LEARNING ABOUT NEW TECHNOLOGIES: EVIDENCE FROM MOBILE PHONE USAGE IN RWANDA

DANIEL BJÖRKEGREN*

HARVARD UNIVERSITY

Although the spread of new technologies is vital for economic development, it is difficult to study with traditional sources of data. The mobile phone represents a new technology which automatically records every potential learning experience, and nearly every remote interaction with peers who could share their own learning experiences. In 2006, a Rwandan mobile phone operator introduced a new plan that represented substantial savings for over 85% of subscribers. This project uses operator data to investigate how individuals learned about this new plan, and aims to differentiate between learning by doing, from the experience of social network neighbors, and from official sources.

Keywords: social networks, learning, information technology, big data

1. Introduction

Learning about new technologies is vital for economic growth. However, it is difficult to study: seldom is it possible to gather data rich enough to both describe behavior and differentiate between channels of learning. The mobile phone is an economically important technology that has seen widespread adoption. It has the relatively unique feature that operators must maintain a log of all actions taken on the network in order to provide service. These records provide a realtime window into how individual subscribers learn to use the network. This project uses rich data from a mobile phone network to track how subscribers learn about a new cheaper phone plan, to determine how profitable technologies diffuse through society.

This connects to several literatures: it connects to the literature on technology adoption in developing countries, which finds that people learn about using profitable technologies based on own and neighbors' experimentation (Foster and Rosenzweig, 1995; Conley and Udry, 2010), and that heterogeneity can explain some patterns of nonadoption (Suri, 2011). It also connects to work on social learning (Banerjee, 1992; Ellison and Fudenberg, 1993). It also connects to the multidisciplinary literature on diffusion on networks (Jackson and Yariv, 2008).

---

*E-mail: danbjork@fas.harvard.edu, Web: http://dan.bjorkegren.com
Revision April 2, 2014. Preliminary and incomplete.
Thank you to Nathan Eagle for providing access to the data, computing facilities, and helpful conversations. In Rwanda, I thank the staff of my telecom partner, RURA, RDB, NISR, and BRD for helpful conversations and access to data. This work was supported by the Stanford Institute for Economic Policy Research through the Shultz Fellowship in Economic Policy.

This setting has three features beneficial for studying learning over networks.

First, mobile phone networks provide a rich, passively collected source of data on both actions and social networks. The data represents nearly the entire network of mobile phone subscribers; to the extent that subscribers learn about mobile phone plans from other subscribers, there is no bias from sampling nodes (Chandrasekhar and Lewis, 2012). The temporal resolution of the data makes it possible to use micro-level event studies for identification: it is possible to determine whether actions preceded or followed the provision of information. Also, transaction records are passively collected; their collection does not influence observed behavior, which would be a concern with surveys gathered at sufficient frequency to track learning over time (Zwane et al., 2011).

Second, subscribers are making economically significant decisions. Telephony represents 5% of expenditure in subscribing households.[1] Plan choice itself is not a trivial decision: on average, subscribers saved almost half of their spending on domestic calls by switching to the newly offered plan; they could have saved approximately 10% more by switching earlier. The operator introduced the plan as part of a sequence of billing changes that made the network more attractive to poor consumers. Although the introduction of the plan also affects operator revenues, this paper focuses on consumers, for which the new plan is analogous to a new technology with large but heterogeneous benefits.

Third, the context is information-poor: the mobile phone network in this country represents the vast majority of remote communication apart from radio. This feature makes it easier to distinguish between confounding sources of information relative to settings with more complex communication networks.

The paper proceeds as follows: the next section details the context, with a particular eye to the communication environment. Section 3 introduces the data used in the project. Section 4 describes the new plan that was made available to subscribers. Section 5 models the decision to adopt the plan. Section 6 presents descriptive results. Section 7 presents robustness checks. Section 8 concludes.

2. Context

The setting is Rwanda, a small, low income African country. It is information-poor, with low penetration of communication technology.

The fixed line network is insignificant, with 0.2% penetration that was fairly constant over time. The mobile phone network dwarfs the landline network: in 2005 the operator I study had a penetration of 2%, and by 2009 the penetration reached 13%. During this period, the regulator restricted the market to two mobile phone operators. The operator from which I have data dominated the market, with 88% market share. There is a network of payphones,

[1] According to the government's household survey, EICV 2010-2011.

but these run on the mobile phone network so their transactions are reported just like other mobile phones.

Not only is the mobile phone network the primary telephone network, it also represents the majority of remote communication apart from radio. The volume of mail sent through the postal service is insignificant (0.2 pieces per person (2007), relative to 2.4 pieces in Kenya and 538.8 pieces in the US (2011)[2]). Most households have radio, but few have electricity, televisions, newspapers, or access to the internet, as shown in Table 1.

Households with mobile phones tend to be richer and more urban than households in general, though less wealthy individuals subscribe as the network expands and handset prices drop.

3. Data

This project makes use of two main data sources:

Call Detail Records. As a side effect of providing service, mobile phone operators maintain transaction records, called Call Detail Records (CDRs). This project uses an anonymized CDR from the dominant Rwandan operator. This data includes nearly every call, SMS, top up, and plan choice made over 4.5 years by the country's mobile phone subscribers, numbering approximately 200,000 in January 2005 and 1.5 million in May 2009. The CDR provides identifiers for the sender and recipient of each transaction, the time stamp, duration, the cell towers used to transmit the call at the beginning and end, and for transactions before July 2008, the incurred charge. The cell tower identifiers can be linked to coordinate locations provided by the operator. The locations of some tower identifiers are missing from this data, so I infer them based on call handoffs with known towers, using a procedure described in Appendix A.

Operator Billing Policies. Details on the operator's historical billing policies were obtained from several sources, including archived versions of the operator's web site, reports from the government regulator, and news articles. The resulting billing model was checked against billing records and adjusted until it fit, in a process described in Appendix B.

4. A New Plan

The vast majority of subscribers (99%) use prepaid plans. These plans have no monthly fee; instead, subscribers maintain a balance that is depleted with use. This balance can be refilled by purchasing an airtime scratch card from any of the operator's agents, who are conveniently located throughout the country (in 2008 there were roughly 20,000 agents).[3] The party that initiates a call pays for the call; receiving a call is free.

[2]Sources: National Institute of Statistics Report 2008, Communications Commission of Kenya "Study on ICT Access Gaps in Kenya", U.S. Postal Service, U.S. Census

[3]Source: Operator

Initially, the operator offered one prepaid plan ("PerMin"), which was billed by the first minute and then every 30 seconds thereafter. As the network expanded to reach poorer consumers, the operator made a sequence of changes in billing policy that reduced prices. The most notable of these was the introduction of a plan billed by the second ("PerSec"). The pricing of the new plan was set so that a 45 second call cost the same under either plan; shorter calls were cheaper under PerSec. The top panel of Figure 1 compares pricing between the two plans.

