

**THE ADOPTION OF NETWORK GOODS:
EVIDENCE FROM THE SPREAD OF MOBILE PHONES IN RWANDA**

DANIEL BJÖRKEGREN*

BROWN UNIVERSITY

This paper develops a method to estimate and simulate the adoption of a network good. I estimate demand for mobile phones as a function of individuals' social networks, coverage, and prices, using transaction data from nearly the entire network of Rwandan mobile phone subscribers over 4.5 years. Because subscribers pay on the margin, the calls placed reveal the value of communicating with each contact. This feature allows me to overcome traditional difficulties in measuring network effects, by estimating the utility of adopting a phone based on its eventual usage. I use this structural model to simulate the effects of three governmental policies. An adoption subsidy had a high social rate of return, and spillovers accounted for a substantial fraction of its impact. A requirement to serve rural areas lowered the operator's profits but increased net social welfare. Taxing telecom represents a public finance opportunity for developing countries but welfare costs are underestimated when network effects are ignored. Shifting from adoption to usage taxes would have substantially increased consumer surplus for most users.

JEL CLASSIFICATION CODES: O33, L96, O180, L51

KEYWORDS: network goods, infrastructure, information technology

*E-mail: danbjork@brown.edu, Web: <http://dan.bjorkegren.com>

Revision November 23, 2015. Preliminary and incomplete.

I am grateful to Michael Kremer, Greg Lewis, and Ariel Pakes for guidance and encouragement. Thank you to Nathan Eagle for providing access to the data, computing facilities, and helpful conversations. For helpful conversations I am also grateful to Natalie Bau, Arun Chandrasekhar, Brian Dillon, Michael Dinerstein, Ben Golub, Rick Hornbeck, Sabrina Howell, Max Kasy, Divya Kirti, Daniel Pollmann, Martin Rotemberg, Heather Schofield, Jesse Shapiro, Mark Shepard, Rob Stewart, Laura Trucco, and Tom Zimmerman. In Rwanda, I thank the staff of my telecom partner, RURA, RDB, CGIS-NUR, NISR, Minicom, PSF, EWSA, and BRD for helpful conversations and

1. INTRODUCTION

Many modern goods are network goods, whose benefits depend on the network of other users. These include technologies for communication (such as telephones, e-mail, and social networks), payment (digital wallets, mobile money), platforms (office productivity software), and systems that learn from their users (recommendation systems, crowdsourcing). While these goods can generate large efficiency gains (Jensen, 2007; Jack and Suri, 2014), their allocations are likely to be inefficient. Individuals are unlikely to internalize all the benefits their adoption generates, so adoption is likely to be suboptimal unless the firms operating the network use sophisticated pricing mechanisms.¹ Also, if markets are competitive and standards are compatible, any single firm will internalize only a small share of the benefits it generates. If instead a market is so concentrated that these benefits are internalized by a small number of firms, the ability of these firms to exert market power raises standard welfare concerns.

Firms and governments use many different policies to guide the provision and adoption of network goods. While theoretical work provides intuition about network effects, there is little empirical work to guide policy choices.² Empirical work has been limited for three reasons. It is costly to measure an entire network using traditional data sources. It is also difficult to identify network effects: one individual may adopt after a contact adopts because the contact provides network benefits, or because connected individuals share similar traits or are exposed to similar environments. And even if these two issues are overcome, it is difficult to evaluate policies, which can cause effects to ripple through the entire network. As a result, there remain open

access to data. This work was supported by the Stanford Institute for Economic Policy Research through the Shultz Fellowship in Economic Policy. Supplemental Appendix available from the author upon request.

¹An individual's adoption benefits immediate contacts because it makes it possible for them to interact using the good. Adoption also makes these contacts more likely to adopt, and thus benefits the contacts of contacts. These benefits ripple through the entire network of potential users, and are unlikely to be internalized by the initial adopter.

²Early theoretical work includes Rohlfs (1974), Katz and Shapiro (1985), and Farrell and Saloner (1985). Most empirical work on network goods measures the extent of network effects; see for example Saloner and Shepard (1995), Goolsbee and Klenow (2002), and Tucker (2008). The paper closest in spirit to this one is Ryan and Tucker (2012), which estimates the adoption of a videoconferencing system over a small corporate network, and evaluates policies of seeding adoption.

questions about how to design policies that better capture the spillover benefits associated with network effects, as well as policies that overcome suboptimal provision arising from high concentrations in industries providing network goods.

This paper overcomes these limitations by combining a new empirical approach with rich data from nearly an entire country's remote communication system. I use 5.3 billion transaction records from Rwanda's dominant mobile phone operator, which held over 88% of the market, during a period of dramatic expansion.³ I estimate a structural model of demand for mobile phones, and then demonstrate how this model can be used to simulate the effects of policies.

My empirical approach has three parts:

First, acknowledging that the utility of owning a mobile phone is derived from its usage, I model the utility of using a phone. In the data I use, I observe every connection between subscribers, as well as the calls placed across each connection. This allows me to overcome two fundamental identification problems associated with network effects. I overcome simultaneity in consumer adoption decisions by inferring the value generated by each connection from subsequent interaction across that connection. This is similar to Ryan and Tucker (2012) who infer the value of a videoconferencing link by the number of calls placed across it. However, in that setting individuals bear no monetary cost, while in the Rwandan system I study, 99% of accounts are prepaid: the person placing a call pays for it on the margin, by the second.⁴ This provides a direct measure of value: a subscriber must value a connection at least as much as the cost of calls placed across it.⁵ I also overcome simultaneity between consumer and firm decisions. In Ryan and Tucker (2012) firm policies are static, but in practice and in many models, firm policies change as the network expands. For instance, as the Rwandan network expanded, the marginal adopter became poorer and more remote, and the firm lowered prices and increased the quality of rural service. This variation makes it possible to identify the underlying demand curve for communication

³This concentration appears to result from a regulator restriction on entry to two firms rather than other features of this setting: the regulator has since allocated more licenses and the dominant operator's market share has declined to 54% (RURA, 2013).

⁴In the first 14 months of the data, calls are billed by the first minute and every following 30 seconds.

⁵In contrast, most empirical studies of network goods use coarse measures of the value of joining the network; exceptions that use individuals' local network are Tucker (2008) and Birke and Swann (2010).

across each link as a function of prices and coverage, but introduces a simultaneity problem: consumer responses to changing firm policies tend to be confounded with the changing composition of consumers. My method separates these effects by using within-link variation, tractably estimating 415 million link fixed effects.

Second, I model the decision to adopt a mobile phone. The utility of having a phone in a given period is given by the utility of communicating with contacts that have phones. Consumers choose when to adopt by weighing the increasing stream of utility from communicating with the network against the declining cost of purchasing a handset.⁶ This model allows me to compute the utility an individual would have obtained if he had adopted at a different time under the observed adoption sequence, but also the utility he would have obtained had the rest of the network adopted in a different order.

Third, to evaluate the impact of policies, I use a simulation method that allows each individual to react directly to a policy change, and to each other's responses, capturing effects that ripple through the network and across physical space. An equilibrium in this context must reconcile nearly 1 million interconnected adoption decisions. I make simulation tractable by defining an equilibrium in publicly announced adoption dates; I then bound the full set of equilibria by exploiting the supermodularity of the adoption decision, in a manner similar to Jia (2008). I turn this method to policy questions facing developing countries.

The spread of mobile phones across the developing world has been dramatic: between 2000 and 2011, the number of mobile phone subscriptions in developing economies increased from 250 million to 4.5 billion (ITU, 2011). Improvements in communication, through mobile phones as well as associated services such as mobile money and mobile internet, have the potential to knit even remote villages into the global economy. But in addition, these technologies are easily taxed and thus represent a public finance opportunity: the mobile industry contributed an average of 7% of government revenues in sub-Saharan Africa as early as 2007 (GSMA, 2012). Developing countries thus face a tension between generating revenue and extending service, particularly to rural and low income areas ('a paramount concern' in the words of

⁶The utility of having a phone increases as coverage improves, calling prices are reduced, and contacts join the network.

former World Bank ICT Director Mohsen Khalil). Governments typically manage this tension with a set of telecom-specific taxes, and regulations and programs that encourage access to the rural poor. However, there is little evidence to guide the design of these policies, and standard approaches that do not account for network effects can give misleading estimates.

I use my approach to evaluate three policies.

Because individuals tend not to internalize adoption spillovers, it is common for firms or governments to subsidize adoption of network goods. I analyze a rural adoption subsidy program implemented by the Rwandan government in 2008, using the simulation method to determine how the policy affected the entire network. I find that a substantial fraction of the subsidy's impact arises from its impact on nonrecipients, who account for more than 63% of the effect on revenue. Although the bounds are wide, the subsidy improved welfare, in a low case by \$301,980 (representing a social return of 53%), and in a high case by \$4.9 million (a social return of 855%).⁷

I also analyze the welfare implications of providing coverage to rural areas. A social planner would expand coverage until the point where building any marginal set of towers would not improve welfare. Firms may stop building before reaching this point: in a competitive market, some of the benefits of expanding coverage will spill over into competitors' networks.⁸ And regardless of market structure, firms are unlikely to internalize all of the value generated for consumers: price discrimination is limited practically, and often also by regulation. Depending on the shape of private and social benefits from expansion, it may be optimal for a government to require the provision of coverage to areas that are unprofitable to serve. I find that in Rwanda, a government coverage obligation led to the building of roughly 6% more rural towers that were unprofitable for the firm but slightly welfare improving for the country.

