her initial three-part tariff choice.

We assume that the parameter δ_i is normally distributed across the population with mean and variance $(\mu_\delta, \sigma^2_\delta)$ and takes the same value for any three-part tariff. An individual-specific parameter η_i captures differences among customers' demand that are constant over time and known to the customers but unknown to the researcher. We assume that it is normally distributed with mean and variance $(\mu_\eta, \sigma^2_\eta)$.

It is possible that similar factors drive customers' choice of and usage under a three-part tariff. This would induce a correlation between the three-part-tariff choice preference, λ_i, and the valuation of the three-part tariff's allowance that affects usage, δ_i. Such a correlation could come from self-selection of customers who switch to a three-part tariff because they plan to use more, from tariff-specific marketing activities (e.g., handset subsidies), or from differences in tariff-specific services (e.g., included text messages). As we discussed previously, our conversations with the provider confirm that there are no such policies, and our analyses of the data provide no support for self-selection. Nevertheless, we revisit this possibility in the results section and examine the correlation between posterior estimates for parameters λ_i and δ_i.

We allow for uncertainty about usage at the time of tariff choice (Lambrecht, Seim, and Skiera 2007; Narayanan, Chintagunta, and Miravete 2007). We model an unobserved usage shock ϕ_{it} that reflects random usage variation. We observe in the data that individual-level usage variation is correlated with individual average consumption (correlation = .77, $p < .0001$); that is, heavy users have a higher variance of usage than light users. Thus, we assume a multiplicative usage shock, ϕ_{it}, which is gamma distributed with equal shape and scale parameter (e.g., with mean 1 and variance 1/r). At the moment of tariff choice, customers know this shock only in distribution. After the tariff choice but before making their usage decision, customers observe their usage shock and consume accordingly. Unobserved tariff-specific preferences, ε_{ijt}, drive tariff choice but do not affect the distribution of demand. We assume that the two sets of unobservables, ε_{ijt} and ϕ_{it}, are independent. Correlation could arise from user- and plan-specific advertising or from customer-specific promotions, but we know from the provider that such campaigns were not present.

To ensure a positive demand intercept, we specify d_{ijt} in exponential form as follows:

$$(8) \qquad d_{ijt} = \phi_{it} e^{h_t a_1 + \eta_i + \delta_i I_j^{3pt}},$$

where h_t is a dummy for holiday periods and a_1 is a parameter to be estimated. Note that though the variance of the usage shock is homogeneous across customers, the effect of the shock on usage is heterogeneous because of its multiplicative interaction with mean usage, $e^{h_t a_1 + \eta_i + \delta_i I_j^{3pt}}$.

Demand Uncertainty and Tariff Choice

A customer chooses the tariff that yields the highest expected indirect utility. Customers are experienced users of a two-part tariff: They know their two-part tariff usage preferences and the distribution of the usage shock, ϕ_{it}, but are uncertain about its exact realization. In other words, at tariff choice, customers do not know their exact usage on each tariff, q^*_{ijt}, because the usage shock ϕ_{it} has not yet been realized. However, they know their *expected* usage on each tariff because they know the distribution of the usage shock. We obtain the expected indirect utility of a two-part tariff by taking expectations over the unknown usage quantity in Equation 5—that is, over the distribution of the uncertain shock ϕ_{it}(see Goettler and Clay 2011; Iyengar, Ansari, and Gupta 2007; Lambrecht, Seim, and Skiera 2007):

$$(9) \qquad \begin{aligned} E[V_{ijt}] &= P(q^*_{ijt} = 0)E[V_{ijt} \mid q^*_{ijt} = 0] \\ &\quad + P(q^*_{ijt} > 0)E[V_{ijt} \mid q^*_{ijt} > 0] \\ &= P(q^*_{ijt} = 0_j)c_i(y_{it} - F_j) + P(q^*_{ijt} > 0) \\ &\quad \times c_i\left[y_{it} - F_j + \frac{1}{2}b_i p_j^2 \right. \\ &\quad \left. - E(\phi_{it} e^{h_t a_1 + \eta_i} \mid q^*_{ijt} > 0)p_j\right] + \varsigma_{ijt}. \end{aligned}$$

Similarly, we obtain the expected indirect utility of a three-part tariff by taking expectations of the three-part tariff optimal usage over the unknown quantity. Although customers are experienced users of two-part tariffs, they do not have past experience under three-part tariffs. As such, when they experience a three-part tariff, they observe a realization of $\phi_{it}e^{\delta_i}$ but cannot separate the effect of the usage shock, φ_{it}, from that of the new tariff structure. In other words, their three-part tariff choice is guided by uncertainty about the usage shock, φ_{it}, and the value of δ_i. Thus, we obtain the expected indirect utility of a three-part tariff by taking expectations of Equation 6 with respect to the distribution of $\phi_{it}e^{\delta_i}$:

$$(10) \qquad \begin{aligned} E[V_{ijt}] &= P(q^*_{ijt} \le \tilde{q}_j)E[V_{ijt} \mid q^*_{ijt} \le \tilde{q}_j] \\ &\quad + P(q^*_{ijt} > \tilde{q}_j)E[V_{ijt} \mid q^*_{ijt} > \tilde{q}_j] \\ &= P(q^*_{ijt} \le \tilde{q}_j)c_i(y_{it} - F_j) + P(q^*_{ijt} > \tilde{q}_j) \\ &\quad \times c_i\left[y_{it} - F_j + p_j\tilde{q}_j \frac{1}{2}b_i p_j^2 \right. \\ &\quad \left. - E(\phi_{it} e^{h_t a_1 + \eta_i + \delta_i} \mid q^*_{ijt} > \tilde{q}_j)p_j\right] + \varsigma_{ijt}. \end{aligned}$$

Note that the expected indirect utility of an *initial* three-part tariff choice is slightly different because, as we discussed previously, δ_i does not affect a customer's initial three-part tariff choice. Thus, when we compute the expected indirect utility of a three-part tariff for customers who have not yet experienced it, δ_i drops from Equation 10 and we take expectations over the shock ϕ_{it}. In the Appendix, we derive the closed-form expressions for the expected indirect utility in all three cases: a two-part tariff, an initial three-part tariff, and subsequent three-part-tariff choices.

Finally, if customers decide to leave the provider, their expected indirect utility is as follows:

$$(11) \qquad E[V_{i0t}] = c_i y_{it} + \varsigma_{i0t}.$$

In summary, customers evaluate the expected indirect utility of each available option and choose the option with the highest expected indirect utility. When evaluating a two-part tariff, customers know their preferences and the distribution of the shocks that affect future demand, but not the exact consumption next period. The same considerations affect their initial three-part tariff choice. When customers have experienced a three-part tariff, they evaluate subsequent three-part-tariff choices taking into account beliefs about their three-part tariff usage, that is, their beliefs about their own valuation of free minutes. We next explain how customers learn about that value over time.

Learning

At the start of our data period, customers had been with the provider for an average of 23.5 months, exclusively on two-part tariffs. Consistent with prior research that has found that customers learn about their usage within approximately 9 months (Iyengar, Ansari, and Gupta 2007), we assume that by the time we observe them, customers have already learned about their two-part-tariff usage. However, after they switch to a three-part tariff, usage behavior changes, and customers cannot easily infer their three-part-tariff usage from prior two-part-tariff consumption.