Most subscribers make very short calls (the median call length is 22 seconds), and as a result PerSec was cheaper for most subscribers. The bottom panel of Figure 1 overlays the distribution of call durations on the pricing of the two plans. Over all domestic transactions, PerSec was cheaper for 97% of subscribers. Even the charges incurred while subscribed to PerMin would have been cheaper under PerSec for 85% of subscribers.[4] The savings associated with switching were substantial: switchers saved on the order of half of their domestic calling expenditure.

The process of switching was simple and did not incur a fee: to switch plans, a subscriber would enter a short code into the phone, not dissimilar from the codes entered when topping up airtime. Thus in order to switch, a subscriber must know both that a different plan is available and the short code to enter into the phone.

Despite the large savings associated with switching to per second billing, it took years for the plan to diffuse through society, as shown by the evolution of subscriber counts in Figure 2. The remainder of this paper analyzes the process through which this plan diffused through society.

5. Model

Each individual i has a choice of using a mobile phone, or not and receiving utility zero. Individuals receive heterogeneous benefits from using a phone. If i uses a phone with the existing technology (PerMin), he receives net benefit:

$$b_i(X) - c^M(X) - a_i$$

where X is a choice variable describing how the technology is used, $b_i(X)$ represents i's benefit of using a mobile phone under choice X, $c^M(X)$ represents the cost under the PerMin plan, and a_i is a cost of adoption. In this case X represents a distribution of call durations.[5]

The optimal X is given by:

[4] I have not yet considered spending on international calls, which are a small fraction of calls. International calls tend to be a bit longer: the median is 42 seconds; there may be a small fraction of users making many international calls who would not find it profitable to switch.

[5] More precisely, a distribution of call durations across each link of i's full social network. This choice variable is analogous to the amount of fertilizer applied to a crop in Foster and Rosenzweig (1995).

$$X_i^{*M} = \arg\max_X b_i(X) - c^M(X)$$

and the individual will use the technology if:

$$b_i(X_i^{*M}) - c^M(X_i^{*M}) - a_i \geq 0$$

Now, a new technology (PerSec) has been introduced which has the same benefit but a different cost function, $c^S(X)$. The adoption cost is the same, but for current subscribers there is a switching cost of k to move to the new technology. The optimal choice under the new technology is:

$$X_i^{*S} = \arg\max_x b_i(X) - c^S(X)$$

The introduction of the new technology changes behavior in two ways. Some individuals who subscribed under the old plan find it optimal to switch. It is optimal for i to switch if:

$$b_i(X_i^{*S}) - c^S(X_i^{*S}) - k \geq b_i(X_i^{*M}) - c^M(X_i^{*M}) \tag{1}$$

And other individuals who did not find it optimal to join under the old plan may find it optimal to join under the new plan: individuals for whom the adoption cost was less than utility under the new plan but more than utility under the old plan ($b_i(X_i^{*M}) - c^M(X_i^{*M}) < a_i \leq b_i(X_i^{*S}) - c^S(X_i^{*S})$). In this paper I focus on existing subscribers' decisions to switch to the new plan.

If subscribers knew all of these objects, adjustment to the new plan would happen immediately, as the plan itself was technically made available instantaneously to all subscribers. However, if there is imperfect information, individuals may delay switching. Uncertainty could arise from several sources. Individuals may not be aware of the existence of the new plan or the short code. Or, they may have uncertainty about the switching cost k, the new cost structure $c^S(x)$, or the benefit function $b_i(X)$ in the region near the optimal usage under the new plan, X_i^{*S}.

Since I do not observe the function $b_i(X)$, I am unable to evaluate Equation 1 empirically. However, I can compute whether it is optimal to switch given a behavior X. In particular, it is unambiguously optimal for subscriber i to switch if:

$$c^S(X_i^{*M}) + k \leq c^M(X_i^{*M})$$

I introduce a convenient metric to describe the optimal plan choice, the ratio of charges under zero switching cost:

$$ChargeRatio_{PM/PS}(X) := \frac{c^M(X)}{c^S(X)}$$

Given an observed behavior X, one of these costs represents the incurred charge under the current plan and one represents the counterfactual under a different plan. This ratio is one for behaviors X that incur the same cost under either plan; it is below one for behaviors that incur higher costs under the PerSec plan and above one for behaviors that incur higher costs under the PerMin plan. Individuals with overall charge ratios above one would save under the PerSec plan, as long as the switching cost k is sufficiently low.

6. Descriptive Results

It was optimal to switch to PerSec for most subscribers. Over all usage, PerSec was cheaper for 97% of subscribers. Even the charges incurred while subscribed to PerMin would have been cheaper under PerSec for 85% of subscribers.[6] Here I present results that describe how the new plan diffused.

The diffusion process is organic. The time series of switches can reveal more about how information is dispersed. Time limited mass marketing would generate spikes. Figure 3 shows the number of switches from PerMin to PerSec by day, over the course of the data. Takeup is quite rapid upon the introduction of the plan, suggesting that some people were initially informed. This period also coincides with the operator's marketing campaign for the new plan. Otherwise the process has few large spikes, suggesting a more organic diffusion process.

Agents matter. Roughly half of switches occur immediately following a top up. This suggests that the operators' agents are involved in the switching process, either by providing information or by lowering the switching cost k. It is difficult to distinguish between these.

Subscribers settle on the optimal plan, which for most is PerSec. Plans can be switched multiple times by entering a short code, at the user's convenience. We might expect subscribers to switch between plans multiple times for two reasons: subscribers may switch to learn about which plan is best for them, or to reduce costs on calls of a certain type (e.g., one might switch to PerMin before making a long call). The evidence that follows suggests that for the vast majority of subscribers it is better to stick to PerSec than to manipulate the plan choice, and that most subscribers learn this after very few switches.

A not insignificant number of subscribers switch back and forth between plans; those who do mostly settle on PerSec (see Table 2).

Subscribers adjust their behavior in tandem with switching: they make shorter calls when subscribed to PerSec and longer calls when subscribed to PerMin. Figure 4 shows the median

[6] Based on the comparison $c^M(X_i^{*M}) \geq c^S(X_i^{*M})$.

charge ratio by the current plan for subscribers with different sequences of plans. Subscribers make calls that are cheaper under PerSec when subscribed to PerSec (higher ratio), and calls that are cheaper under PerMin when subscribed to PerMin (lower ratio). However, the most striking feature of the graph is that all of the median ratios lie above 1, denoted by the red line: the median subscriber would be better off simply sticking with PerSec rather than engaging in complex switching behavior.