Finally, I use my model to evaluate the potential of telecom taxation to generate government revenue. I compute a social welfare improving tax policy in a class of instruments that includes dynamic (time-varying) taxes on adoption and usage, for a

⁷Due to the cost of computing an equilibrium with interdependent demand, I report impacts as changes in the bounds of outcomes rather than bounds on the changes. I discuss this further in Section 7.

⁸A fraction of these benefits can be internalized using interconnection fees, but some will spill into the interiors of competitor networks.

given revenue requirement. I find that baseline tax regime had a substantial welfare cost that would be underestimated if network effects were ignored. I find that had the government shifted taxes to usage rather than adoption it could have more than doubled the consumer surplus accruing to the bulk of users, while raising the same amount of government revenue. The previous two policy simulations make conditions for adoption less favorable, so that the individuals adopting in my simulations will be a subset of those I observe in the data. The taxation simulation makes conditions more favorable, so that the true benefits of switching policies will be larger than what I estimate.

Although the method I present uses network structure revealed by adoption and usage, it can also be applied to goods that have yet to diffuse. An analyst can gather data about the adoption of a good from a context where exogenous factors have induced adoption to be high, simulate the effects of policies, and use the conclusions to inform policy in a context with lower adoption. As an example of this strategy, I exploit the fact that Rwandan government regulations resulted in most of the country receiving cellular coverage to predict the effects of expanding coverage as a function of population density. An analyst can also gather data about a good that has already diffused on a network of interest, and then simulate the adoption of a good that has yet to diffuse. As an example of this strategy, I describe how the adoption of mobile phones can inform policy for mobile internet service.

This paper connects with several literatures:

This paper studies classic network goods, whose value depends directly on the network of other users. Several empirical studies have measured the existence or extent of network effects in various environments (for example, Brynjolfsson and Kemerer (1996); Goolsbee and Klenow (2002); Akerberg and Gowrisankaran (2006); Tucker (2008)). Conceptually related are goods with indirect network effects, whose value may increase with additional users not because they provide direct benefits, but because they induce a response from the other side of the market. There has been more substantial empirical work on indirect network effects; related papers include Ohashi (2003) on video cassette recorders, Lee (2013) on video game platforms, and Gowrisankaran et al. (2010) on DVD players.

The diffusion of technologies is essential for the productivity of developing economies. While many studies have explored aggregate trends in adoption or individual adoption for a sample of users, this study models how nearly an entire network of users adopts a technology with rich data on how the technology is ultimately used.⁹ The paper also connects to a literature analyzing the rapid spread of information and communication technologies (ICTs) across developing countries (Aker and Mbiti, 2010).

A growing literature analyzes the impact of social networks on economic behavior (see Jackson, 2009). My paper is conceptually related to Banerjee et al. (2013), which estimates and simulates the diffusion of microfinance over a network following an injection of information. While the authors primarily model the transmission of information about a good over a social network, I model the adoption and subsequent usage of a good whose benefits are derived from the network itself.

This paper also contributes to an emerging literature that uses passively collected transaction records to analyze developing economies. These records overcome some limitations of traditional sources of data (e.g., Zwane et al., 2011), and can also answer questions that could not be answered with equivalent data from a developed country. In developed economies, transaction data from any one source typically represents only a small slice of an agent's economic activities because agents generally face many alternatives.¹⁰ Within a developing economy, a single data source can be comprehensive: in Rwanda during the period of interest, records from a single mobile phone operator represent the vast majority of remote communication.

The next section describes the expansion of mobile phone networks worldwide and in Rwanda. Section 3 describes the data. Section 4 presents stylized facts about mobile phone usage in Rwanda. Section 5 introduces a model of phone adoption and usage. Section 6 describes the procedure I use to estimate the parameters of this model and the country's communication graph. Section 7 introduces a method that is used to simulate the effects of two counterfactual policies: Section 8 analyzes the

⁹See, for example: Griliches (1957); Foster and Rosenzweig (1995); Conley and Udry (2010); Comin and Hobijn (2010).

¹⁰For example, the full remote communication behavior of a consumer in the U.S. may be spread over mail, home e-mail, work e-mail, a work phone, a personal mobile phone, fax, chat, Skype, Facebook, and other more specialized channels. Even complete data from any one of these channels will be heavily selected and difficult to interpret.

effect of an adoption subsidy, and Section 9 analyzes operator incentives to provide service in rural areas. Section 11 concludes.

2. CONTEXT

The expansion of mobile phone networks across the developing world has had several common features. Initial networks were built in cities and served elites. Handset prices were initially expensive, but fell dramatically with reductions in component costs and economies of scale, making phones accessible to poor consumers. Operators adapted to this broader base of potential subscribers by expanding coverage beyond urban centers and reducing usage prices. The empirical strategy presented in this paper will disentangle the impact of these factors for the spread of mobile phones in Rwanda, and simulate the spread under alternate scenarios.

Rwanda between 2005-2009 is an attractive setting to study the spread of mobile phones in developing countries. Because the Rwandan regulator restricted entry, the market during this period was extremely concentrated: the mobile operator whose data I use held above 88% of the market, and its records reveal nearly the entirety of the country's remote communication. There are few alternatives for remote communication: the fixed line network is small (with penetration below 0.4%), and mail service is insignificant.¹¹ The data on which this project is based is long enough to capture both adoption and use decisions for a substantial fraction of the population, as well as substantial variation in prices and provision of service.

Rwanda. Rwanda is a small, landlocked country in East Africa. It is predominantly rural; most households live off of subsistence farming. The country's experience with mobile phones is similar to that of other sub-Saharan African countries, apart from three main differences. First, Rwanda is less developed than the African average and most of its neighbors: per capita consumption in 2005 was \$265, while the World Bank reported a sub-Saharan African average of \$545 (WDI, 2013). Second, it has two opposing features that affect the profitability of building a mobile phone network:

¹¹The average mail volume per person was 0.2 pieces per year in Rwanda, relative to 2.4 pieces in Kenya and 538.8 pieces in the US (Sources: National Institute of Statistics Report 2008, Communications Commission of Kenya, U.S. Postal Service 2011, U.S. Census).

it is very hilly, which interferes with signal propagation, but it also has a high population density, which allows each tower to cover more potential subscribers. Third, the Rwandan market was slow to develop competition, due to fewer licenses being allocated by the regulator and initial snags in the performance of the second licensee. During the period of limited competition, prices were relatively high and penetration was relatively low.

Network Rollout. In combination with other reconstruction efforts after the 1994 Rwandan Genocide, the new government spurred the development of a mobile phone network. An exclusive license was given to a multinational operator, which started operations in the capital, Kigali, in 1998. Service quickly spread from Kigali to other urban centers. Two changes influenced further rollout.

Global handset prices began to decline, making mobile phones accessible to larger segments of the population. In 2005, the cheapest mainstream handset in Rwanda cost roughly \$70, or three and a half months of the mean person's consumption; by 2009 handsets were available for \$20.

Regulatory changes induced a change in market structure. In 2003, the government announced it would provide a license to a second operator, which entered the market in 2005. This second operator was not very successful: it reached a maximum of 20% of market share for a brief period after the end of my data, and in 2011 its license was revoked for failure to meet obligations. In combination with providing a second license, the government attached minimum coverage obligations to the first operator's license.¹²

The dominant operator changed pricing structures to accommodate lower income users and expanded into rural areas. At the beginning of 2005, holding an account on the dominant operator in Rwanda entailed paying a monthly access fee of roughly \$2, paying a minimum of \$0.27 per call, and topping up a minimum of \$4.53 when credit ran low.¹³ By the middle of 2008, essentially all nonmarginal charges had been removed, talk time was billed by the second, and the minimum top up amount

¹²A third operator entered the market at the end of 2009 and has been quite successful, taking a third of the market by 2012.

¹³Prepaid balances must be refilled or 'topped up' when are depleted in order to continue making calls.

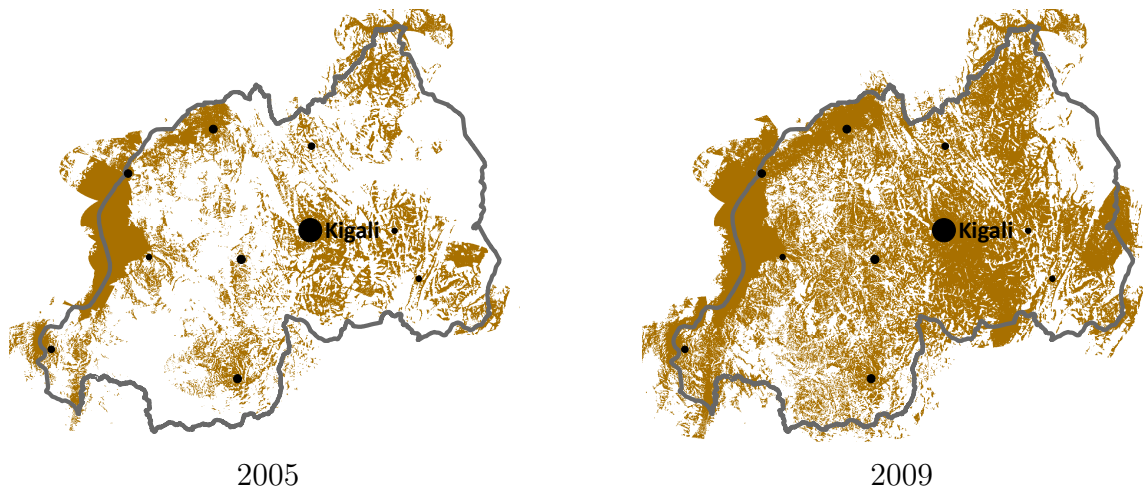
TABLE 1. Household Characteristics (Nationally Representative)