Two mechanisms could describe the process by which customers become aware of their three-part-tariff usage. Either they become instantly knowledgeable as they make their first usage decision under a three-part tariff, or they learn gradually as they observe their usage behavior. The data show that customers who switch from three-part tariffs do so after an average of 2.7 usage periods, which suggests that they learn about their three-part-tariff usage over time.

Therefore, we extend our model to accommodate customer learning. Specifically, we allow customers to learn about how their usage under a three-part tariff differs from their usage under a two-part tariff. This corresponds to learning about their preference for free minutes and, thus, about the parameter δ_i, which captures the differential usage.

We assume a Bayesian learning process (Erdem and Keane 1996) in which customers learn about their true value of δ_i using their own usage as the signal of preferences. Econometrically, customers know that the unknown usage quantity under a three-part tariff, $\phi_{it}e^{\delta_i}$, is gamma distributed with parameters (r, β_i), where $\beta_i = r/e^{\delta_i}$.[5] Given that customers know the value of r—because 1/r is the variance of the usage shock—learning about δ_i translates into finding the true value of the scale parameter β_i. Thus, after switching to a three-part tariff, customers form a belief about the true value of β_i (we denote the time-varying individual beliefs by $\tilde{\beta}_{it}$). Then, at the end of each period, customers observe their three-part-tariff usage and update their beliefs before making their next tariff choice. Figure 2 summarizes the decision process.

We assume that the initial beliefs are gamma distributed. This is a less restrictive specification than the commonly used normal distribution because the gamma distribution allows for nonsymmetry in the distribution of the unknown usage quantity. In addition, normally distributed beliefs would not be easily tractable in our multiplicative setting. Because the noise affecting the learning process (usage shock) is not normally but rather gamma distributed, normal beliefs would not be conjugate priors. Moreover, the gamma specification allows customers to learn at different rates in a relatively parsimonious model given that the variance of their beliefs, and thus their speed of learning, depends on their own signal (see Equation 16). More generally, although most learning models reflect learning about an additive shock or shift, we suggest an approach that allows customers to learn about a shock with a multiplicative nature.

After switching to a three-part tariff, customers form an initial belief over the distribution of β_i such that

$$\tilde{\beta}_{i0} \sim \text{gamma}(\alpha_0, \beta_0). \tag{12}$$

Customers know the tariffs' characteristics $(p_j, \tilde{q}_j)$, preferences a_1 and η_i that affect their demand intercept, and their demand slope, b_i. At the end of the first period under a three-part tariff, τ_1, they observes their consumption $q^*_{ij\tau_1}$ and receive the signal $S_{i\tau_1}$

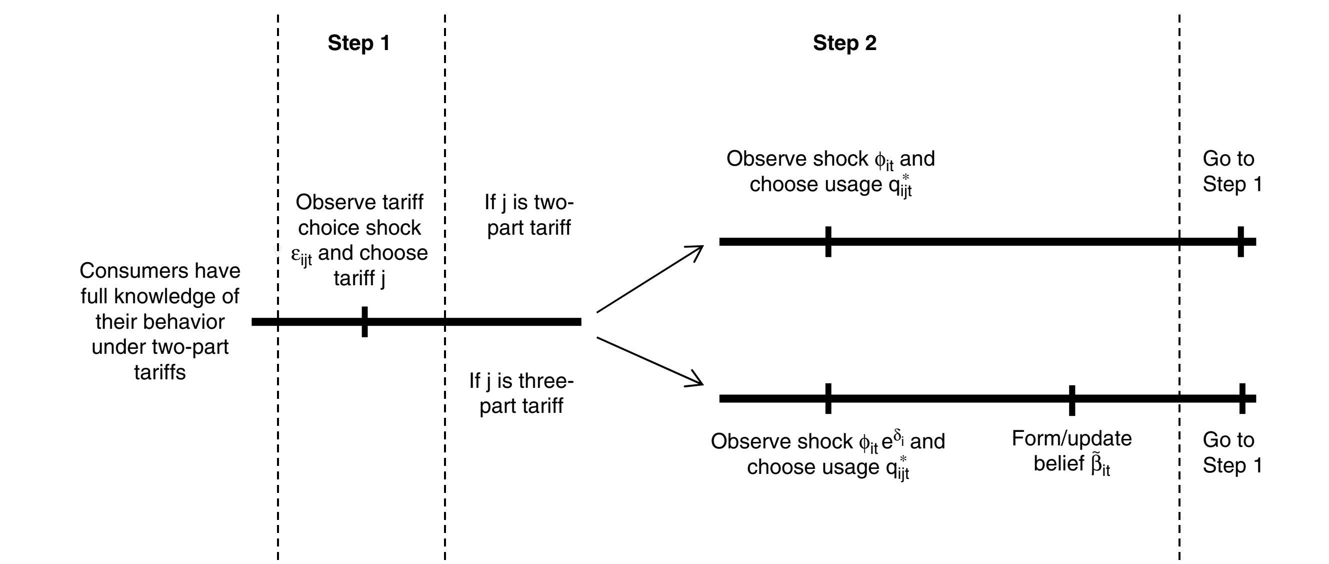
$$S_{i\tau_1} = \begin{cases} \dfrac{q^*_{ij\tau_1}}{e^{h_t a_1 + \eta_i}} & \text{if } q^*_{ij\tau_1} \leq \tilde{q}_j \\[2ex] \dfrac{q^*_{ij\tau_1} + b_i p_j}{e^{h_t a_1 + \eta_i}} & \text{if } q^*_{ij\tau_1} > \tilde{q}_j, \end{cases} \tag{13}$$

which they know is gamma distributed, with known shape parameter r and scale parameter β_i.[6] They update their prior

[5]This result follows from the properties of the gamma distribution: Let $X \sim \text{gamma}(\alpha_1, \alpha_2)$ with $\alpha_1, \alpha_2 > 0$. For any $k > 0$, $kX \sim \text{gamma}(\alpha_1, \alpha_2/k)$.

[6]If a customer's usage is zero or equal to the allowance, $\tilde{q}_j$, no one-to-one relationship exists between the shock and the realized usage so the customer cannot infer the signal s_{it}. We have no observations of zero usage under three-part tariffs. In 11 instances, we observe usage equal to the allowance, resulting in a many-to-one mapping between φ_{it} and q^*_{ijt}. We assume that customers do not update their beliefs in such cases.

Figure 2
SEQUENCE OF MODEL DECISIONS



belief about β_i, $\tilde{\beta}_{i1}$, which then enters the next period's tariff choice,

$$(14) \qquad \tilde{\beta}_{i1} \sim \text{gamma}(\alpha_0 + r, \beta_0 + s_{i\tau_1}).$$

More generally, customers who have spent n periods under a three-part tariff have the following belief about the scale parameter:

$$(15) \qquad \tilde{\beta}_{i\tau_n} \sim \text{gamma}\left(\alpha_0 + nr, \beta_0 + \sum_{t=1}^{n} s_{i\tau_t}\right),$$

with mean and variance

$$(16) \quad E(\tilde{\beta}_{i\tau_n}) = \frac{\alpha_0 + nr}{\beta_0 + \sum_{t=1}^{n} s_{i\tau_t}} \qquad \text{Var}(\tilde{\beta}_{i\tau_n}) \frac{\alpha_0 + nr}{\left(\beta_0 + \sum_{t=1}^{n} s_{i\tau_t}\right)^2}.$$

Because customers are unaware of the value of free minutes before experiencing the three-part tariffs, we set $\alpha_0 = r\beta_0$. This leads to an expected value of the initial belief of r and reduces Equation 15 to Equation 16. The variance of the initial belief is r/β_0. For any value of (α_0, β_0), the expected value of the belief converges to the true value, and the variance goes to zero as the customer gains more experience under a three-part tariff (for proof, see the Web Appendix at www.marketingpower.com/jmr_webappendix).