In fact nearly the entire distribution of subscribers would be better off sticking to PerSec than switching back and forth: Figure 5 shows the charge ratio quantiles; again, most quantiles lie above 1 under either plan.

Two features of Table 2 suggest that subscribers do learn that it is better to stick with PerSec: first, that the counts are declining in switches, and second, that most subscribers settle on PerSec (more subscribers starting with PerMin switch an odd number of times, ending up with PerSec, and more subscribers starting with PerSec switch zero or an even number of times).

Peers matter. To analyze the peer effects, I estimate the hazard rate of switching. Specifically I estimate regressions of the following form:

$$Switch_{igt}|NotSwitched_{igt-1} = 1\{\alpha X_{igt} + \beta f(X_{N_{igt}}) + \gamma Y_{gt} + \epsilon_{igt} > 0\}$$

where X_{igt} is a vector of individual characteristics, $X_{N_{igt}}$ is a matrix of characteristics of i's neighbors on the call network, and Y_{gt} is a vector of geographic and temporal controls. I run the regression by month.[7]

Information from mass media will be either sharp in time (radio, television) or sharp in space (billboard), and thus it should be possible to absorb these with temporal and geographic fixed effects (though so far I have not included them at sufficient resolution to completely absorb these).

There are two important potential confounds in estimating peer effects:

Network connections tend to be homophilous: neighbors may take similar actions because they share correlated unobservables rather than due to a peer effect. There are several ways to address this: by controlling for observable characteristics like calling patterns, by looking for flows of information, or by looking at asymmetric edges.

Connected individuals may also be exposed to correlated shocks: we may worry that connected individuals may see the same billboard, or be informed by the same agent. This is more difficult, but could be addressed if the peer effect can be explained through information channels.

See Table 3 for the first set of results, which suggest an increase in the fraction of my social network neighbors who have switched from 0% to 100% is associated with between 38-67 percentage point increase in my hazard rate. I report marginal effects from a logit

[7] I define i and j to be neighbors if there have been at least 3 calls between them over the data set.

regression, including both an aggregate and individual time trend.[8] Column 1 shows that subscribers who would have saved more by switching based on past usage are more likely to switch. Column 2 shows that I am more likely to switch when I have more neighbors who have switched; however, this could be due to learning from peers or due to the confounds mentioned above. We can control for the most obvious form of homophily: I am likely to be clustered with others who make similar types of calls, and thus would obtain similar benefits from switching. Column 3 shows indeed I am more likely to switch when my neighbors would have saved more by switching (based on usage while they are subscribed to PerMin). However, when I include the direct peer effect as well as a control for neighbors' savings, the peer effect barely moves (Column 4). It becomes smaller but still large when including controls for my own savings and time fixed effects (Columns 5 and 6).[9]

These results are consistent with peer effects being an important element in learning about per second billing.

Peers provide information. Information about plans is generated either by the operator (and shared through agents or marketing material), or through direct experience on the network. Since accounts are prepaid, subscribers receive no monthly bill; instead, the precision of the information an individual gets from experience on the network depends on how much feedback he requests from the network. If an individual submits a balance inquiry after every transaction, he can learn from every transaction; if an individual seldom requests balance information, he will have much less precise feedback about the costs of his usage. The model would thus predict that the more precise information that an individual's neighbors generate, the better his decision. This also suggests a placebo test: when the old plan is sufficiently well understood, only the precision of neighbors who have switched to the new plan should matter.

The results of this test are presented in Table 4. Column 1 replicates Column 5 from the first regression table as a baseline. Column 2 introduces a control for how much information is generated by neighbors who have switched, the ratio of balance inquiries to calls. This information control absorbs some of the variation that had been explained by the fraction of neighbors switched, suggesting that at least part of the peer effect could operate through transmission of information learned by experience. As a check, Column 3 adds the corresponding control for neighbors who have not switched (who cannot generate any information about the new plan from their experience) and finds it to be near zero.

Switchers reoptimize immediately for the new pricing schedule. Subscribers' patterns of transactions immediately following a switch yield information about the learning process. If after switching it takes a while before subscribers adjust their behavior, they may be learning about prices under the new plan from their own actions; if they adjust

[8]Standard errors are based on a faulty approximation and should be corrected.

[9]I was not able to complete the regression with geographic controls in time for this draft.

immediately, it suggests that they were already aware of the prices under the new plan prior to switching.

The next few graphs describe transactions immediately preceding and following a switch, for subscribers who switched from PerMin to PerSec once. The pattern of calls sent adjusts immediately following a switch. The top panel in Figure 6 shows that the median call duration drops immediately following a switch, consistent with subscribers reoptimizing behavior for the new plan (since shorter calls are cheaper). A better measure of this change in behavior is the charge ratio, which is shown in the bottom panel of Figure 6; after switching, subscribers change their call behavior towards calls that are cheaper under PerSec (higher ratio). These results suggest that subscribers are aware of the pricing policy under PerSec at the point of switching. This is puzzling: if they were aware of the prices under PerSec they should have switched earlier, since the majority of subscribers would have saved significantly under the new plan. It suggests that the information being shared is not simply that the new plan is cheaper, but a richer description of the billing difference.

As a control, we can also look for breaks in calls received. Since the network is caller pays, my switching has no effect on the billing of calls received, so I wouldn't expect an effect. We do see a gradual transition, in Figure 7. This could be because my switching is correlated with neighbors switching (I have not controlled for that here), or because I exert some influence over the call length.

A further check is the pattern in balance inquiries. If subscribers are learning about prices under the new plan, we would expect to see a jump in balance inquiries immediately following a switch. In Figure 8, looking at the fraction of calls that were preceded by at least one balance inquiry, do we see a jump, but the increase starts before the switch. To get a sense of temporal scale of these graphs, Figure 9 how the transactions are spaced out over days. One explanation for the rise in balance inquiries prior to switching is that calls are more spaced out in time: if subscribers don't compensate by reducing the frequency of balance inquiries, the fraction of calls preceded by a balance inquiry would increase. Then the run up before a switch would be overstated relative to that after the switch.

As a comparison we can look at the behavior of subscribers who switched more times. Figure 10 shows the pattern for subscribers who switched three times from PerMin. We see a similar pattern prior to a switch, but there are fewer balance inquiries when a subscriber returns to a plan they have used before.

7. Robustness

I have a robustness test:

Asymmetry in edges. A significant concern is that edges in the call graph represent homophily– a tendency for nodes to share correlated features. When these features are observed, they can be controlled for, but unobservable features could appear to be peer effects.