	All Households		Households with Mobile Phones	
	2005	2010	2005	2010
Consumption per capita (real)	\$264.81	\$288.06	\$925.14	\$429.77
Monthly spending on airtime	-	\$2.65	-	\$5.75
Rural	0.85	0.86	0.23	0.75
Has electricity	0.05	0.10	0.62	0.22
Owns fixed line phone	0.008	0.003	0.14	0.007
Owns mobile phone	0.05	0.40	1.00	1.00
Owns radio	0.46	0.63	0.93	0.84
Owns television	0.02	0.05	0.41	0.12
Proportion of households	1.00	1.00	0.05	0.40

Sources: consumption: EICV 2005-2006 (N=6,900), 2010-2011 (N=7,354), National Institute of Statistics Rwanda; remainder of rows: DHS 2005 (N=10,272) and 2010 (N=12,540). Nationally representative sampling weights applied. Consumption per capita deflated to January 2006 prices; the deflator in 2010 was 1.42. Dash indicates that that question was not asked.

was reduced to \$0.90. From 2005 to 2009, the number of cell towers tripled, and the fraction of the country's land area with coverage increased from 60% to 95%, as shown in Figure 1. Reduced prices and improved coverage induced rural and poor households to adopt. Although 85% of Rwandan households live in rural areas, in 2005 only 23% of households with mobile phones were rural; by 2010, 75% were. In 2005 households with mobile phones had a mean consumption per capita of 3.5 times the average; by 2010 the mean consumption of phone owning households was 1.5 times the average. Table 1 shows the baseline characteristics of the Rwandan population and these changing demographics of phone owners, and Table 2 shows usage patterns. Figure 2 shows the trend of mobile and fixed line telephone subscriptions. Figure 3 shows the changes in prices, coverage, and network adoption.

FIGURE 1. Coverage



Locations with coverage are shaded. Cities are denoted by points sized by population.

3. DATA

This project uses several data sources:¹⁴

Call detail records: As a side effect of providing service, mobile phone operators record data about each transaction, called Call Detail Records (CDRs). This project uses anonymous call records from the dominant Rwandan operator, which held above

¹⁴More details about measurement are described in the Supplemental Appendix.

TABLE 2. Usage Profile (1.2005-7.2008)

	Monthly Usage			Charge per Transaction
	Median	Mean	S.D.	Median
Calls	9.4	56.5	114.7	\$0.10
Missed calls	40.8	187.0	381.6	\$0.00
SMS	1.0	10.2	91.8	\$0.09
Balance inquiries	5.3	40.3	65.0	\$0.00
Balance recharges	0.5	3.6	7.1	\$0.83
Calling charges	\$1.93	\$4.34	\$9.33	

Calling charges exclude SMS, international calls, and service fees. Computed using the billing data, is available January 2005 - July 2008.

FIGURE 2. Telephone Subscriptions in Rwanda



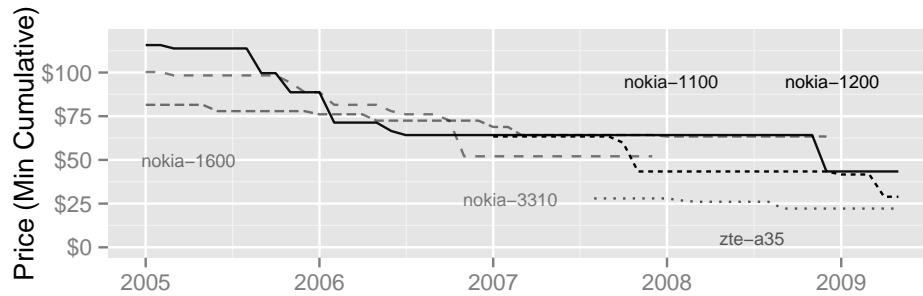
88% of the market during this period. This data includes nearly every call, SMS, and top up made over 4.5 years by the operator’s mobile phone subscribers, numbering approximately 300,000 in January 2005 and growing to 1.5 million in May 2009. For each transaction, the data reports: anonymous identifiers for sender and receiver, corresponding to the phone number and handset, time stamps, call duration, the incurred charge (for transactions before August 2008), and the location of the cell towers used.¹⁵

Coverage: I create coverage maps by computing the areas within line of sight of the towers operational in each month. I use a method suggested by the operator’s network engineer. Elevation maps are derived from satellite imagery recorded by NASA’s Shuttle Radar Topography Mission and processed by the Consortium for Spatial Information (Jarvis et al., 2008; Farr et al., 2007). I also compute two instruments for coverage, incidental coverage from the placement of the electric grid and the slope of the surface.

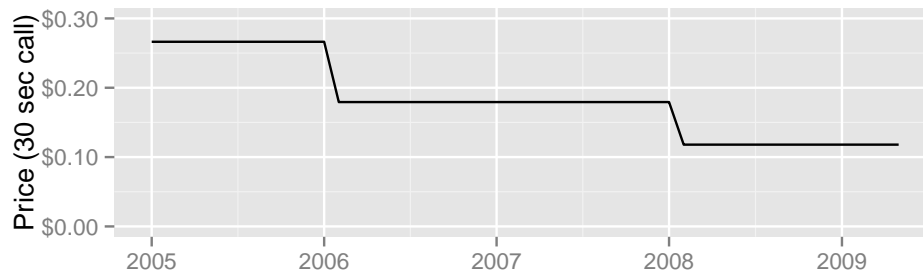
Individual locations and coverage: I infer each subscriber’s set of most used geographical locations using an algorithm analogous to triangulation, a version of Isaacman et al. (2011)’s ‘important places’ algorithm that I have modified to improve performance in rural areas. Around each individual’s most used locations, I compute

¹⁵Some months of data are missing; from the call records: May 2005, February 2009, and part of March 2009, and from the billing records: October 2006 and the months following August 2008. The records of some tower identifiers are missing from this data. I infer the location of missing towers based on call handoffs with known towers using a procedure I have developed, described in the Supplemental Appendix.

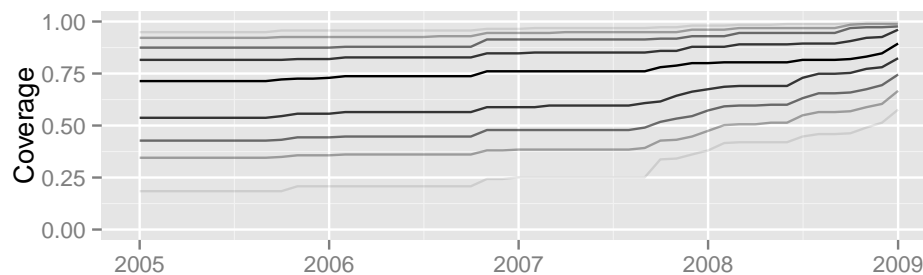
FIGURE 3. Variation in Data
Handset Prices (top 5 retail models)



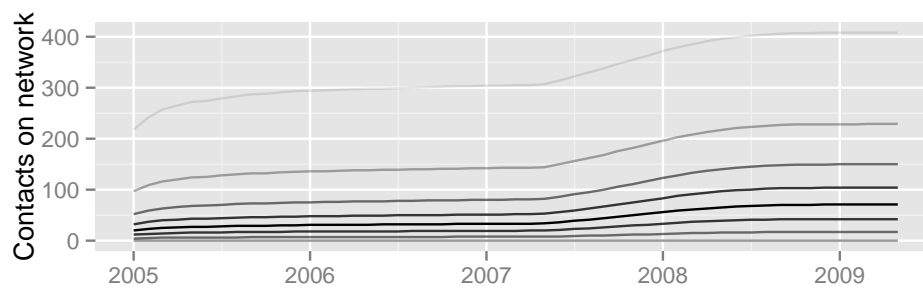
Calling Price (30 second call at peak rate)



Coverage (quantiles for eventual subscribers)



Contacts on the network (quantiles for eventual subscribers)



Quantile graphs graph the 10th through 90th percentile of the given quantity over time for all individuals who eventually subscribe, irrespective of whether that individual had subscribed by that time. Contacts graph omits 90% quantile.

the fraction of area receiving coverage in a given month using a two-dimensional Gaussian kernel with radius 2.25 km. I then compute the coverage available to each individual during each month by averaging this fraction over the individual's locations, weighting each location by the number of days calls were placed from that location.

Handset prices: I create a monthly handset price index $p_t^{handset}$ based on 160 popular models in Rwanda, weighting each model by the quantity activated on the network. I account for the introduction of new handsets by filling in missing prices with prices from handsets of comparable quality.

Household surveys: I use several nationally representative household surveys to provide background information: DHS and government surveys (EICV) from 2005 and 2010, and a technology usage survey (Stork and Stork, 2008).

4. PATTERNS OF MOBILE PHONE USE

Subscribers use the network creatively to relay information at low cost. Calls are extremely short: 58% of accounts have never placed a call longer than five minutes, and the mean length is 37.5 seconds. According to a representative household survey, 92% of subscribers report that the main purpose of the last 10 calls was social (Stork and Stork, 2008).

The primary unit of observation is an account, which corresponds to a phone number. Although accounts are prepaid and not explicitly linked to individuals, few individuals had more than one account in Rwanda during this period.¹⁶ I assume that each account is associated with a unitary entity such as an individual, firm, or household. This assumption would be violated if the composition of people sharing a handset changed over time.¹⁷ For ease of exposition in the rest of the paper I will refer to accounts as individuals or nodes.

¹⁶There are an average of 1.03 accounts per user (Gillwald and Stork, 2008). There was little reason to change accounts: there was one majority operator, opening an account cost roughly \$1, and the asymmetry in billing increased the hassle of changing your phone number. Prepaid accounts are not explicitly closed; if unused, they become inactive and are reactivated when credit is next added.