Identification and Estimation

There are three groups of parameters: (1) preferences for usage and sensitivity to the usage price (η_i, b_i, δ_i); (2) preferences for tariffs, marginal utility of income, and switching costs (λ_i, c, ρ_1, ρ_2); and (3) parameters capturing the distribution of uncertainty, which includes beliefs (β_0) and usage shocks (r). For each customer, we observe tariff choice and usage. Usage is governed by the preference for minutes, the price sensitivity, and the distribution of the usage shock, and tariff choice is governed by *expected* usage (determined by known preferences and beliefs about the unknowns) and tariff choice–specific parameters. For 13.7% of customers, we observe usage under at least two tariffs.

Given that many customers stay with the same tariff for the entire observation period (i.e., they face the same marginal price), we cannot identify unobserved individual-level heterogeneity in both η_i and b_i. Therefore, we estimate η_i with unobserved individual-level heterogeneity, which we identify from differences in usage across customers, and specify the price coefficient b_i as a linear function of individual-level (gender and location) and district-level (labor and literacy levels) demographic variables. That is, $b_i = b + a_2 d_i$, where d_i is a vector containing demographic information and b and a_2 are parameters to be estimated.

We identify the price coefficient from differences in usage prices across customers (i.e., the different usage prices across tariffs, ranging from MU 0 to .079) and within customers (i.e., the change in usage price when a customer switches between tariffs, and the different marginal prices under a three-part tariff, ranging from MU 0 for usage below and MU .050 for usage above the allowance). Observing usage under a variety of different usage prices enables us to precisely pin down the price coefficient. Previous research is restricted to a much lesser degree of price variation, even across customers (Iyengar, Ansari, and Gupta 2007; Lambrecht, Seim, and Skiera 2007; Narayanan, Chintagunta, and Miravete 2007).

We can disentangle δ_i and λ_i because we observe both tariff choice and usage behaviors: λ_i does not enter the usage decision, and δ_i does not affect the initial three-part-tariff choice. We identify λ_i from observed *choices* between two-part and three-part tariffs because we observe that customers who are otherwise similar in their parameter values have a different propensity to choose a three-part tariff. For switchers to a three-part tariff, we identify δ_i from the individual-level differences in *usage* levels on two-part versus three-part tariffs that are not explained by changes to the budget constraint. Note that we can separate δ_i from η_i for all three-part-tariff switchers because we observe the same set of customers under both pricing regimes: first under a two-part tariff and then under a three-part tariff.

Churn is not prominent, so we are unable to identify the marginal utility of income, c, which we set to 1 (Narayanan, Chintagunta, and Miravete 2007). We identify the sensitivity to switching costs, ρ_1, from differences in the propensity to switch between customers with otherwise similar parameters. We can identify the sensitivity to switching costs, ρ_1, separately from the three-part-tariff preference, λ_i, because some customers switch between two-part tariffs only.

The parameter β_0, which reflects the variability of the initial belief about δ_i, is identified from differences in tariff choice after customers have switched to a three-part tariff (11.7% of the three-part tariff switchers later switch to a different tariff). Variability across (*whether* they switch again) and within (*when* they switch) three-part-tariff customers enables us to pin down β_0. Finally, we identify the parameter r, which drives usage uncertainty, from tariff choice and usage variation across customers and periods.

We derive the expression of the likelihood function in the Appendix. It entails the joint probability of tariff choice and usage decisions. We estimate the model in a hierarchical Bayes framework. We use a data augmentation approach to model both the unobserved individual-level parameters and the time-variant beliefs (see the Web Appendix at www.marketingpower.com/jmr_webappendix for details).

In addition to the full model (denoted by Model 3 hereinafter), we also estimate two restricted versions: Model 1, which assumes that customers do not value three-part-tariff minutes any differently from two-part-tariff minutes and so the same parameter set governs behavior on two-part and three-part tariffs, and Model 2, which accounts for a greater valuation of minutes under a three-part tariff but assumes that, after switching to a three-part tariff, customers immediately acquire full knowledge about their value of δ_i.

RESULTS

Estimation Results

Table 3 summarizes the posterior distributions of the parameter estimates of the three models (Model 1, in which we set δ_i to zero; Model 2, which includes δ_i but no learning; and Model 3, which accounts for individual-level learning about δ_i). Using each set of estimates, we predict usage levels for all customers in the data (Table 4). The

Table 3

POSTERIOR DISTRIBUTIONS OF PARAMETER ESTIMATES

	Model 1			Model 2			Model 3		
	M	*95% Interval*		*M*	*95% Interval*		*M*	*95% Interval*	
Demand Intercept									
Individual Level-Intercept									
Mean, μ_η	5.471	5.455	5.487	5.466	5.450	5.482	5.472	5.457	5.488
Standard deviation, σ_η	.694	.682	.705	.694	.683	.705	.687	.676	.699
Dummy for holiday season	.039	.035	.044	.036	.032	.041	.036	.032	.040
Demand Slope, b_i									
Intercept	239.913	223.661	256.461	244.791	229.500	261.232	271.026	255.085	285.935
Female	22.626	9.936	34.704	22.590	9.876	35.279	22.829	8.479	37.983
District-level literacy rate	−25.876	−34.992	−17.653	−26.412	−34.524	−17.274	−28.456	−37.191	−19.073
District-level employment rate	15.206	7.262	23.452	14.207	5.782	23.345	14.573	5.991	23.229
District is capital district	−16.042	−28.230	−3.084	−17.138	−30.327	−3.931	−16.755	−29.627	−3.839
Variance of usage shock, 1/r	.217	.214	.220	.213	.210	.216	.205	.203	.208
Variance of initial belief, r/β_0							.212	.147	.296
Valuation of Free Units									
Mean, μ_δ				.211	.175	.246	.218	.176	.257
Standard deviation, σ_δ				.373	.346	.403	.384	.351	.416
Preferences in Tariff Choice, ζ_{ijt}									
Switching cost between tariffs, ρ_1	−1.025	−1.037	−1.012	−1.051	−1.065	−1.037	−1.025	−1.036	−1.014
Switching cost to other provider, ρ_2	−47.651	−48.027	−47.290	−47.766	−48.094	−47.397	−47.573	−47.949	−47.181
Preference for the Three-Part Tariff									
Mean, μ_λ	−4.916	−5.485	−4.417	−5.257	−5.853	−4.676	−6.024	−6.738	−5.367
Standard deviation, σ_λ	5.714	5.392	6.076	6.310	5.937	6.705	6.412	6.032	6.832
Log Marginal Density	−394,837			−394,197			−394,325		

Notes: N = 5831 customers, 63,449 usage, and 63,616 choice observations.

predicted two-part-tariff usage of 282 minutes in Model 3 compares well with observed usage of 294 minutes. The predicted three-part-tariff usage of 434 minutes accurately reflects the observed three-part-tariff usage of 434 minutes. In contrast, Models 1 and 2 predict considerably lower three-part-tariff usage (369 and 410 minutes, respectively). Similarly, fit measures suggest that Model 3 best reflects usage behavior (Table 5). Compared with Model 1, Model 3 reduces the mean squared error (MSE) for usage, conditional on observed tariff choice, by 3.0%. If we consider three-part-tariff observations only, the MSE reduces by 72.1%. In contrast, Model 2 reduces the MSE by 1.3% for all observations and by 21.5% for three-part-tariff observations compared with Model 1. The mean absolute percentage error confirms that Model 3 performs better than either Model 1 or Model 2. The percentage of correctly predicted choices (hit rate) indicates that all three models predict choice well. All fit measures confirm that Model 3 captures customers’ tariff choice and usage behavior better than Models 1 and 2.