The data provides one way to test this, by exploiting the weights and directionality of the edges. Consider two connected nodes on the network. If the nodes equally share the cost of the calls between them (that is, the volume of calls initiated are balanced), this suggests that the nodes share correlated features. Asymmetry in cost sharing suggests less correlation in features. Specifically, we can test whether the learning effects we measure are asymmetric over the two nodes connected by asymmetric links (e.g., is a poor rural farmer more affected by his rich cousin's experience than vice-versa?). A model of learning could result in either asymmetric or symmetric effects across asymmetric links. However, since homophily must be bidirectional, it would predict symmetric effects for two nodes connected by asymmetric links. If the homophily were along the same dimension as measured by cost sharing, asymmetric links should have lower effects for both nodes.

Table 5 shows the results of this test. Column 1 repeats Column 5 from the first regression table as a baseline. Column 2 decomposes the fraction of neighbors who have switched into three elements based on the asymmetry of edges: Duration Ratio represents the ratio of total duration paid for by the neighboring node to the duration paid for by the ego node. More symmetric edges have a higher peer effect coefficient, which could be due to homophily. However, on an asymmetric edge, a switch by the member who pays for more of the calls is associated with a higher likelihood of switching for the other member than the reverse, which is difficult to explain entirely with homophily.

Columns 3 and 4 decompose neighbors' charge ratios in a similar manner, and find a slight asymmetric effect.

This test could be made more robust by varying the cutoffs used to determine asymmetry.

8. Conclusion

Although the diffusion of new technologies are important for economic growth, it has been difficult to study how economic actors learn about these innovations. This paper analyzes the diffusion of one specific technology—a new mobile phone plan—that represented a substantial improvement for most subscribers, using detailed records on usage and communication. I find that diffusion of the new plan took time, but most subscribers eventually settle on the plan that is best for them. The operator's agents are important in the switching process, though it is difficult to distinguish whether they provide information or just help subscribers enter the code needed to switch plans. Subscribers learn from their social network, especially from others that closely monitor usage. Although it would be optimal for most subscribers to switch plans even without adjusting behavior, most subscribers immediately adjust their behavior after switching in the direction that reoptimization would suggest. This suggests that individuals learn the new pricing structure before switching. In a future refinement of this project I aim to finely distinguish the channels through which information about the new plan spreads.

This paper demonstrates techniques for tracing information flows through a communication network. These techniques could easily be combined with small experiments on the network to obtain more robust identification. It is my hope that this line of work can improve our understanding of how information flows through societies and result in policies that improve the targeting of information.

References

Banerjee, A. V. (1992): "A Simple Model of Herd Behavior," *The Quarterly Journal of Economics*, 107, 797–817.

Chandrasekhar, A. and R. Lewis (2012): "Econometrics of Sampled Networks," *Working Paper*.

Conley, T. G. and C. R. Udry (2010): "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, 100, 35–69.

Ellison, G. and D. Fudenberg (1993): "Rules of Thumb for Social Learning," *Journal of Political Economy*, 101, 612–643.

Foster, A. and M. R. Rosenzweig (1995): "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*.

Jackson, M. and L. Yariv (2008): "Diffusion, Strategic Interaction, and Social Structure," in *Handbook of Social Economics (forthcoming)*, Elsevier.

Suri, T. (2011): "Selection and Comparative Advantage in Technology Adoption," *Econometrica*, 79, 159–209.

Zwane, A. P., J. Zinman, E. Van Dusen, W. Pariente, C. Null, E. Miguel, M. Kremer, D. S. Karlan, R. Hornbeck, X. Giné, E. Duflo, F. Devoto, B. Crepon, and A. Banerjee (2011): "Being surveyed can change later behavior and related parameter estimates," *Proceedings of the National Academy of Sciences*.

Appendix

Appendix A. Estimating missing cell tower locations

We have a set of cellular towers whose coordinates are known, K, and a set whose coordinates are unknown and to be estimated, U. We also have usage data from the mobile phone operator which references all towers. Specifically, let us consider using anonymized call detail records (CDRs) which list the tower used at the beginning and end of a transaction for both the sending and receiving phone.

One way to infer the locations of the missing towers is to take advantage of calls that were handed off from one tower to another during a call. This can happen if a person moves during a call, or if a tower is overloaded. If many calls of short duration are handed off from tower X to tower Y, this suggests that X and Y are near each other. Thus, one straightforward way to infer the missing tower locations is to perform a weighted sum of the coordinates of known towers, where weights are derived from the nearness of the towers along a metric.

More formally, we predict the coordinates of an unknown tower $x_u = (x_u^{\text{long}}, x_u^{\text{lat}})$ by computing a weighted average of the coordinates of the set of known towers K, with weights given by the standardized metric w_{ku} of the relationship between k and u:

$$\hat{\mathbf{x}} = \sum_{k \in K} w_{ku} \mathbf{x_k}$$

A simple metric based on the number of handoffs between towers k and u works quite well[10]:

$$w_{ku} = \frac{N_{ku}^{\text{Handoffs}}}{\sum_{j \in K} N_{ju}^{\text{Handoffs}}}$$

To gauge the precision of tower estimates, I estimate the locations of known towers using leave-one-out sampling. Results are summarized by tower density in Figure 1.

The mean error in predicted location is 7.0 km, with a standard deviation of 12.3 km. The cluster of towers with densities around 140 represent towers in the capital city.

The method lacks a force that pushes predicted locations away from other towers; it might be possible to address this with a correction factor.

Appendix B. Inferring plan details and choices

The operator reported plan charges on its website as well as to the regulator; however, there are some discrepancies between these reported charges and the charges actually incurred. To infer the incurred plan charges, I use the following process:

[10]Other metrics could certainly be used, such as the distribution of lengths of calls that were handed off (when the duration of a call is short, it is less likely that a call was handed off due to travel). In a first pass I found that this information was less useful than the raw number of handoffs, presumably because even short calls can be handed off for load balancing at long distances (up to 35 km for GSM networks).

FIGURE 1. The method performs poorly in areas of low tower density, but becomes more precise as tower density increases.

(1) Plan and timeband information gathered from archive.org historical scrapes of operator's website and regulator's annual reports
(2) Plans and timebands are assembled into a model of call charges. Most plans have unique charges.
(3) The model is run against call charge data
    (a) The current plan is inferred from the charge
    (b) Unclassified charges, and switches between plans are noted
    (c) Patterns in unclassified charges and plan switches are explored to identify inaccuracies. These are corrected and the process is iterated.

Using this process I am able to fit 93% of charges to plans. I do not have direct evidence on plan switches, and so infer a switch whenever a charge does not match the current plan but does match another plan available at the time. Since most plans have unique charges, inferred switches should match actual switches in the vast majority of cases.