¹⁷I expect most *changes* in sharing to be across households, because much of expansion represents new households adopting phones rather than households purchasing additional phones. Among Rwandan households with mobile phones in 2010, 60% had one phone and 27% had two (EICV). For more discussion of this assumption, see Appendix A.

Calls reveal a social network. A call from one individual to another reveals a desire to communicate. Taken together, observed calls trace out the links of a latent social network for remote communication, which I refer to as the communication graph. I model the utility of communicating with a fixed potential set of contacts, which may represent family, friends, or business contacts. I assume that by the end of the data I have observed the full communication subgraph for the individuals who subscribe by May 2009: that the contacts I observe an individual call represent all of the contacts they would like to call among those who subscribe by this date.¹⁸

The prepaid billing structure is empirically convenient in that the calling party always pays on the margin for a call, so that the calling decision reflects willingness to pay for communication with a given contact. Due to the asymmetry in billing, the direction of the call is important: in the absence of a side contract, a call from i to j reveals that i is willing to pay at least the cost of the call, but does not reveal how much j would be willing to pay. Because of the potential importance of direction, I take the communication graph to be a fixed, directed network. I will present results under different assumptions of the value of incoming calls.

Dependence between links. A typical demand model would suggest links are substitutable: when my friend Jacques buys a phone, I may call him more and my brother less. An information sharing model would suggest complementarities: Jacques and my brother may share additional information, and as a result I may call both more.

One simple test of dependence is whether the volume of calls across a link changes as more of the sender's and receiver's contacts join the network. To test this, I estimate a simple gravity model, regressing each link's monthly call volume on the sender's and receiver's number of subscribing contacts, controlling for price changes and coverage, and including fixed effects for each link. If links were substitutable, as new contacts join the network a subscriber would reduce calls to existing contacts; barring any confounds this would result in a negative coefficient on number of contacts. Complementarity would result in a positive coefficient. As shown in the first two rows of Table 3, results are consistent with dependence between links being small, and on

¹⁸For this project, the call graph is exactly the object of interest; it may differ from a social network revealed through survey methods. For more discussion see Appendix A.

TABLE 3. Determinants of Calling

Dependent Variable: Duration (seconds per month, outgoing)	I	II	III	IV	V
Price USD/minute (event study)	-48.10 (0.63)	-46.17 (0.63)	-46.86 (0.63)	-46.00 (0.63)	-46.19 (0.63)
Sender's Coverage	15.99 (0.37)	12.42 (0.38)	14.75 (0.37)	12.39 (0.38)	0.20 (0.69)
Receiver's Coverage	24.67 (0.38)	22.54 (0.38)	22.21 (0.39)	22.07 (0.39)	10.17 (0.69)
Sender's \times Receiver's Coverage					19.63 (0.94)
# Sender's Subscribed Contacts		0.025 (0.001)		0.022 (0.001)	0.022 (0.001)
# Receiver's Subscribed Contacts			0.016 (0.001)	0.004 (0.001)	0.004 (0.001)
$N_{observations}$	44,108,852	44,108,852	44,108,852	44,108,852	44,108,852
N_{links}	1,663,018	1,663,018	1,663,018	1,663,018	1,663,018
R^2	0.0011	0.0011	0.0011	0.0011	0.0012

All regressions include link, month, and price regime fixed effects. Estimates computed using incremental least squares, on a 1% sample of nodes and all their links. The price coefficient is estimated based on an event study around the two price changes, in February 2006 and February 2008, using a two month window before and after. Dummies are included for the other months within each price regime. The top 1% degree nodes have been omitted; their inclusion attenuates the contact coefficients. Standard errors reported in parentheses. R^2 's omit contributions of fixed effects.

net complementary. The change in call volume along a given link associated with 10 other contacts joining the network is the same as that associated with a calling price decrease of \$0.005 per minute. For comparison, the median number of contacts is 61, and the final peak calling rate is \$0.23 per minute.¹⁹ To simplify the model, I assume the utility obtained from a contact is independent of the state of other contacts on the network.

Adoption. In Rwanda as in many developing countries it would be difficult to enforce service contracts at scale, so nearly all accounts are prepaid. Joining the network entails opening an account, which was easy and cheap (about \$1), and investing in

¹⁹Several factors could lead to low substitutability in this setting. Subscribers spend little time on the phone (the median usage is 25 minutes per month), so phone use is unlikely to crowd out other activities.

a handset, which was expensive (offered at retail price; about \$70 in 2005). Most handsets were mainstream, imported models, with little differentiation in terms of features. Handsets could be purchased from the operator or independent sellers; local price trends are consistent with global trends. I treat the handset market as perfectly competitive and handset prices as exogenous.

The data covers a period of continual declines in handset and calling prices, and continual improvements in coverage and network size, as shown in Figure 3. Individuals appear to plan ahead when considering adoption: when asked in 2007, 89% of individuals without phones planned to purchase a phone in the future (Stork and Stork, 2008). I model adoption as a dynamic decision, where individuals incorporate expectations of future improvements into the adoption decision.

Additional Simplifications. I make a number of simplifications for tractability and due to data limitations. During this period, there were two operators licensed in Rwanda. My partner operator always had the vast majority of the market, with over 88% of subscriptions during the first 4 years of data; I ignore the other operator.²⁰

I focus on voice calls and do not explicitly model the utility from nonvoice transactions. To the extent these transactions are important, when I estimate a model of the utility from calls, it will represent a proxy for total communication utility. I will still be able to estimate this total utility by considering the adoption decision. Usage and availability of mobile internet during the period of interest was negligible, and mobile money was not available on the Rwandan network until 2010. Though important in other contexts, in Rwanda text messaging or SMS was high priced (\$0.10 per message) and represents less than 13% of revenue and 16% of transactions.²¹ From the data it is not possible to match the sender and receiver of a given SMS; for this reason I do not explicitly model SMS. I also omit any utility from missed calls. Only calls that are answered incur a charge; subscribers may exploit this feature of billing, communicating simple information by leaving missed calls ('beeps' or 'flashes', see Donner 2007). Because it is difficult to distinguish between missed calls that provide utility (communicating information) and those that provide disutility (due to network

²⁰See Supplemental Appendix for details on the evolution of the market.

²¹Revenue statistic based on the period of data where charges are reported, which covers January 2005-August 2008. The price of sending an SMS did not change from 2005-2009 for standard plans.

problems or inability to connect), I do not explicitly model missed calls.²² Since I have no information about foreign subscribers, I will not model the small fraction of international calls. I also omit calls from payphones.²³ An individual may learn about the benefits of using a phone from observing the usage of others; I do not model this. For more discussion about these simplifications, see Appendix A.

5. MODEL

In this section I describe a model of handset adoption. The utility of owning a phone is derived from making calls, so I begin with a model of usage. The model of usage will also account for changes that improved communication across links, specifically the expansion of coverage and reduction of calling prices.

Let G be the communication graph (a directed social network). The nodes of the graph, N , represent individuals who eventually adopt phones. At each period, each individual $i \in N$ may have a phone or not; let $S_t \subseteq N$ be the set of individuals with phones in month t . A directed link $ij \in G$ indicates that i has a potential desire to call j over the phone network; I assume this link exists if i has ever called j . I assume these links are fixed over time. As shorthand let $G_i = \{j | ij \in G\}$ be i 's set of contacts.

Calling Decision. At each period t where he has a phone, individual i can call any contact j that currently subscribes, $j \in G_i \cap S_t$, to receive utility u_{ijt} . Each month, i draws a communication shock ϵ_{ijt} representing a desire to call contact j ; this desire might be high after an important event or to coordinate a meeting, or low if there is little information to share. The shock is drawn from a link-specific distribution, $\epsilon_{ijt} \stackrel{iid}{\sim} F_{ij}$ that will be specified later.²⁴

²²An early iteration of the usage model included missed calls; however, it was difficult to estimate as changes in the price of calls induce little substitution between calls and missed calls.

²³Payphones place approximately 12% of call durations but receive only 0.8%. Because payphones receive so few outgoing calls from the rest of the network, omitting them would have little effect on the preferred usage model which uses outgoing calls. However, their presence would affect the adoption decision, in a manner similar to handset sharing—see Appendix A.

²⁴In the true data generating process, communication shocks are likely correlated across time and across links: e.g., while organizing a business deal I may call a set of contacts intensively for a few months. As long as this correlation structure does not have a time trend, it will simply be smoothed out in the adoption decision.

Given the shock, i chooses a total duration $d_{ijt} \geq 0$ for that month, earning utility:

$$u_{ijt} = \max_{d \geq 0} v_{ij}(d, \epsilon_{ijt}) - c_{ijt}d$$

where $v(d, \epsilon)$ represents the benefit of making calls of total duration of d given communication shock ϵ , and c_{ijt} represents the per-second cost.

I model the benefit of making calls as:

$$v_{ij}(d, \epsilon) = d - \frac{1}{\epsilon} \left[\frac{d^\gamma}{\gamma} + \alpha d \right]$$

where the first term represents a linear benefit and the second introduces decreasing marginal returns. $\gamma > 1$ controls how quickly marginal returns decline. α is a cost-dependent censoring parameter that controls the intercept of marginal utility, and thus affects the fraction of months for which no call is placed.²⁵

The marginal cost includes the per second price as well as a hassle cost of obtaining coverage:

$$c_{ijt} = \beta_{price} p_t + h(\phi_{it}, \phi_{jt})$$

where β_{price} represents call price sensitivity, p_t is the per-second calling price (including any tax), and $h(\phi_{it}, \phi_{jt})$ represents the hassle cost when the caller or receiver have imperfect coverage. An individual's coverage $\phi_{it} \in [0, 1]$ is derived from the fraction of the area surrounding his most used locations receiving cellular coverage in month t . The reduced form evidence suggests that the hassle depends primarily on the interaction of caller and receiver's coverage, so I parameterize the hassle cost as: $h(\phi_{it}, \phi_{jt}) = \beta_{coverage} \phi_{it} \phi_{jt}$.²⁶

²⁵There is little in the data to differentiate between the distribution of shocks and the precise shape of the utility function. My strategy is to impose restrictions from theory and intuition on the utility function, and then select a distribution that matches the data well. I specify 8 properties that a reasonable functional form of utility from telephone calls should satisfy (see Appendix B), which led to the selected form.