We next turn to the parameter estimates of Model 3 in more detail. The individual-level preference for free minutes, δ_i, has a mean of .218 and a standard deviation of .384. It is positive for 83.9% of three-part-tariff customers, indicating that a large majority of customers value the free minutes beyond their direct cost implications. Similarly, Model 2 finds a positive valuation for 84.8% of customers. The variance of the usage shock, r, has a posterior mean of .205, which translates into a significant effect on usage

Table 4

(IN-SAMPLE) PREDICTED VERSUS ACTUAL USAGE LEVELS (IN MINUTES)

	Model 1: Prediction	*Model 2: Prediction*	*Model 3: Prediction*	*Observed*
Two-Part Tariffs	*283*	*283*	*282*	*294*
Tariff_2_1	154	154	153	158
Tariff_2_2	230	230	229	221
Tariff_2_3	267	266	266	269
Tariff_2_4	267	266	266	269
Three-Part Tariffs	*369*	*410*	*434*	*434*
Tariff_3_1	292	323	326	322
Tariff_3_2	600	642	722	721
Tariff_3_3	812	931	1084	1125

Table 5
INDIVIDUAL-LEVEL FIT MEASURES

	Model 1	*Model 2*	*% Difference*[a]	*Model 3*	*% Difference*[a]
MSE (in Thousands)					
Full sample	46.94	46.33	−1.3	45.54	−3.0
Sample of two-part-tariff usage	46.67	46.38	−.6	46.3	−.8
Sample of three-part-tariff usage	56.85	44.6	−21.5	15.84	−72.1
Mean Absolute Percentage Error					
Full sample	73.38	72.4	−1.2	71.87	−2.1
Sample of two-part-tariff usage	73.91	72.9	−1.2	72.59	−1.8
Sample of three-part-tariff usage	53.97	54.15	−1.0	43.54	−19.3
Hit rate (%)[b]	.98	.98	—	.98	—

[a]Percentage change compared with Model 1.
[b]We compare actual and predicted individual tariff choice in each period and report the average across all observations.

volatility. For example, a value of the usage shock that is equal to its standard deviation shifts the average two-part-tariff usage by 45.3%. The negative coefficient for the sensitivity to switching costs between tariffs, ρ_1, indicates that the switching fee notably reduces switching. Similarly, the negative coefficient for the sensitivity to costs of switching to other providers, ρ_2, shows that customers also have high nonpecuniary costs of leaving the provider.

The parameter β_o pertains to the learning process: The variance of customers' initial belief, r/β_o, reflects the extent of over- or underestimation of the true value of δ_i when a customer first switches to a three-part tariff. For the sample of 247 customers who stay under a three-part tariff for at least four periods, we compute the individual-level deviation between the true value of $\beta_i = r/e^{\delta_i}$ and the belief, $\tilde{\beta}_{it}$. In Period 1, the deviation is 32.8%, and it reduces to 25.5% after customers experience the three-part tariff for three months. This result indicates that customers learn about their three-part-tariff usage over time. Not surprisingly, estimation of the learning process leads to a slightly lower variance of the usage shock, 1/r, than in Model 2. Variation in usage that we previously attributed to the usage shock is now partly captured by learning about δ_i.

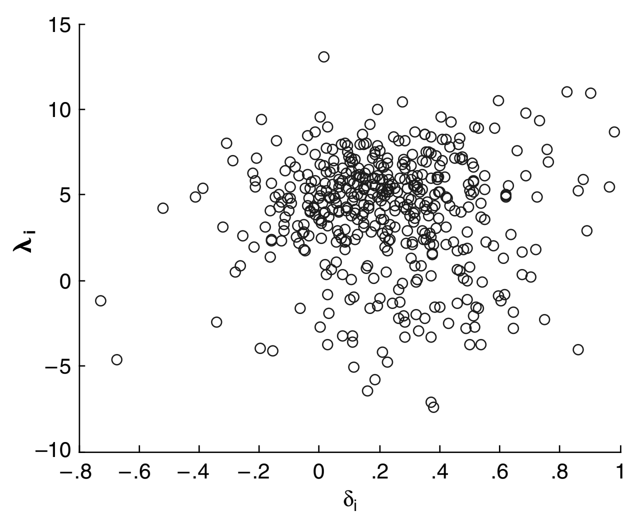
Finally, we examine the correlation between the posterior means of the individual-level parameters λ_i and δ_i (Figure 3). The additional valuation of consuming under a three-part tariff is uncorrelated with the customer's choice preference for three-part tariffs (correlation = .004, p = .930). As a robustness check, we compute the correlation between the posterior means of λ_i and δ_i, considering only customers for whom we observe at least five, six, or seven periods under a three-part tariff.[7] In all cases, we find that the correlation is small and insignificant (Table 6).

In summary, the results provide strong evidence that accounting for the value of free minutes explains behavior significantly better than ignoring such an effect. We show that a large majority of customers value free units above their costs implications and that customers learn about this valuation over time.

Customer Sensitivity to Prices and Allowances

We evaluate the sensitivity of customers' behavior with respect to tariff attributes. From the estimates of Model 3, we compute the elasticity of choice and usage to changes of prices and allowances (Table 7). The two-part tariff's choice elasticity with respect to changes in the access price is −2.13. The three-part tariff's choice elasticity with respect to the access price is −13.28. That is comparable to previously estimated three-part-tariff elasticities of up to −10.15 (Lambrecht, Seim, and Skiera 2007). These estimates reflect that the elasticity increases in the access price and imply that the access price elasticity is, on average, about four times larger for the three-part tariff than the two-part tariff. For the two-part tariff, we find a particu-

Figure 3
CORRELATION AMONG THE POSTERIOR INDIVIDUAL-LEVEL ESTIMATES FOR δ_i AND λ_i



[7]By doing so, we confirm that the absence of correlation is not driven by the prior assumption of independence, because we exclude customers whose individual-level estimates would mostly rely on the prior distribution.