Table 1. Household Characteristics (Nationally Representative)

	All Households	Households with Mobile Phones
Fraction of Households	1.00	0.05 (2005), 0.40 (2010)
Consumption per capita (real)	$264.81	$925.14
Household size	6	7.5
Phones per household	0.1	1.3
Rural	0.83	0.30 (2005), 0.57 (2008)
Has radio	0.58	0.90
Has electricity	0.06	0.39
Has television	0.03	0.24

	All Individuals
Newspaper circulation	0.01
Internet users	0.02

Sources: EICV 2005-2006 (N=6,900), 2010-2011 (N=7,354), National Institute of Statistics Rwanda, Financial Survey 2008, Demographic and Health Survey 2007, Media Sustainability Index, and World Bank/ITU.

Table 2. Number of Switches by Initial Plan

Initial Plan	Switches	Number of Subscribers
PerMin	0	88,226
	1	**283,510**
	2	10,766
	3	**44,643**
	4	2,883
	5	**15,427**
	6 or more	*42,640*
PerSec	0	**46,727**
	1	348
	2	**3,032**
	3	54
	4	**692**
	5	16
	6 or more	*1,026*

Numbers in **bold** indicate final plan was PerSec.

Table 3. Determinants of Switching

	$Switch_{ijt}$ conditional on not switching prior					
	1	2	3	4	5	6
log(Charge Ratio PerMin/PerSec)	0.392***				0.301***	0.216***
	(0.001)				(0.001)	(0.001)
Fraction of Neighbors Switched		0.671***		0.667***	0.375***	0.481***
		(0.002)		(0.002)	(0.002)	(0.002)
log(Median Neighbor Charge Ratio PerMin/PerSec)			0.568***	0.424***	0.162***	0.154***
			(0.006)	(0.005)	(0.005)	(0.004)
Years since start of data	-0.100***	-0.252***	-0.082***	-0.255***	-0.196***	
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	
Years of experience	0.120***	0.163***	0.159***	0.158***	0.126***	0.104***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	-0.166***	-0.216***	-0.164***	-0.240***	-0.208***	-0.250***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Fixed Effects						Month x Year
McFadden R-sq.	0.289	0.249	0.195	0.252	0.304	0.343
N	2,889,109	2,889,109	2,889,107	2,889,107	2,889,107	2,889,107

Notes

Table 4. Determinants of Switching

	$Switch_{ijt}$ conditional on not switching prior		
	1	2	3
log(Median Switched Neighbor Ratio Balance Inquiries/Calls)		0.145***	0.142***
		(0.001)	(0.001)
log(Median Unswitched Neighbor Ratio Balance Inquiries/Calls)			0.007***
			(0.000)
log(Charge Ratio PerMin/PerSec)	0.301***	0.278***	0.278***
	(0.001)	(0.001)	(0.001)
Fraction of Neighbors Switched	0.375***	0.280***	0.276***
	(0.002)	(0.002)	(0.002)
log(Median Neighbor Charge Ratio PerMin/PerSec)	0.162***	0.117***	0.117***
	(0.005)	(0.005)	(0.005)
Years since start of data	-0.196***	-0.178***	-0.175***
	(0.001)	(0.001)	(0.001)
Years of experience	0.126***	0.118***	0.118***
	(0.000)	(0.000)	(0.000)
Intercept	-0.208***	-0.224***	-0.226***
	(0.001)	(0.001)	(0.001)
McFadden R-sq.	0.304	0.314	0.314
N	2,889,107	2,889,107	2,889,107

Notes

TABLE 5. Determinants of Switching: Asymmetry

	$Switch_{ijt}$ conditional on not switching prior			
	1	2	3	4
#Neighbor[Duration Ratio < 0.5].Switched / #Neighbor		0.342***		0.317***
		(0.002)		(0.002)
#Neighbor[0.5 <= Duration Ratio <= 2].Switched / #Neighbor		0.427***		0.377***
		(0.003)		(0.003)
#Neighbor[Duration Ratio > 2.0].Switched / #Neighbor		0.374***		0.343***
		(0.002)		(0.002)
log(Median Neighbor[Duration Ratio < 0.5] Charge Ratio PerMin/PerSec)			0.084***	0.079***
			(0.004)	(0.004)
log(Median Neighbor[0.5 <= Duration Ratio <= 2] Charge Ratio PerMin/PerSec)			0.076***	0.078***
			(0.003)	(0.003)
log(Median Neighbor[Duration Ratio > 2.0] Charge Ratio PerMin/PerSec)			0.065***	0.069***
			(0.003)	(0.003)
log(Charge Ratio PerMin/PerSec)	0.301***	0.297***	0.300***	0.298***
	(0.001)	(0.001)	(0.001)	(0.001)
Fraction of Neighbors Switched	0.375***		0.342***	
	(0.002)		(0.002)	
log(Median Neighbor Charge Ratio PerMin/PerSec)	0.162***	0.166***		
	(0.005)	(0.005)		
Years since start of data	-0.196***	-0.195***	-0.187***	-0.187***
	(0.001)	(0.001)	(0.001)	(0.001)
Years of experience	0.126***	0.127***	0.125***	0.125***
	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	-0.208***	-0.208***	-0.205***	-0.205***
	(0.001)	(0.001)	(0.001)	(0.001)
McFadden R-sq.	0.304	0.305	0.300	0.300
N	2,889,107	2,889,107	2,803,767	2,803,767

Notes

FIGURE 1. Price schedule and density of observed calls by plan type, 1.2005-1.2008