²⁶Rwanda is geographically small enough that, even at the beginning of the data, the signal from urban towers extends into even remote areas, but it is also hilly, so that the resulting coverage is quite spotty. When coverage is poor it is often possible to walk to a nearby hilltop to make a call; this hassle cost is reduced as coverage improves. In principle the caller and receiver's coverage could enter the hassle cost asymmetrically, but because it proved to be difficult to estimate a more complex specification, I include only the strongest term, which is the interaction (see Table 3).

Given this functional form, calling prices, and coverage of both sender and receiver affect both the frequency and duration of calls. The marginal benefit of an additional second of duration across a link is decreasing, so i will call j until the marginal benefit equals the marginal cost. This implies an optimal duration of:

$$d(\epsilon, p_t, \phi_{it}, \phi_{jt}) = [\epsilon(1 - \beta_{price}p_t - \beta_{coverage}\phi_{it}\phi_{jt}) - \alpha]^{\frac{1}{\gamma-1}}$$

which is larger when the desire to communicate that month (ϵ) is larger. If the desire to communicate is not strong enough, the individual would prefer not placing a call across that link: $d_{ijt} = 0$ when $\epsilon_{ijt} \leq \underline{\epsilon}_{ijt} := \frac{\alpha}{1 - \beta_{price}p_t - \beta_{coverage}\phi_{it}\phi_{jt}}$.

Then, the expected utility i receives from being able to call j in period t is:

$$Eu_{ij}(p_t, \phi_t) = \int_{\underline{\epsilon}_{ijt}}^{\infty} \left[d(\epsilon, p_t, \phi_t) \cdot \left(1 - \beta_{price}p_t - \beta_{coverage}\phi_{it}\phi_{jt} - \frac{\alpha}{\epsilon}\right) - \frac{1}{\epsilon} \frac{d(\epsilon, p_t, \phi_t)^\gamma}{\gamma} \right] dF_{ij}(\epsilon)$$

where ϕ_t represents the vector of coverage for all individuals.

Adoption Decision. Each month i is on the network, he receives expected utility from each contact who is also on the network:

$$Eu_{it}(p_t, \phi_t, \mathbf{x}_{G_i}) = \sum_{j \in G_i \text{ and } x_j \leq t} Eu_{ij}(p_t, \phi_t) + w \cdot Eu_{ji}(p_t, \phi_t)$$

where x_j represents j 's adoption time, u_{ij} represents the utility of calls from i to j (which i pays for), u_{ji} represents calls from j to i (which j pays for), and $w \in [0, 1]$ specifies how much recipients value incoming calls. Each month that i is not on the network he receives utility zero.

Conditional on the adoption decisions of others, an individual's adoption decision represents an optimal stopping problem. Individual i chooses when to adopt by weighing the discounted stream of benefits against the price of a handset, represented by index $p_t^{handset}$ (including any tax). Then, if i believes that his contacts will adopt at times $\hat{\mathbf{x}}_{G_i}$, he will consider the utility of adopting at time x_i to be:

$$U_i^{x_i}(\hat{\mathbf{x}}_{G_i}) = \delta^{x_i} \left[\sum_{s=0}^{\infty} \delta^s Eu_{ix_i+s}(p_{x_i+s}, \phi_{x_i+s}, \hat{\mathbf{x}}_{G_i}) - \beta_{price}p_{x_i}^{handset} + \eta_i \right]$$

where an individual's type η_i captures heterogeneity in the utility of adoption that is unobserved to the econometrician. While I explicitly model key determinants of adoption, there remains important heterogeneity between individuals that I do not

observe. Particular individuals may obtain more or less utility from using a handset than that suggested by the usage model, or individuals may receive direct utility from owning a handset from features like the clock or flashlight. At the point of adoption, individuals may forecast their per-period utility with heterogeneous error. And individuals may face different fixed costs of adoption: some may purchase a cheaper, used handset; others may have to learn new skills to operate a handset. I incorporate these three forms of heterogeneity in a manner that will prove tractable for estimation and simulation. Since I am unable to differentiate between the three sources, an individual's type refers to their sum: $\eta_i := \frac{1}{1-\delta} \left[\eta_i^{utility} + \eta_i^{forecast} \right] + \eta_i^{adoption}$. I do not restrict the distribution of η_i (specifically, it need not be mean zero), but do require that each individual's type is constant over time to make simulation tractable. This assumption would be violated if, for example, a person who was pessimistic about improvements in the network later became optimistic, a handset provided status value that changed over time, or a person preferred to purchase a handset in a certain month because he was flush with cash or had more calling needs. I expect any changes in the value of status to be dwarfed by the large changes in fundamentals over this period (prices, coverage, and individuals on the network). Since individuals are likely to face idiosyncratic shocks, I evaluate the performance of the model in their presence using Monte Carlo simulations in the Supplemental Appendix, and find that it performs well for up to moderate shocks.

Network Adoption Equilibrium. The model of individual decisions presented thus far, and the estimation procedure introduced in the next section, are compatible with many definitions of equilibrium. However, counterfactual simulations will require a specific, tractable definition of equilibrium.

Initial adopters (S_0) are held fixed. Each other individual i decides on an adoption time $x_i \in [1, \dots, \bar{T}]$ to maximize his payoff $U_i^{x_i}(\mathbf{x}_{G_i})$, which depends on his contacts' adoption decisions (\mathbf{x}_{G_i}). The number of potential states of the network is large ($2^{|S \setminus S_0|} > 2^{1,000,000}$); I maintain tractability with a simplified concept of equilibrium:

First, I simplify individuals' expectations about the future. I avoid populating and managing a vast tree of potential states of the world by assuming that in equilibrium,

individuals compute payoffs based on a correct anticipation of the dates their contacts adopt ($\hat{\mathbf{x}}_{G_i} = \mathbf{x}_{G_i}$), with any forecast error fully captured in the constant term η_i .²⁷

Second, I simplify the strategies individuals can employ: individuals choose only one action, their adoption time x_i ; they may not condition their strategy on the actions of others in prior periods. This will result in a form of naiveté: individuals do not anticipate how the rest of the network will respond to their actions.²⁸

An equilibrium corresponds with a Nash equilibrium of the game where each individual simultaneously announces their adoption date x_i at the beginning of time (a complete information static game). An equilibrium Γ is defined by adoption dates $\mathbf{x} = [x_i]_{i \in S}$ such that the adoption date of each individual $i \in S \setminus S_0$ is optimal given their contacts' adoption dates: $x_i = \arg \max_t U_i^t(\mathbf{x}_{G_i})$.

Despite the simplicity of this definition, it will allow for rich behavior. In simulations, a perturbation of utility that causes one individual to change their adoption date can shift the equilibrium, inducing ripple effects through potentially the entire network. There are also likely to be multiple equilibria. In a stylized theoretical model of network adoption, the structure of equilibria can range from an equilibrium where no individuals ever adopt, to one where all individuals adopt immediately. In this setting, equilibria will not include this full range. First, the lowest equilibrium tends to include at least some adoption, because there is a stock of initial adopters and the net benefit of connecting to them increases over time, as prices decline and coverage improves. Second, coordinating on immediate adoption is unlikely to be optimal, because the economic costs of producing handsets decline over time. If over a period the decline in handset cost exceeds the social benefits a potential adoptee

²⁷These adoption dates can be derived from the perceived adoption utility $U_i^{x_i}(\cdot)$ for each individual in the network; if there are multiple equilibria, individuals would also need to know the method by which one is selected. (The perceived adoption utilities can be further decomposed into model primitives that include link utilities $\{u_{ijt}\}$ and individual types $\{\eta_i\}$. However, while it is conceivable that a person might know how another perceives the benefits of adopting ($U_i^{x_i}(\cdot)$), it would be unreasonable for individuals to decompose this object into primitives, as type η_i includes forecast error.)

²⁸In the Supplemental Appendix I test how an individual's adoption decision would differ if it took into account others' responses; I find differences in this setting are minor.

would generate, it is socially optimal for them to delay adoption.²⁹ The social benefits generated by an individual’s adoption are themselves nontrivial, because they can depend on the decisions of all other nodes.

The next section describes how the parameters of the model are estimated, and Section 7 describes a simulation procedure that identifies equilibria.

6. ESTIMATION

Individuals choose when to adopt a mobile phone and, if they adopt, how to use the phone. I work backwards, starting with the model of usage. The usage decision identifies the expected duration of voice calls across each link of the communication graph, $Ed_{ij}(p_t, \phi_t)$, the associated expected utility, $Eu_{ij}(p_t, \phi_t)$, and the price coefficient β_{price} that maps utility to valuation in dollars. In the adoption decision, individuals weigh the price of a handset against the discounted stream of call utility it provides. This decision overidentifies the price coefficient; I use plausibly exogenous variation affecting the utility of adoption to check the first estimate.

Calling Decision. I use data on phone calls to estimate the country’s latent communication graph (the parameters of the call shock distributions F_{ij}), the shape of the utility function (γ and α), and how usage responds to prices and coverage (β_{price} and $\beta_{coverage}$). I estimate responsiveness to prices and coverage using time series variation in both quantities.