Table 6
CORRELATION BETWEEN λ_i AND δ_i

Number of Three-Part-Tariff Observations per Customer	*Correlation*	p-*Value*	*Number of Customers*
5	.007	.917	195
6	−.015	.864	139
7	−.153	.135	97

Table 7
MODEL 3: SUMMARY OF ELASTICITIES

Elasticity of…	*With Respect to…*	*All Tariffs*	*Two-Part Tariffs*	*Three-Part Tariffs*
Tariff choice	Access price	−2.955	−2.127	−13.276
Tariff choice	Usage price	−10.924	−11.395	−5.052
Tariff choice	Allowance			4.300
Usage	Usage price	−.094	−.100	−.027
Usage	Allowance			.843

larly high choice elasticity with respect to the usage price of −11.39. This is likely due to the structure of the two-part-tariff menu, in which the individual usage prices are relatively comparable. Three-part-tariff choice is less elastic to the usage price (−5.05). Its elasticity to changes in the usage allowance is 4.30, reflecting that the allowance plays an important role in consumer tariff choice.

We now turn to the effect on usage. Similar to previous results (Park, Wetzel, and Mitchell 1983), two-part-tariff usage is relatively inelastic to changes in the usage price (−.10). Usage is even less elastic on three-part tariffs (−.03) as not every usage observation lies above the usage allowance. The elasticity of usage with respect to the allowance is .84. Overall, the prices and allowances are more important in determining choice than usage conditional on choice.

Impact of Free Minutes on the Provider's Forecasts and Revenues

Thus far, our results show clear evidence of the additional value of free minutes. A firm benefits from knowing that customers value usage under a three-part tariff more than under a two-part tariff along several dimensions. First, if a firm were not aware of such greater usage, it would underestimate usage under a three-part tariff, with likely significant impacts on capacity planning. As Table 4 illustrates, a model that ignores the effect of a greater valuation for free minutes (Model 1) underestimates three-part-tariff usage by 14.9%. If a provider is operating close to capacity limits, such a prediction error may easily cause a drop in call quality and customer satisfaction, which eventually could lead to higher churn. Finally, systematic errors in usage predictions will lead to misleading revenue forecasts, resulting in a negative impact on management decisions, such as budgeting, customer resource allocation (based on, for example, customer lifetime value calculations), and targeting.

To measure the effect of δ_i for our provider's revenue, we simulate three-part-tariff revenue based on the individual-level estimates of the full model (Model 3) and compare it with the revenue the company would obtain if customers did not attach greater value to the free minutes (i.e., we set δ_i to zero). We find that the mean expected revenue per three-part-tariff customer decreases from MU 21.3 to MU 17.1. This result means that the preference for free minutes, δ_i, accounts for 19.7% of the revenues obtained from three-part-tariff customers and thus represents a significant fraction of a firm's three-part-tariff revenues.

RECOMMENDATIONS TO THE FIRM

Our results indicate that three-part tariffs can increase customer usage and firm revenues. We next explore whether and how the firm can exploit this insight by encouraging customers to switch to three-part tariffs. We investigate the impact of lowering the switching fee on the firm's revenues, whether the firm can increase revenues by increasing customers' choice preferences for three-part tariffs, and whether the firm would benefit from eliminating two-part tariffs. Because we do not have data on customer acquisition, we restrict the analysis to the current set of customers.

Lowering the Switching Fee

Our estimates show that the switching fee strongly affects switching ($\rho_1 = -1.025$). To check whether the firm could further increase its revenues by reducing the switching fee, we simulate revenues in the period following our observation window under different levels of the switching fee using the estimates from Model 3. Figure 4 shows that the firm would maximize its revenues if it reduced the switching fee from the current level of MU 10 to MU 3.6. We find that this change would increase the firm's revenues by 3.9% (95% posterior interval [1.9%, 5.8%]).[8]

We next examine what share of this revenue increase is due to the effect of δ_i. We compute expected revenues using the same model estimates but now set δ_i to zero. Figure 4, Panel A, illustrates that the revenue increase from lowering the switching fee is largely due to customers' preference for free minutes. Figure 4, Panel B, illustrates that at the optimal level of the switching fee of MU 3.6, this effect is significantly different from zero. In the absence of δ_i, reducing the switching fee to MU 3.6 would not significantly increase revenues, as shown by the posterior mean of the revenue increase is 1.5% with a 95% posterior interval of [−.2%, 3.3%] (Figure 4, Panel C). In conclusion, the increase in revenue obtained by reducing the switching fee, thus encouraging customers to switch to three-part tariffs, is mainly driven by customers' greater valuation for three-part-tariff minutes.[9]

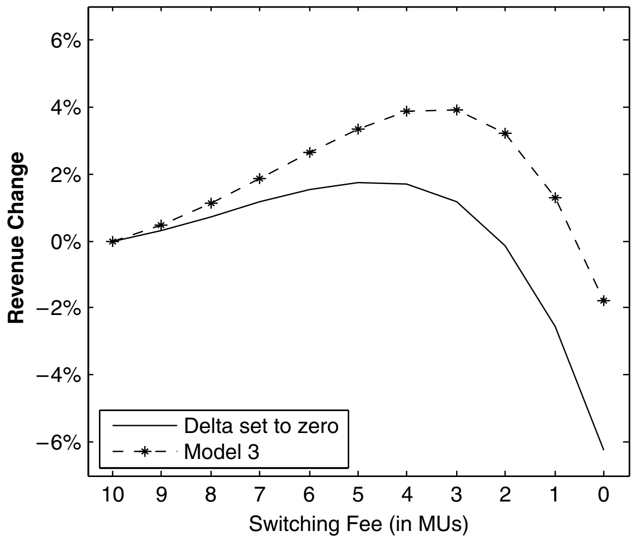
[8]An increase in revenue may not fully translate into an increase in profits if the firm does not have excess capacity. However, in this case the company's network was not operating at or close to capacity constraints.

[9]Note that our econometric model assumes that customers' choice decisions are based on the next period only. If customers were forward looking, our model would potentially overestimate customers' sensitivity to switching costs, and the effect of lowering the switching fee on provider revenues may be lower than we predict here. We run a sensitivity analysis in which we decrease the estimate for customers' sensitivity to the switching fee, ρ_1, by 5%, 10%, and 20%. Our findings are robust to those changes (see the Web Appendix at www.marketingpower.com/jmr_webappendix).

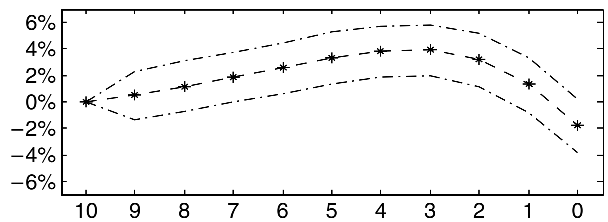
Figure 4

CHANGE IN REVENUE AND SURPLUS DUE TO REDUCTION OF THE SWITCHING FEE

A: Posterior Mean for Revenue Change



B: Posterior Interval for Model 3



C: Posterior Interval for Delta Set to Zero

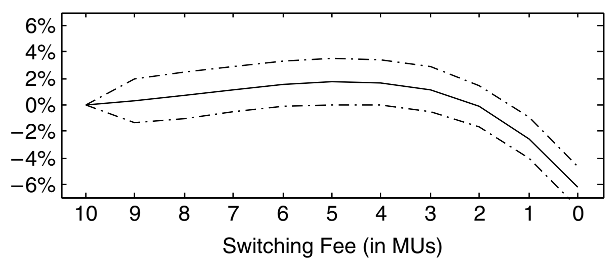
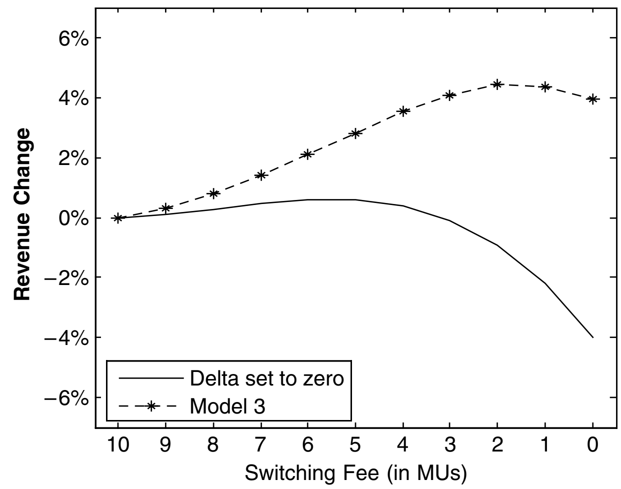