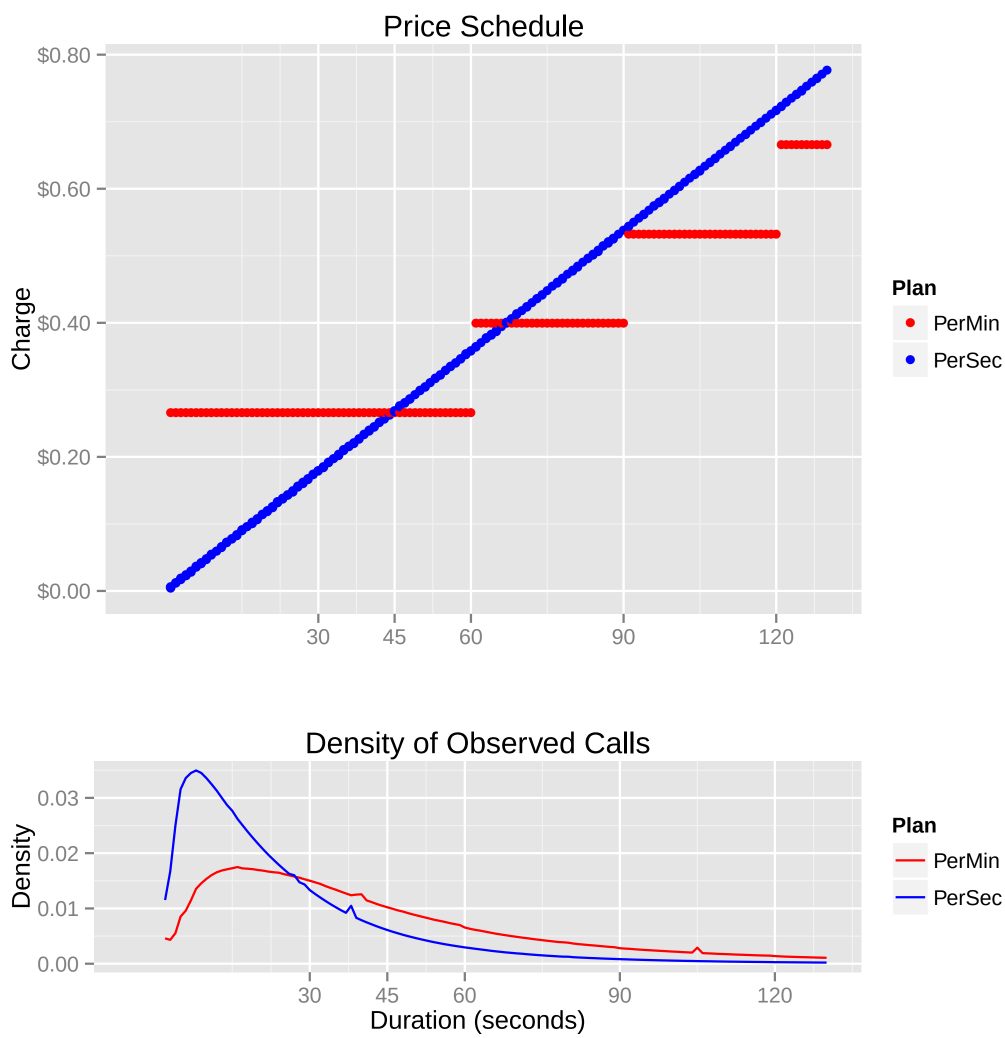


Figure 2. Subscribers by plan and date

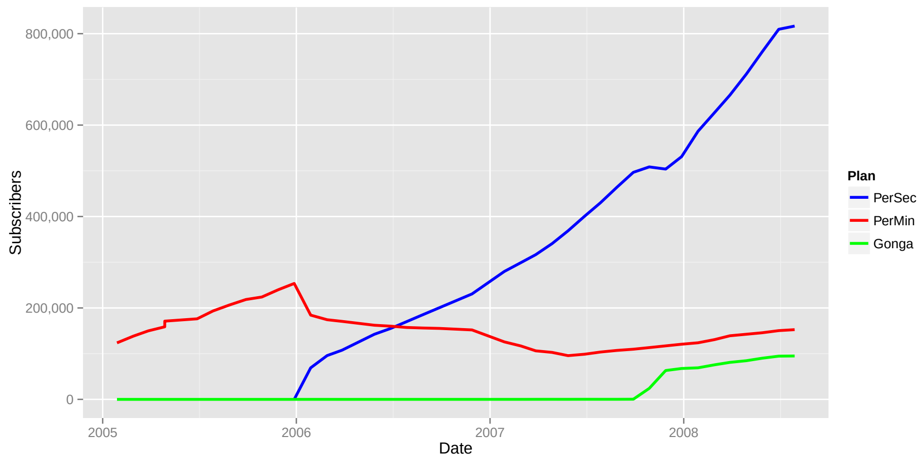
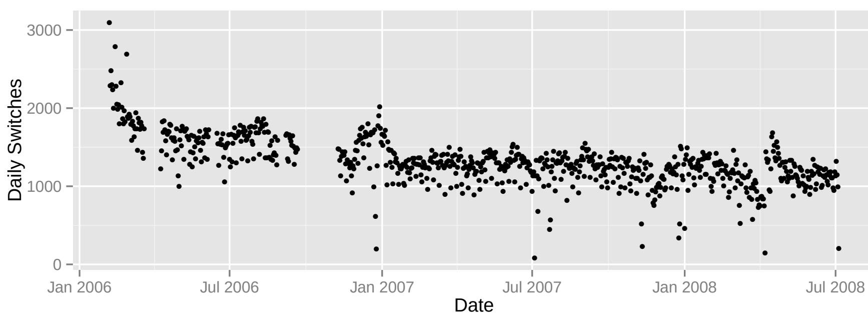
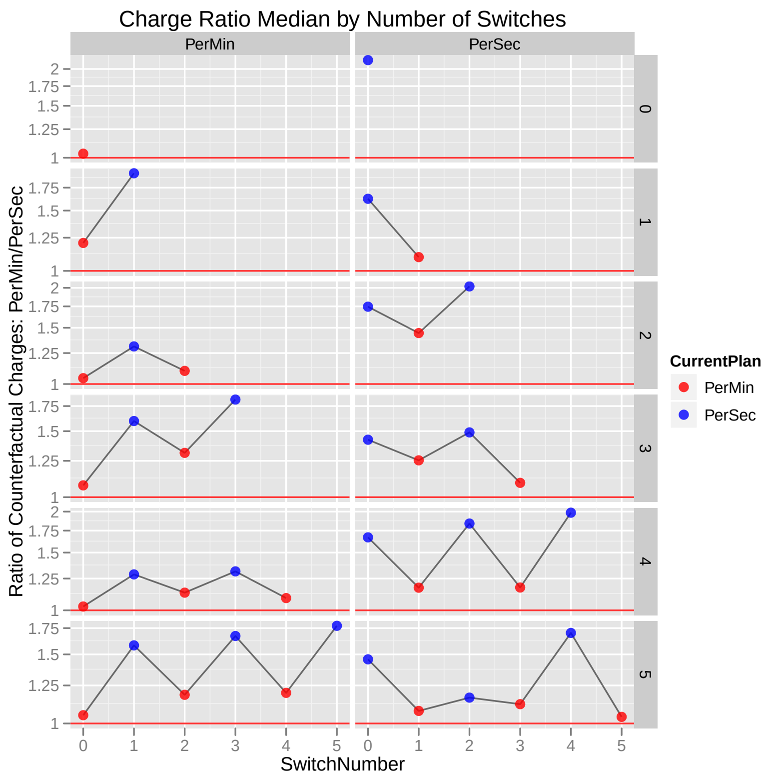