Identification. It is generally difficult to disentangle peer effects across social network links from nonpeer effects that are correlated among network neighbors. I am able to overcome this problem because I observe the actual behavior of interest across links: the communication I observe represents the source of utility derived from the network.

I determine the value of this communication by estimating its elasticity to price. The operator steadily reduces calling prices and improves coverage over time, while

²⁹Some of these changes, the increase in coverage and reduction in calling prices, arise as the operator encourages and adapts to the growth of the network. In a richer definition of equilibrium that included operator decisions, these changes may be attenuated in a low equilibrium. However, even in the absence of these changes, consumers would still face dramatically declining global handset prices, so that the lowest equilibrium would tend to include some adoption.

the network expands.³⁰ The changing composition of the network would bias a simple time series estimate of the response of duration to prices and coverage: less talkative individuals subscribe later, causing average durations to decrease as communication costs are reduced. Supply is also endogenous: the operator adjusts prices and expands coverage in response to the changing composition of active links. I address both of these concerns by measuring the response of durations to costs using within-link variation: how usage between pairs of nodes change as coverage improves and prices decline. I do this in a manner analogous to using fixed effects, by estimating link specific call shock distribution parameters.³¹

Estimation Procedure. First, I specify a distribution for call shocks ϵ_{ijt} . To account for the large fraction of months on a given link without a call, I use a mixture of a lognormal distribution, $\ln N(\mu_{ij}, \sigma_i^2)$, and a mass point at negative infinity with probability $1 - q_i$. Thus, across each link some censoring is explained by cost (controlled by α in the utility function), and some would occur regardless of cost (controlled by the individual parameter q_i).³²

The calling decision has 7 types of parameters. I allow the means of the shock distribution to vary at the link level (μ_{ij}), I allow the standard deviation of the shock distribution and cost-independent censoring parameter to vary at the individual level

³⁰Prior to February 2006, calls were billed by the first minute and each subsequent half minute; after, subscribers could opt in to per second billing (and most quickly did). Modeling the per-minute charges would add significant complexity, so instead I assume these calls were billed at an equivalent per second price, selected to approximate both marginal and average prices. I set the per second price to the equivalent charge under the per minute rate when calls are of length 30 seconds.

³¹In contrast, Ryan and Tucker (2012) group nodes into 64 different types based on characteristics, and groups links by which types they connect. This more aggregated approach works in a context with static firm policies; however, when firm policies change over time, it would confound the response to the policy with unobserved heterogeneity within types. For example, if within one type half adopts before an increase in quality and half after, it is unclear how much of usage is explained by the level of quality and how much by fundamental differences between early and late adopters. To disentangle this heterogeneity one could include fixed effects. Fixed effects based on date of adoption would have a manageable number (52 in my context) but this would entail partitioning the parameter space based on an endogenous agent decision. More disaggregated fixed effects avoid this. The subsample approach I develop tractably estimates a nonlinear model with 415 million fixed effects (and would remain tractable with orders of magnitude more). This allows me to separately identify network parameters and consumer sensitivity to firm policies (price and quality: coverage).

³²On average, there is a call across a given link only 12% of months. If I used only a familiar continuous distribution, most of the mass of the distribution would be to the left of the censoring point, and distribution parameters would be estimated primarily off of this censoring point.

(q_i, σ_i) , and I assume that the shape and sensitivity parameters are common to all links $(\gamma, \alpha, \beta_{price}, \beta_{coverage})$. I assume that both individuals i and j know the parameters of link ij at the beginning of time. I estimate these parameters using maximum likelihood.

In each period t , for each pair of contacts i and j , I observe a duration $d_{ijt} \geq 0$. The model maps each duration d to an underlying call shock ϵ , conditional on prices and coverage:

$$\epsilon(d | p_t, \phi_{it}, \phi_{jt}) = \frac{d^{\gamma-1} + \alpha}{1 - \beta_{price}p_t - \beta_{coverage}\phi_{it}\phi_{jt}}$$

There will be a month without a call ($d_{ijt} = 0$) if the call shock was not high enough to place a call. i will choose duration zero for the set of epsilons mapping just below duration 1 second, so that a month without a call has likelihood $F_{ij}[\epsilon(1 \text{ second} | p_t, \phi_{it}, \phi_{jt})]$. If the shock is large enough, i will place a call; the higher ϵ , the longer the duration. The likelihood of calls of total duration d_{ijt} from i to j in month t is $F_{ij}[\epsilon(d_{ijt} + 1 | p_t, \phi_{it}, \phi_{jt})] - F_{ij}[\epsilon(d_{ijt} | p_t, \phi_{it}, \phi_{jt})]$.

The full sample has 1,525,061 nodes and 414.5 million links (representing a total of 15 billion link-month observations). I estimate the 4 common parameters and the distribution parameters defining the communication graph (two for each node and one for each link, 418 million altogether) using three steps. First, I jointly estimate common and distribution parameters for a random 2% subset of nodes and their full set of links.³³ Because the number of parameters is large, I find that the maximum likelihood estimates of the common parameters tend to be biased ($\hat{\alpha}$ tends to be biased downwards).³⁴ In the second step, I measure and correct this bias with a parametric bootstrap, by estimating from data simulated from the model. Third, I impose the bias-corrected common parameters estimated in the previous two steps to

³³The 2% sample is 24,849 nodes with 4,056,654 links. This represents approximately 122 million link-month observations; 8 million with calls.

³⁴The number of parameters grows with the number of links, so for asymptotics I take the number of observations to grow in the time dimension. For individuals who adopt late in the data, I have few observations of usage; and links with few calls have mostly censored observations, which are less informative (the median number of observations per node is 1,791 and per link is 46; see Table 4 for the quantiles of the observations per link, node, and observation). These links lead to an incidental parameter problem, affecting the estimate of α in a fairly predictable way. See Monte Carlo simulations in the Supplemental Appendix.

estimate the remaining distribution parameters for the full sample.³⁵ The individual likelihoods are separable conditional on the common parameters, so this last step is computationally much less demanding than performing a full joint estimation.

I use the estimated model to compute the expected duration and utility along each link ij , $Ed_{ij}(p_t, \phi_t)$ and $Eu_{ij}(p_t, \phi_t)$, using the common time path of calling prices, and the paths of coverage specific to caller i and receiver j .³⁶

Adoption Decision. The adoption decision reveals individual types η_i , and provides a second check of the estimate of β_{price} . In choosing when to adopt, individuals weigh the cost of adoption against the discounted stream of benefits from being on the network. The main cost of adoption is the price of a handset, taken as exogenous.³⁷

I assume that individuals have perfect foresight, apart from additive forecast error which absorbed into η_i . I also assume that individuals make adoption decisions independently, without taking into account how others in the network will respond. Under perfect foresight, i knows that his contacts will adopt at \mathbf{x}_{G_i} , and he will adopt at the time $x_i = \arg \max_x U_i^x(\mathbf{x}_{G_i})$. If time were modeled as continuous, under regularity conditions the optimum would be obtained from the first order condition $\left. \frac{\partial U_i^x(\mathbf{x}_{G_i})}{\partial x} \right|_{x=x_i} = 0$. I compute a discrete time analogue using differences.

I observe each individual's month of adoption, x_i , and consider the utility he would have received had he adopted a different month, conditional on the actions of others. At time x_i , i faced the decision of buying a handset and obtaining utility $U_i^{x_i}$, or

³⁵Extremely long calls can lead to numerical issues because they may result in draws from the extreme tails of the normal distribution. Because they can hinder convergence, I omit the 1% of nodes that have talked to a contact longer than one hour in a given month in the first step. In the third step, convergence is less sensitive so I am able to estimate the parameters of all nodes. The ratio of links to nodes differs in the first and third steps because nodes that have placed long calls tend to have more links.

³⁶Integrals are evaluated using Monte Carlo draws. Note that both integrals are nonlinear functions of estimated parameters, so uncertainty in parameter estimates could bias the estimates of these expectations; however, I find these biases seem to be small in Monte Carlo simulations (reported in the Supplemental Appendix).

³⁷Nonmarginal fees associated with using the network did change over this time, but these were small relative to the price of a handset. Before June 2007, subscribers needed to add roughly \$4.53 in credit per month to keep their account open. The lifting of this policy led to a large increase in account openings. Actually opening an account entails purchasing a SIM card, which cost roughly \$1 itself plus the cost of an initial top up. The initial top up amount changed over time but the cost of the SIM remained relatively constant. Available top up amounts also changed during this period, which I do not model.

postponing adoption by K months for utility $U_i^{x_i+K}$. Since he adopted at x_i , revealed preference implies $U_i^{x_i} \geq U_i^{x_i+K}$. The utility of being on the network during the following K months must have exceeded the value of the drop in handset prices.³⁸

$$\sum_{k=0}^{K-1} \delta^k E u_{i x_i+k}(p_{x_i+k}, \phi_{x_i+k}, \mathbf{x}_{G_i}) + (1 - \delta^K) \eta_i \geq \beta_{price} (p_{x_i}^{handset} - \delta^K p_{x_i+K}^{handset})$$

Similarly, i could have purchased a handset K months earlier. At time $x_i - K$, i chose to postpone adoption to obtain expected utility $U_i^{x_i}$ instead of buying a handset and getting utility $U_i^{x_i-K}$. This implies $U_i^{x_i} \geq U_i^{x_i-K}$. Because i chose to postpone adoption by K months at $x_i - K$, the utility from the K months prior to purchase must have been worth less than the drop in handset prices:

$$\sum_{k=1}^K \delta^{K-k} E u_{i x_i-k}(p_{x_i-k}, \phi_{x_i-k}, \mathbf{x}_{G_i}) + (1 - \delta^K) \eta_i \leq \beta_{price} (p_{x_i-K}^{handset} - \delta^K p_{x_i}^{handset})$$

These conditions are necessary for equilibrium and are valid in the presence of multiple equilibria. If a group of individuals make joint adoption decisions, the bounds on η_i that I back out are likely higher than the true bounds.³⁹ The results also follow if individuals forecast future paths of utility with constant amounts of error; for example if each individual's level of pessimism or optimism about future utility is persistent.