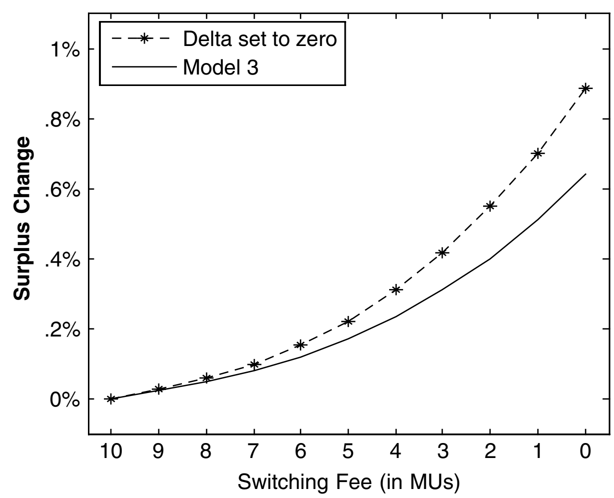
Figure 5

CHANGE IN REVENUE AND SURPLUS DUE TO CHANGES IN SWITCHING FEE AND TARIFF OFFERINGS

A: Posterior Mean for Revenue Change



B: Posterior Mean for Surplus Change



Adjusting the Tariff Structure

The firm could also consider removing two-part tariffs from its product offering while allowing current two-part-tariff customers to remain with their tariff. We investigate whether such a decision would be beneficial if the firm also lowered the switching fee. We find an optimal switching fee of MU 2.2, which implies an expected revenue increase of 4.1% (95% posterior interval [1.8%, 6.4%]) and an average increase in customer surplus of .4% (95% posterior interval [.2%, .6%]). The revenue effect would be negligible, and even negative, if customers did not have a preference for free minutes (Figure 5). This result highlights the risk for a firm to misalign its pricing structure if it were to abstract from the value of free minutes to consumers.

Note that the optimal level of the switching fee is lower than the optimal switching fee if the firm continued to offer two-part tariffs. This difference is driven by customers who previously would have switched to a different two-part tariff, provided the switching fee was sufficiently low. Now, because two-part tariffs are no longer available, these customers either remain with their current tariff, keeping their revenues constant, or switch to a three-part tariff, in which revenue is likely to increase because of the preference for free minutes.

Because customers may appreciate a simplification of the provider's pricing offering, we also examine the consequences of eliminating the switching fee entirely. If, as Figure 5 illustrates, the firm eliminated the switching fee and also removed the two-part tariffs, expected revenue would increase by 4.0% (95% posterior interval [1.5%, 6.5%])—only slightly less than at the optimal level of the switching fee—and customer surplus would increase by .6% (95% posterior interval [.5%, .8%]). In contrast, if two-part tariffs were still under offer, eliminating the switching fee would lead to a revenue loss because customers would be more likely to switch to two-part than to three-part tariffs (Figure 4).

CONCLUSION

Compared with two-part tariffs, three-part tariffs introduce free units of consumption. Behavioral research suggests that free products affect customers' behavior beyond what the change in the budget constraint would predict. In this research, we examine how demand changes when customers switch from a two-part to a three-part tariff. We argue that the free component of a three-part tariff leads to a positive affective response, thus increasing the valuation of the service, an effect that persists even after consumers have exceeded their usage allowance.

To identify the effect of tariff structure separately from the shift to the budget constraint, we jointly estimate tariff choice and usage. We explicitly model the effect of the usage allowance on the valuation of a three-part tariff. The results confirm that the structure of the three-part tariff affects behavior. That is, the majority of customers value free units beyond the change to their cost structure, significantly increasing the firm's revenue. The additional valuation of free minutes accounts for 19.7% of the revenue generated from three-part-tariff customers. We provide evidence that after switching to a three-part tariff, customers learn about their valuation over time. We find that the provider would benefit from reducing the switching fee, which would lead more customers to switch to three-part tariffs, and from discontinuing the option to switch to two-part tariffs. Thus, customers' valuation for free units is crucial to any revenue increase.

Our findings have important implications. First, they suggest that companies should conduct field tests of new tariff structures to better understand how new tariffs affect customer behavior. This is important to set optimal prices and switching fees. Second, it may be beneficial for providers to advertise free minutes more intensely because of their high value to consumers. Third, companies that face capacity constraints should take into account that changing the tariff structure could lead to greater-than-expected usage of current capacity.

More broadly, our results highlight the dual role of a nonlinear pricing plan: It not only determines the cost to the customer but also alters the perceived characteristics of the service and thus influences customers' choice and usage. Overall, our findings add a new dimension to nonlinear pricing research that typically assumes that the difference in tariff structure affects usage exclusively through the budget constraint.

Our results motivate a more extensive study of how different tariff structures affect usage. A limitation of our study is that the data only include customers who were with the firm before it introduced three-part tariffs. Future work could address customer acquisition and market expansion effects of introducing three-part tariffs and could examine how a firm can optimally combine three- and two-part tariffs. Furthermore, research should investigate behavior when customers switch back to two-part tariffs. Because we have few periods in which we observe customers switching back, our data do not address whether usage declines beyond what a change in the budget constraint would predict or whether greater usage alters preferences to such an extent that it becomes persistent, even when a customer leaves a three-part tariff. In addition, although we ran multiple robustness checks and sensitivity analyses, we cannot conclusively rule out that other factors contribute to the phenomenon we observe. Further research could address this issue with experimental data, in which all alternative accounts can be controlled for at the same time. Finally, research could examine the effect of other tariff structures (e.g., bucket pricing) that strictly limit consumption to the usage allowance (Schlereth and Skiera 2012) on consumers' valuation of a service.