Figure 3. Time series of switches from PerSec to PerMin



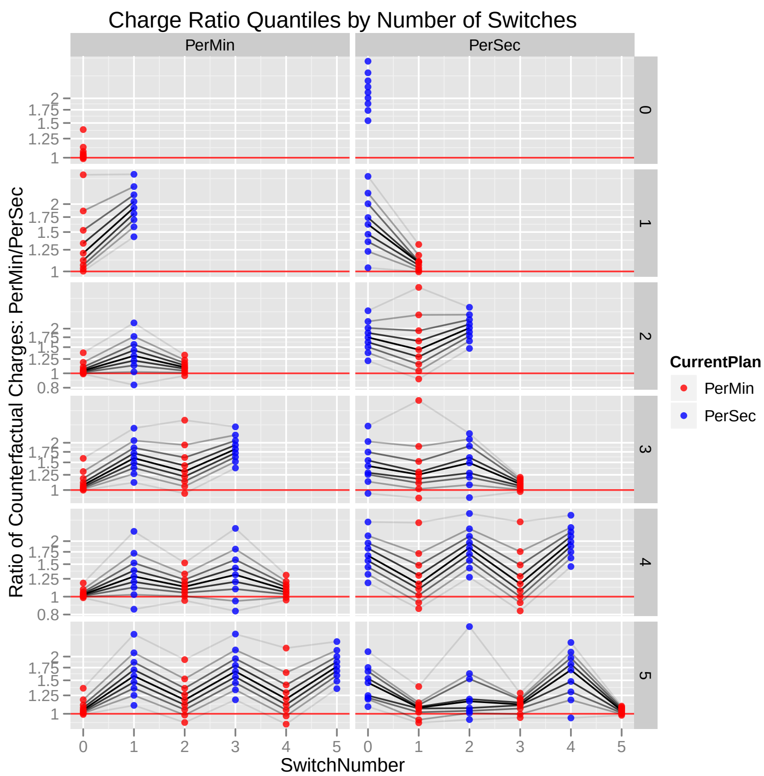
The week following a period of missing data is omitted, since switches occurring during the missing data would be miscoded as occurring at the next transaction, which would lead to artificial spikes in switches.

Figure 4. Switch Patterns: Median Charge Ratio by Current Plan



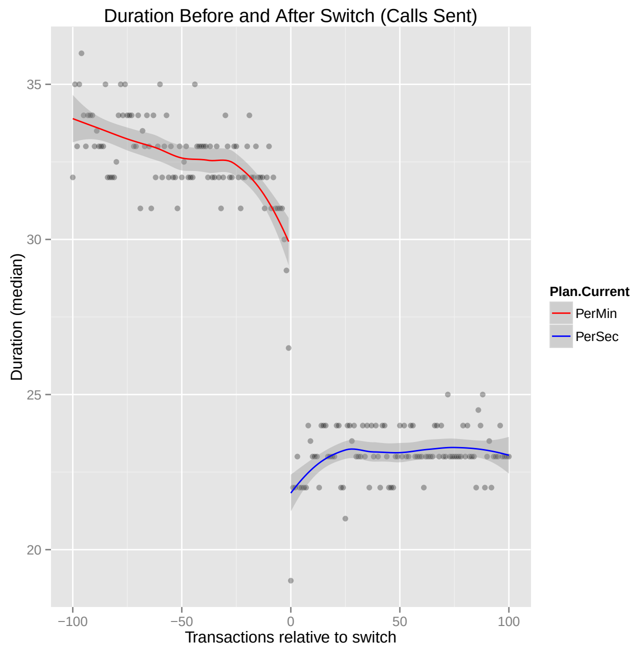
A separate graph is drawn for subscribers of different initial plans and total number of switches. Subscribers change behavior upon switching, but PerSec is still optimal for the median.

Figure 5. Switch Patterns: Distribution of Charge Ratios by Current Plan

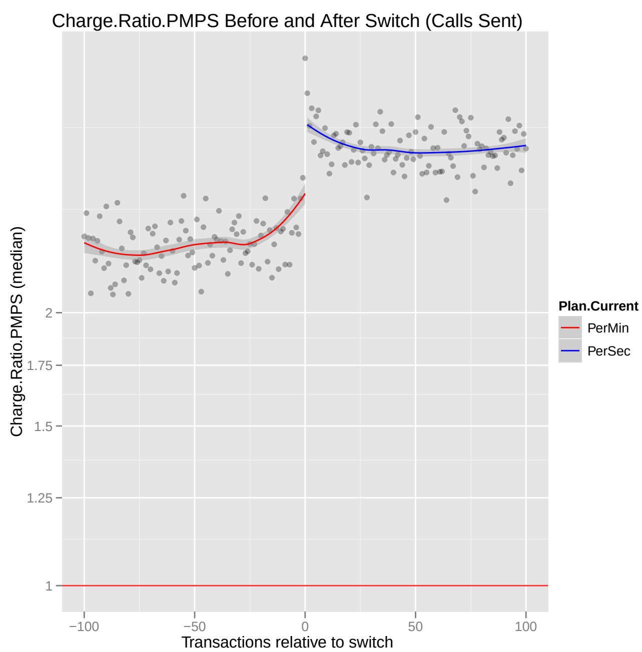


Median is shown in black; other deciles are shaded in gray.

FIGURE 6. Subscribers immediately adjust their calling to take advantage of the different prices under PerSec



(a)



(b)