If a unit of utility in the call model corresponds with a unit in the adoption model, then individual types η_i can be backed out from the adoption decision and no additional parameters need be estimated. However, the adoption decision overidentifies β_{price} , so as a check I estimate the parameter using an instrumental variables moment inequalities strategy. I exploit variation in the cost of providing coverage to different areas due to Rwanda's hilly geography, in a manner similar to Yanagizawa-Drott (2014), as well as in the number of contacts who join in response to a government adoption subsidy program. (For more details, see Appendix C.)

³⁸Under perfect foresight, i correctly forecasts the first K months of utility and his expectation of the continuation flow does not change between x_i and $x_i + K$. Both options provide the same continuation flow of utility after $x_i + K$, so they differ only in the utility provided in the first K months.

³⁹See Supplemental Appendix.

Given β_{price} , the adoption inequalities imply bounds for each individual's realized type: $\eta_i \leq \eta_i \leq \bar{\eta}_i$, where:

$$\eta_i = -\frac{1}{1-\delta^K} \left[\sum_{k=0}^{K-1} \delta^k E \left(\sum_{j \in G_i \cap S_{x_i+k}} u_{ijx_i+k} + w \cdot E u_{jix+k} \right) - \beta_{price} (p_x^h - \delta^K p_{x+K}^h) \right]$$

$$\bar{\eta}_i = -\frac{1}{1-\delta^K} \left[\sum_{k=1}^K \delta^{K-k} E \left(\sum_{j \in G_i \cap S_{x_i-k}} u_{ijx_i-k} + w \cdot E u_{jix-k} \right) - \beta_{price} (p_{x-K}^h - \delta^K p_x^h) \right]$$

for a deviation of K months.⁴⁰ During months extra fees were charged, I incorporate the fee schedule.⁴¹ I fix the discount factor $\delta = 0.9916 \sim (0.9)^{1/12}$.

Results. Parameter estimates are reported in Table 4. The model explains months that no calls are placed across a link partly with the cost of communication (since the cost-dependent censoring parameter α is above zero) and partly because there would have been no communication regardless of cost (since most estimated q_i 's are below 1).

Predicted durations are highly correlated with observed durations (correlation by month is 0.95, by node is 0.91, and by link is 0.67), and levels match closely: total predicted duration is 3% higher than observed duration.⁴²

⁴⁰While each individual chose one out of many potential adoption dates, I include only moments for one pair of comparisons (one value of K) to avoid a selection problem that would otherwise arise from the finiteness of the data. To balance precision with smoothing, I select $K=2$ months. Note that the error structure cannot rationalize all adoption decisions: there are some observed decisions (28%) for which the inequalities of η_i cross. This suggests that the true adoption model has time-varying heterogeneity that I do not observe. A seemingly straightforward way to rationalize these decisions would be to include a second, time-varying error term. However, if this were a random effect from a distribution, simulation would quickly become intractable. In many other empirical settings, errors can be sampled independently, but here a shock to individual i 's utility of adopting at time t can potentially affect the decisions of the entire network. Then, an equilibrium would be conditioned on the entire set of errors, an object of dimension $\bar{T} \cdot |S \setminus S_0| \sim 89m$. Even if this space were sampled intelligently, it would be difficult to sample a sufficient region given that computing one equilibrium takes approximately 12 hours on current hardware. However, Monte Carlo results reported in the Supplemental Appendix suggest that under reasonable parameter values, the model without time-varying shocks does capture policy impacts quite well even when the underlying model includes these shocks. When the inequalities for η_i cross, I assume η_i is the mean of the two bounds. I back out these bounds differently for subsidy recipients; see Supplemental Appendix.

⁴¹See Supplemental Appendix.

⁴²There are a small number of nodes with many links and high variation between links for which the model obtains poor fit. I omit the 10 nodes for which the estimates are most biased; for these 10 nodes the predicted duration is 6506 times too large.

TABLE 4. Parameter Estimates

Classification	Parameter	Estimate						
Common Parameters	γ	2.310						
	α	158.294 (bias corrected)						
	β_{price}	0.199						
	$\beta_{coverage}$	-0.900						
	Quantile:	1%	25%	50%	75%	99%	Number	
Link Parameters	μ_{ij}	0.43	3.59	4.89	5.47	7.59	414.5m	
	$SE(\mu_{ij})^a$	1.22	1.89	1.98	2.74	3.82		
	N per link	8	19	46	51	51		
Node Parameters	σ_i	0.00...	0.34	0.53	0.89	2.03	1.5m	
	$SE(\sigma_i)^a$	0.16	0.36	0.66	1.06	4.81		
	q_i	0.03	0.10	0.20	0.55	1.00	1.5m	
	$SE(q_i)^a$	0.00...	0.04	0.07	0.11	0.47		
	N per node	37	724	1,791	5,985	63,212		
<i>Backed out from adoption decision</i>	$\bar{\eta}_i/\beta_{price}$	\$	-13.86	-0.45	0.69	1.25	11.48	1.5m
	$\underline{\eta}_i/\beta_{price}$	\$	-17.49	-1.33	0.34	0.99	7.13	1.5m
Overall	N per parameter	8	24	41	45	50		
	$N_{observations}$	15 billion						

a: Standard errors reported in this table assume that there is no covariance between common parameters and communication graph parameters (for computational tractability). Usage decision parameters are estimated in a three step maximum likelihood procedure. The second panel reports the quantiles of estimates, quantiles of standard errors, and quantiles of observations per node and link. Each node has two parameters plus one parameter per link.

The model provides two separate ways of measuring the value of joining the network, the first based on the decision to call a contact and incur a marginal cost per second, and the second based on the decision incur the price of a handset at the time of adoption. The calling decision implies a price sensitivity ($\beta_{price} = 0.199$)

which corresponds with an average price elasticity of -2.16.⁴³ If recipients do not value incoming calls ($w = 0$), the adoption decision implies a very similar estimate: $\beta_{price} \in [0.170, 0.188]$, and under some treatment of outliers the bounds admit the estimate from the call model.⁴⁴ If recipients value incoming calls as much as outgoing calls ($w = 1$), the call utility appears to roughly double count the surplus from calls: I obtain the point estimate $\beta_{price} = 0.466$ which is 2.3 times the estimate of the call model.⁴⁵ I proceed using $w = 0$ as a base specification, and evaluate the case where $w = 1$ for robustness in the Supplemental Appendix. I use the more precise estimate implied by the call model, allowing error to be absorbed into an individual's type η_i . I also estimate a coverage sensitivity of ($\beta_{coverage} = -0.900$) which corresponds with an average elasticity of 1.45 for either sender or receiver's coverage.

The parameters of the communication graph are shown in the second panel of Table 4; I interpret these parameters using comparative statics in Table 5. I show expected outcomes for links with the median estimate of shock variance σ_i and cost-independent censoring parameter q_i , and a range of quantiles of shock means μ_{ij} . The top panel shows expected durations, costs, and utility when both parties have full coverage and prices are the lowest observed in the data. Since coverage is perfect, there is no hassle

⁴³The model implies that elasticities depend on cost and link parameters; average elasticity computed from a sample of 100,000 random links in a month where both sender and receiver subscribed, using Monte Carlo integration. This is comparable to developing country estimates of elasticities of penetration with respect to price, as proxied by average revenue per user: Waverman et al. (2005) finds a price elasticity of -1.50 for a sample of developing countries from 1996-2003; Kathuria and Uppal (2009) find -2.12 in India in 2008. As a check, the price sensitivity can be compared against the results of the OLS regression of the determinants of calling presented in Table 3. It is difficult to compute elasticities in an OLS specification because most link-months are zero, but I can compare estimates of the derivative of expected duration with respect to price. Table 3 presents OLS results from an event study around the two price changes, suggesting an increase of \$0.01 in the per minute price would result in a decline in expected duration of 0.46-0.48 seconds on each link. The structural model implies a smaller response: an average decline of 0.24 seconds for the months of the price changes. There are many reasons these estimates could differ. One is that the event study is unfortunately not as tightly identified as I would like: both price changes each occurred on the first of February of different years, so it is identified off of how calls in February and March change in those years relative to the three other years.

⁴⁴Outlier nodes earning a very high utility would skew the estimate, so my preferred estimate omits the nodes earning the top 1% utility from the network. The full sample estimate is $[0.224, 0.227]$. If I instead omit the 151 nodes with the top 0.01% utility, the bounds admit the estimate from the call decision: $[0.197, 0.208]$.

⁴⁵Omits the nodes earning the top 1% utility from the network. The full sample estimate is 0.534; if I instead omit the 151 nodes with the top 0.01% utility, the estimate is 0.507.

cost. For the median link (middle column), calls would be infrequent: the probability of making in a given month is 0.13, and durations are short: conditional on making a call, the expected duration for that month is 37.5 seconds. The expected monthly cost of communicating across the median link is \$0.02 (0.04% of the average monthly per capita consumption in a phone owning household in 2010), and the link provides an expected utility of \$0.04. Since the median individual has 61 links, the total durations and utilities for each individual will represent the sum from many links.