APPENDIX

Expected Indirect Utility

Equation 9 shows the expected indirect utility of choosing a two-part tariff. In the estimation, we evaluate the probability of consuming zero minutes and taking expectations over the distribution of ϕ_{it}. From the demand Equation 3, it follows that the probability of zero usage is

$$P(d_{ijt} \le b_i p_j) = P\left(\phi_{it} \le \frac{b_i p_j}{e^{h_t a_1 + \eta_i}}\right),$$

which corresponds to the cumulative distribution function (CDF) of a gamma distribution with parameters (r, r) evaluated at $b_i p_j / e^{h_t a_1 + \eta_i}$. To simplify the notation, hereinafter we denote $G(x \mid \alpha_1, \alpha_2)$ as the CDF of a gamma distribution of shape parameter α_1 and scale parameter α_2, evaluated at x, and $g(x \mid \alpha_1, \alpha_2)$ as the probability density function of a gamma distribution defined as $g(x \mid \alpha_1, \alpha_2) = \alpha_2^{\alpha_1} x_2^{\alpha_1 - 1} e^{-\alpha_2 x} / \Gamma(\alpha_1)$. The probability of observing zero usage is therefore

$$P(q^*_{ijt} = 0) = G\left(\frac{b_i p_j}{e^{h_t a_1 + \eta_i}} \mid r, r\right).$$

The term $\phi_{it} \mid \phi_{it} > (b_i p_j / e^{h_t a_1 + \eta_i})$ follows a truncated gamma distribution. Therefore, we can express the expected value of the demand intercept under a two-part tariff, $\phi_{it} e^{h_t a_1 + \eta_i}$, conditional on usage greater than zero, as follows:

$$E(\phi_{it} e^{h_t a_1 + \eta_i} \mid q^*_{ijt} > 0) = e^{h_t a_1 + \eta_i} \frac{1 - G\left(\frac{b_i p_j}{e^{h_t a_1 + \eta_i}} \middle| r+1, r\right)}{1 - G\left(\frac{b_i p_j}{e^{h_t a_1 + \eta_i}} \middle| r, r\right)}.$$

We substitute this equations into Equation 9 and obtain the expected indirect utility of a two-part tariff, which determines a customer's two-part tariff choice:

$$\text{(A1)}\quad E[V_{ijt}] = G\left(\frac{b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, r\right)c_i(y_{it} - F_j) + \left[1 - G\left(\frac{b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, r\right)\right] \times c_i\left[y_{it} - F_j + \frac{1}{2}b_i p_j^2 - e^{h_t a_1 + \eta_i} \times \frac{1 - G\left(\frac{b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r+1, r\right)}{1 - G\left(\frac{b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, r\right)} p_j\right] + \varsigma_{ijt}.$$

Similarly, we derive the expression of the expected indirect utility of choosing a three-part tariff (from Equation 10). Substituting Equation 8 into Equation 4, we obtain

$$P(q^*_{ijt} \le \tilde{q}_j) = P\left(\phi_{it} e^{\delta_i} \le \frac{\tilde{q}_i + b_i p_j}{e^{h_t a_1 + \eta_i}}\right).$$

Thus, the probability that a customer's three-part-tariff usage is equal to or below the allowance is

$$P(q^*_{ijt} \le \tilde{q}_j) = G\left(\frac{\tilde{q}_j + b_i p_j}{e^{h_t a_1 + \eta_i}} \,\middle|\, r, \beta_i\right).$$

The term $\phi_{it} e^{\delta_i} \,|\, \phi_{it} e^{\delta_i} > (\tilde{q}_j + b_i p_j / e^{h_t a_1 + \eta_i})$ follows a truncated gamma distribution with parameters (r, β_i). We express the expected indirect utility of a three-part tariff as follows:

$$\text{(A2)}\quad E[V_{ijt}] = G\left(\frac{\tilde{q}_j + b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, \beta_i\right)c_i(y_{it} - F_j) + \left[1 - G\left(\frac{\tilde{q}_j + b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, \beta_i\right)\right] \times c_i\left[y_{it} - F_j + p_j\tilde{q}_j + \frac{1}{2}b_i p_j^2 - e^{h_t a_1 + \eta_i}\frac{r}{\beta_i}\frac{1 - G\left(\frac{\tilde{q}_j + b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r+1, \beta_i\right)}{1 - G\left(\frac{\tilde{q}_j + b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, \beta_i\right)} p_j\right] + \varsigma_{ijt}.$$

Customers who have not yet switched to a three-part tariff are not aware of its additional value so their choice is unaffected by δ_i. Therefore, at their initial three-part-tariff choice, δ_i does not enter their expected indirect utility of a three-part tariff. They thus take expectations over the shock ϕ_{it}, and the expected indirect utility of a three-part tariff simplifies to

$$\text{(A3)}\quad E[V_{ijt}] = G\left(\frac{\tilde{q}_j + p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, r\right)c_i(y_{it} - F_j) + \left[1 - G\left(\frac{\tilde{q}_j + p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, r\right)\right] \times c_i\left[y_{it} - F_j + p_j\tilde{q}_j + \frac{1}{2}b_i p_j^2 - e^{h_t a_1 + \eta_i}\frac{1 - G\left(\frac{\tilde{q}_j + p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r+1, r\right)}{1 - G\left(\frac{\tilde{q}_j + p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, r\right)} p_j\right] + \varsigma_{ijt}.$$

Overall Likelihood Function

For every customer i and time t, we observe usage, q_{ijt}, and tariff choice, k_{it}. These are outcomes of the customer and time-specific covariates, $Z_{it} = \{h_t, d_t\}$; the tariff-specific characteristics, $X_j = \{p_j, F_j, \tilde{q}_j\}$; the population parameters, $\Phi = \{b, \rho_1, \rho_2, \beta_0, a_1, a_2, \exp(r)\}$; the time-invariant individual parameters, $\varpi_i = \{\eta_i, \delta_i, \lambda_i\}$; and the individual-specific time-variant beliefs $\tilde{\beta}_{it}$.

Likelihood of usage. For a consumer under a two-part tariff, the probability of observing a particular usage level, given the tariff choice, is

$$\text{(A4)}\quad f(q_{it} \mid k_{it}, \Phi, \varpi_i, Z_{it}, X_j) = P(q^*_{ijt} = q_{ijt} \mid k_{it} = j, \Phi, \varpi_i, Z_{it}, X_j) = P(q^*_{ijt} = 0)I^{\{q_{it}=0\}} + P(q^*_{ijt} = \phi_{it}e^{h_t a_1 + \eta_i} - b_i p_j)I^{\{q_{it}>0\}}, = G\left(\frac{b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, r\right) + \Im_{it}\, g\left(\frac{b_i p_j}{e^{h_t a_1 + \eta_i}}\,\middle|\, r, r\right)I^{\{q_{it}>\tilde{q}_j\}}$$

where the first term corresponds to the probability of observing zero usage, expressed as the CDF of a gamma distribution with shape and scale parameter r, evaluated at $b_i p_j / e^{h_t a_1 + \eta_i}$. The term $I^{\{A\}}$ is the indicator function that takes the value of 1 if statement A is true and zero if otherwise. We divide the second term into two parts, the Jacobian of the transformation from q_{ijt} to ϕ_{it}, $\Im_{it} = (1/e^{h_t a_1 + \eta_i})$, and the probability density function of ϕ_{it}, which is distributed gamma (r, r).