FIGURE 7. Calls received change also as a subscriber switches

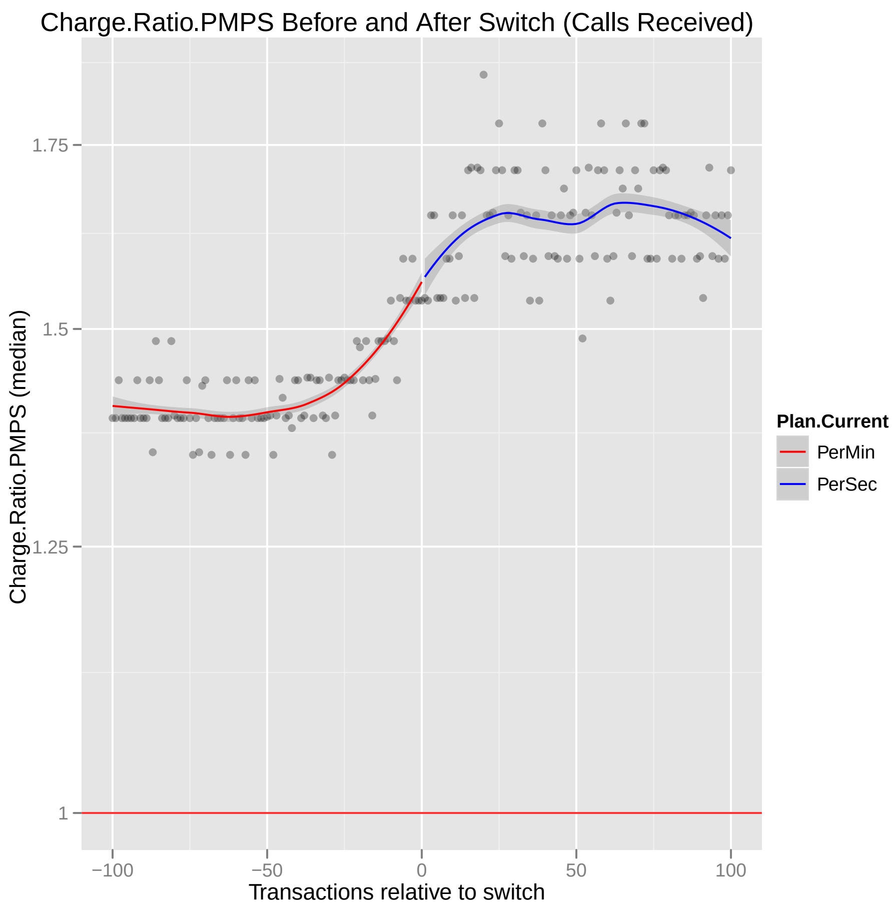


Figure 8. Subscribers check their balance frequently as they are switching plans

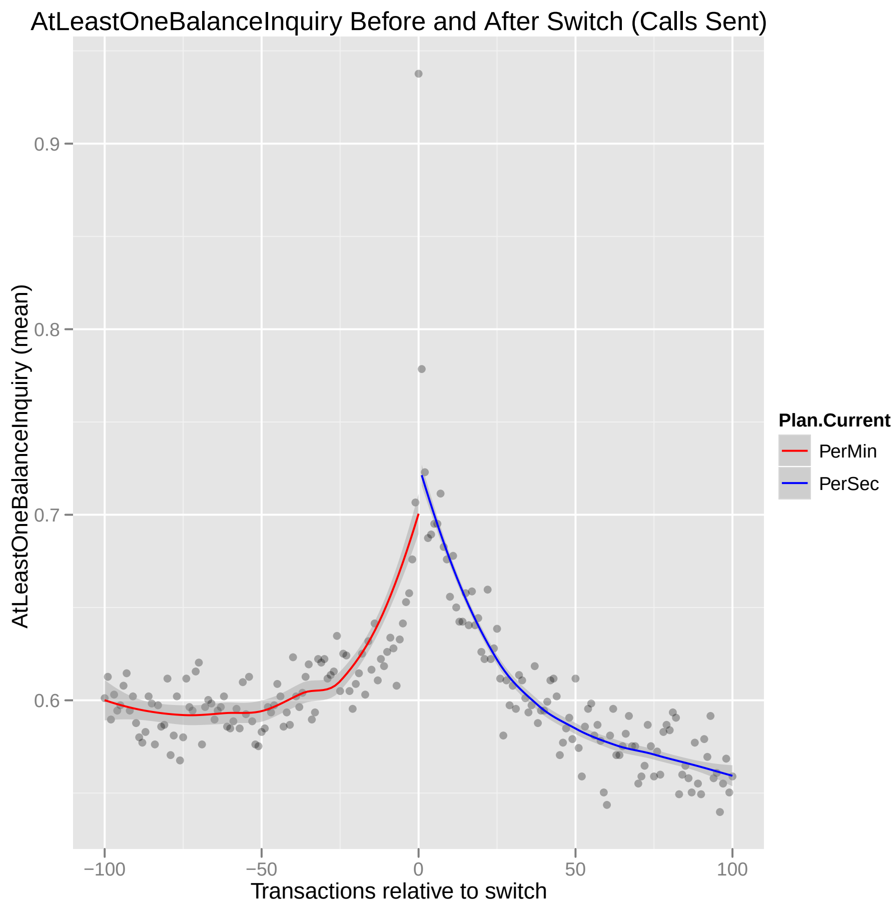


Figure 9. Calls occur at a higher frequency following a switch

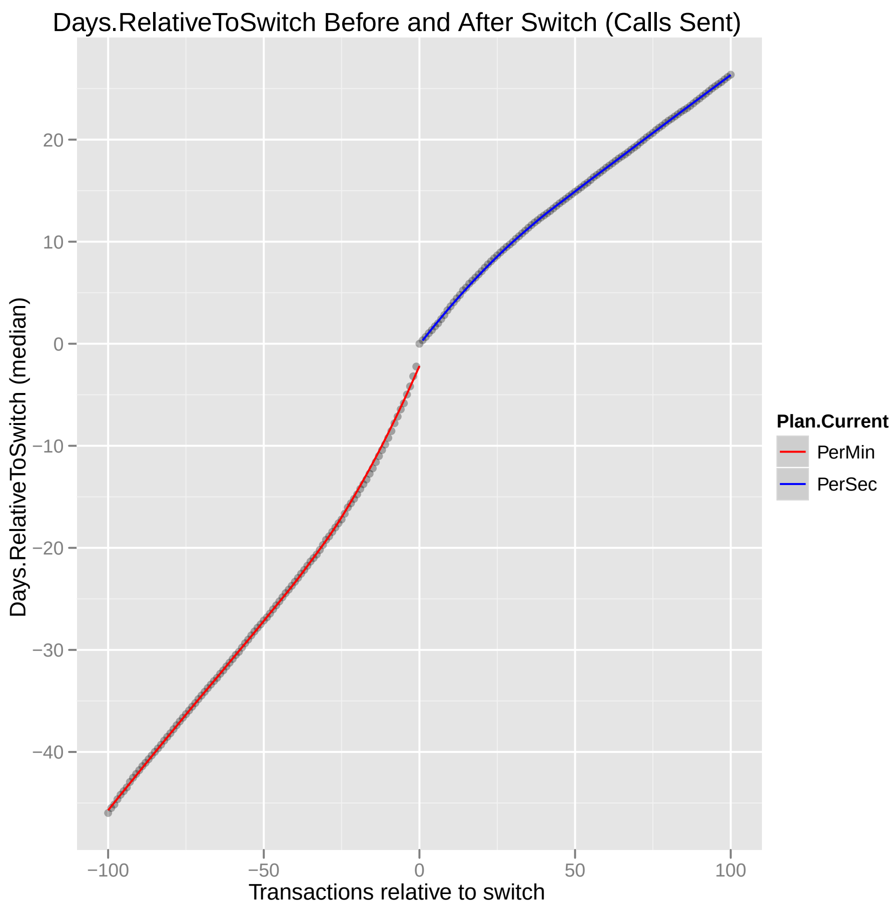


FIGURE 10. Pattern of balance inquiries among subscribers who switched three times, from PerMin