For a consumer under a three-part tariff, the probability of observing a particular usage level is

$$\text{(A5)}\quad f(q_{it} \mid k_{it}, \Phi, \varpi_i, h_t, X_j) = P(q^*_{ijt} = q_{ijt} \mid k_{it} = j, \Phi, \varpi_i, h_t, X_j) = P(q^*_{ijt} = \phi_{it}e^{z'_{it}a_i + \eta_i + \delta_i})I^{\{q_{it}<\tilde{q}_j\}} + P(q^*_{ijt} = \tilde{q}_j)I^{\{q_{it}=\tilde{q}_j\}} + P(q^*_{ijt} = \phi_{it}e^{z'_{it}a_i + \eta_i + \delta_i} - b_i p_j)I^{\{q_{it}>\tilde{q}_j\}} = \Im_{it}\, g\left(\frac{q_j}{e^{h_t a_1 + \eta_i + \delta_i}}\,\middle|\, r, r\right)I^{\{q_{it}<\tilde{q}_j\}} + \left[G\left(\frac{\tilde{q}_j + b_i p_j}{e^{h_t a_1 + \eta_i + \delta_i}}\,\middle|\, r, r\right) - G\left(\frac{\tilde{q}_j}{e^{h_t a_1 + \eta_i + \delta_i}}\,\middle|\, r, r\right)\right]I^{\{q_{it}=\tilde{q}_j\}} + \Im_{it}\, g\left(\frac{q_j + b_i p_j}{e^{h_t a_1 + \eta_i + \delta_i}}\,\middle|\, r, r\right)I^{\{q_{it}>\tilde{q}_j\}}.$$

Likelihood of tariff choice. The tariff-specific shock ε_{ijt} is assumed to follow a Type 1 extreme value distribution.

Therefore, the probability of choosing a particular tariff j is given by

$$\text{(A6)}\quad f(k_{it} \mid \alpha, \varpi_i, \tilde{\beta}_{it}, Z_{it}, X_j) = P(k_{it} = j \mid \Phi, \varpi_i, \tilde{\beta}_{it}, Z_{it}, X_j) = \prod_{j=0}^{J} \frac{e^{\tilde{V}_{ijt}}}{\sum_{g=0}^{J} e^{\tilde{V}_{igt}}},$$

where the term $\tilde{V}_{ijt}$ denotes the expected indirect utility of each tariff j. We obtain the likelihood function by integrating the customer's tariff choice and usage decisions:

$$\text{(A7)}\quad L = \prod_{i=1}^{I}\prod_{t=1}^{T_i}[f(q_{it} \mid k_{it}, \alpha, \varpi_i, Z_{it}, X_j) \times f(k_{it} \mid \alpha, \varpi_i, \tilde{\beta}_{it}, Z_{it}, X_j)]$$

REFERENCES

Anderson, Eric and Duncan Simester (2003), "Effect of $9 Price Endings on Retail Sales: Evidence from Field Experiments," *Quantitative Marketing and Economics*, 1 (1), 93–110.

Bagh, Adib and Hemant K. Bhargava (2008), "Doing More with Less: Price Discrimination with Three-Part vs. Two-Part Tariffs," working paper, Department of Economics, University of Kentucky.

Bertini, Marco and Luc Wathieu (2008), "Research Note: Attention Arousal Through Price Partitioning," *Marketing Science*, 27 (2), 236–46.

Cnet.com (1996), "New Usage Plan Swamps AOL," (December 2), [available at http://news.cnet.com/New-usage-plan-swamps-AOL/2100-1023_3-251324.html].

Darke, Peter R. and Darren W. Dahl (2003), "Fairness and Discounts: The Subjective Value of a Bargain," *Journal of Consumer Psychology*, 13 (3), 328–38.

DellaVigna, Stefano and Ulrike Malmendier (2006), "Paying Not to Go to the Gym," *American Economic Review*, 96 (3), 694–719.

Donovan, Robert J., John R. Rossiter, Gilian Marcoolyn, and Andrew Nesdale (1994), "Store Atmosphere and Purchasing Behavior," *Journal of Retailing*, 70 (3), 283–94.

Dubin, Jeffrey A. and Daniel L. McFadden (1984), "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," *Econometrica*, 52 (2), 345–62.

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

Goettler, Ronald L. and Karen B. Clay (2011), "Tariff Choice with Consumer Learning and Switching Costs," *Journal of Marketing Research*, 48 (August), 633–52.

Golden, Linda and Mary R. Zimmer (1986), "Relationship Between Affect, Patronage, Frequency and Amount of Money Spent with a Comment on Affect Scaling and Measurement," in *Advances in Consumer Research*, Vol. 13, Richard J. Lutz, ed. Provo, UT: Association for Consumer research, 53–57.

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

——— and Matthew Osborne (2011), "Tariff Choice, Usage Uncertainty, Biased Beliefs, and Learning," working paper, MIT Sloan School of Management.

Hanemann, W. Michael (1984), "Discrete/Continuous Models of Consumer Demand," *Econometrica*, 52 (3), 541–62.

Heilman, Carrie M., Kent Nakamoto, and Ambar G. Rao (2002), "Pleasant Surprise: Consumer Response to Unexpected In-Store Coupons," *Journal of Marketing Research*, 39 (May), 242–52.

Isen, Alice M., Margaret Clark, Thomas E. Shalker, and Lynn Karp (1978), "Affect, Accessibility of Material in Memory, and Behavior: A Cognitive Loop?" *Journal of Personality and Social Psychology*, 36 (1), 1–12.

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

———, Kamel Jedidi, Skander Essegaier, and Peter J. Danaher (2011), "The Impact of Tariff Structure on Customer Retention, Usage and Profitability of Access Services," *Marketing Science*, 30 (5), 820–36.

Jensen, Sissel (2006), "Implementation of Competitive Nonlinear Pricing: Tariffs with Inclusive Consumption," *Review of Economic Design*, 10 (1), 9–29.

Lambrecht, Anja, Katja Seim, and Bernd Skiera (2007), "Does Uncertainty Matter? Consumer Behavior Under Three-Part Tariffs," *Marketing Science*, 26 (5), 698–710.

——— and Bernd Skiera (2006), "Paying Too Much and Being Happy About It: Existence, Causes, and Consequences of Tariff-Choice Biases," *Journal of Marketing Research*, 43 (May), 212–23.

Narayanan, Sridhar, Pradeep K. Chintagunta, and Eugenio J. Miravete (2007), "The Role of Self Selection and Usage Uncertainty in the Demand for Local Telephone Service," *Quantitative Marketing and Economics*, 5 (1), 1–34.

Nunes, Joseph C. (2000), "A Cognitive Model of People's Usage Estimations," *Journal of Marketing Research*, 37 (November), 397–409.

Park, Rolla Edward, B.M. Wetzel, and B.M. Mitchell (1983), "Price Elasticities for Local Telephone Calls," *Econometrica*, 51 (6), 1699–1730.

Schlereth Christian and B. Skiera (2012), "Measurement of Consumer Preferences for Bucket Pricing Plans with Different Service Attributes," *International Journal of Research in Marketing*, 29 (June), 167–80.

Shampanier, Kristina, Nina Mazar, and Dan Ariely (2007), "Zero as a Special Price: The True Value of Free Products," *Marketing Science*, 26 (6), 745–57.

Sherman, Elaine and Ruth B. Smith (1987), "Mood States of Shoppers and Store Image: Promising Interactions and Possible Behavioral Effects," in *Advances in Consumer Research*, Vol. 14, Melanie Wallendorf and Paul Anderson, eds. Provo, UT: Association for Consumer Research, 251–54.

Thomas, Manoj and Vicky Morwitz (2005), "Penny Wise and Pound Foolish: The Left-Digit Effect in Price Cognition," *Journal of Consumer Research*, 32 (1), 54–64.

Wilson, Robert (1993), *Nonlinear Pricing*. New York: Oxford University Press.