The middle panel of Table 5 shows the impact of reducing both parties' coverage to 70% (near the initial median coverage): optimal durations and probability of calling decrease, hassle costs increase, and utility is reduced. The bottom panel instead shows the impact of increasing price to the highest observed in the data, but maintaining full coverage: durations and probability of calling both reduce, but due to an increase in price rather than in hassle cost.

Estimates of individual types suggest that the median individual expects to receive \$0.34-0.69 of additional value from a handset each month, beyond the value represented by the call model (shown in Table 4). However, types are heterogeneous: the 25th percentile expects to receive \$0.45-1.33 less value, and the 75th percentile \$0.99-1.25 more. These values would include any additional utility benefit of owning a handset, or annualized fixed cost of adoption or forecast error about the value of joining the network.

7. SIMULATION OF NETWORK GOOD ADOPTION

This section outlines a simulation method to compute a new network equilibrium based on changes to the environment. Since I observe only individuals who were subscribers between January 2005 and May 2009, I consider the impact of counterfactuals on this subset.⁴⁶ Because counterfactuals may induce these individuals to delay adoption, I set the end date for the simulation three years beyond the limits

⁴⁶I hold fixed the adoption dates of initial adopters. Since I infer adoption from transactions, I assume subscribers with transactions between January and March of 2005 are initial adopters. For simulation results to capture the full impact of a policy, I can only compute counterfactuals that do not increase the utility provided to individuals that delayed adoption past May 2009. If a counterfactual does, the results will underestimate adoption.

TABLE 5. Call Model Comparative Statics

		Quantile of Shock Mean						
		1%	25%	50%	75%	99%		
Price	Coverage	Expected:						
	Lowest	100%	Duration conditional on call	4.2 sec	13.6 sec	37.5 sec	64.1 sec	7.9 min
			Probability of call	0.00...	0.004	0.13	0.18	0.20
			Call cost	\$ 0.00...	0.00...	0.02	0.05	0.37
			Hassle cost	\$ 0	0	0	0	0
			Net utility	\$ 0.00...	0.00...	0.02	0.06	0.67
Lowest	70%	Duration conditional on call	0	11.3 sec	27.1 sec	43.3 sec	5.8 min	
		Probability of call	0	0.001	0.07	0.15	0.20	
		Call cost	\$ 0	0.00...	0.01	0.03	0.27	
		Hassle cost	\$ 0	0.00...	0.01	0.03	0.29	
		Net utility	\$ 0	0.00...	0.004	0.02	0.34	
Highest	100%	Duration conditional on call	0	10.7 sec	24.9 sec	38.7 sec	5.3 min	
		Probability of call	0	0.00...	0.06	0.14	0.20	
		Call cost	\$ 0	0.00...	0.01	0.05	0.57	
		Hassle cost	\$ 0	0	0	0	0	
		Net utility	\$ 0	0.00...	0.003	0.01	0.27	

The shock variance σ_i and cost-independent censoring parameter q_i are set to their medians (0.53 and 0.20 respectively), and outcomes from the model are shown for the range of shock distribution means $\mu_{i\tau}$. Statistics computed on on 1% random subsample of nodes.

of the calling data, using aggregate adoption statistics to scale utility to account for expansion in the network after my data ends.⁴⁷

Because individual types η_i are set identified, I consider an equilibrium $\Gamma(\boldsymbol{\eta})$ as a function of the vector of individual types, $\boldsymbol{\eta} = [\eta_i]_i$. To identify an equilibrium, I use an iterated best response algorithm:

- (1) Propose a candidate adoption path \mathbf{x}^0
- (2) Allow each individual to optimize their decision, holding fixed the adoption path of others:

$$x_i^1 = \arg \max_t \delta^t U_i^t(\eta_i, \mathbf{x}_{G_i}^0)$$

- (3) Iterate, using the path from the previous step \mathbf{x}^k to form the next:

$$x_i^{k+1} = \arg \max_t \delta^t U_i^t(\eta_i, \mathbf{x}_{G_i}^k)$$

- (4) Stop when the equilibrium converges: $x_i^{k+1} = x_i^k$ for all i ⁴⁸

There tend to be multiple equilibria, for two reasons. Because each individual's type is backed out as a set rather than a point, different points in the set of types $\{\boldsymbol{\eta} | \underline{\eta}_i \leq \eta_i \leq \bar{\eta}_i\}$ may imply different equilibria. And, even given a vector of types $\boldsymbol{\eta}$, there may be multiple equilibria depending on how optimistic individuals are about others' adoption.

I derive bounds for the entire set of equilibria by exploiting its lattice structure. First, note that there is a monotonic relationship between η_i and i 's optimal adoption date x_i : a higher type η_i weakly decreases i 's optimal adoption date. Second, note

⁴⁷See Supplemental Appendix for more details. I did not extrapolate future utility in this way when estimating the adoption decision because estimates would be sensitive to the extrapolation assumptions; for simulation, these assumptions affect only individuals who end up changing their adoption month to lie outside of the period I have data. (For individuals receiving adoption subsidies the extrapolation also affects their lower bound estimates.)

⁴⁸With the aim of speeding convergence, in practice at each step k I use the path defined by x_j^k for individuals j that have reoptimized in this step and x_j^{k-1} for individuals who have not yet reoptimized in this step, in the same manner as the Gauss-Seidel method. The equilibrium identified is likely to be sensitive to the order that individuals reoptimize when simulating policies with nonmonotonic effects (shifting some individuals' adoption forwards and others backwards). In the paper I only consider policies with monotonic effects, which are less likely to be sensitive. I tested sensitivity by comparing a solved equilibrium to one solved with agents optimizing in reverse order and found small changes likely arising from rounding error (0.2% of nodes had different adoption months, averaging to an average difference in adoption month of -0.0003). The algorithm sometimes reaches a cycle rather than an equilibrium. These cycles tend to be quite small, involving only a handful of nodes. If the algorithm reaches a cycle, I break the cycle and note the number of nodes involved.

that the underlying game has strategic complements: a decrease in i 's adoption date x_i weakly decreases j 's optimal adoption date.⁴⁹ The lowest possible equilibrium, $\underline{\Gamma}$, can be identified by setting each individual's type η_i to its lower bound, and using a pessimistic candidate adoption path: $\mathbf{x}^0 = \bar{T}$ (initially individuals expect everyone else to completely delay adoption). The highest possible equilibrium, $\bar{\Gamma}$, can be identified by setting each individual's type to its upper bound and using an optimistic candidate adoption path: $\mathbf{x}^0 = 0$ (initially individuals expect everyone else to adopt immediately).⁵⁰ These filter through to provide bounds on the adoption date for each individual, $[\underline{x}_i, \bar{x}_i]$. I compute a third equilibrium, Γ^{mid} , by setting individual's types to the mean of the low and high bound estimates ($\eta_i = \frac{\underline{\eta}_i + \bar{\eta}_i}{2}$), and using the observed adoption path as the candidate path (initially individuals expect the observed equilibrium).

The state space is large: there are $\bar{T}^{|S \setminus S_0|}$ or on the order of $89^{1,000,000}$ possible outcomes. Altogether, I find the algorithm identifies an equilibrium in about 12 hours.⁵¹

Revenue and Utility. For each equilibrium I compute the net present value of revenue and utility, as of January 2005. Given a usage tax rate of $\tau_{usage,it}$, the firm revenue from equilibrium Γ is computed by summing the price times the expected duration across each link:

$$R_F^\Gamma = \sum_{i \in S} \sum_{t \geq x_i} \delta^t \frac{p_t}{1 + \tau_{usage,it}} \sum_{j \in G_i \cap S_t} Ed_{ij}(p_t, \phi_{it}, \phi_{jt})$$

Total utility from calls is computed analogously:

$$U_{calls}^\Gamma = \sum_{i \in S} \sum_{t \geq x_i} \delta^t \left[\sum_{j \in G_i \cap S_t} Eu_{ij}(p_t, \phi_{it}, \phi_{jt}) + w \cdot Eu_{ji}(p_t, \phi_{jt}, \phi_{it}) \right]$$

where this utility is net of calling and hassle costs incurred.

⁴⁹This follows from the lattice structure of \mathbf{x} and because $U^{x_i}(\eta_i, \mathbf{x}_{-i})$ has increasing differences in x_i and x_j , or is supermodular in \mathbf{x} ; see Topkis (1978) and Milgrom and Shannon (1994).

⁵⁰This is analogous to how Jia (2008) uses the lattice structure of an entry game to identify a range of equilibria.

⁵¹One factor that contributes to the algorithm's performance is that the benefits to individual i of joining the network at any point in time are bounded, by i 's minimal set of contacts $G_i \cap S_0$ and maximal set G_i .

In order to realize this utility, an individual had to purchase a handset. I assume handsets are provided by a competitive market at marginal cost. The handsets that subscribers purchase would last beyond the end of the data, so I calculate the cost of using the handset during the data by assuming each individual purchases a handset at their adoption month x_i and then sells it back at the end of the data at the prevailing price. This yields the following cost of handset ownership:

$$C_{handsets}^{\Gamma} = \sum_{i \in S} \left[\delta^{x_i} p_{x_i}^h - \delta^{\bar{T}^{data}} p_{\bar{T}^{data}}^h \right]$$

Then, the total net utility in money is given by:

$$U_{net}^{\Gamma} = \frac{1}{\beta_{price}} U_{calls}^{\Gamma} - C_{handsets}^{\Gamma}$$

where I convert the utility from calling into dollars using the price sensitivity. In welfare calculations I omit the type η_i that enters the individual's adoption decision, because this term may pick up a forecast error that does not represent the utility individuals receive. The government earns revenue from taxes on adoption ($\tau_{adoption,it}$) and usage ($\tau_{usage,it}$):

$$R_G^{\Gamma} = \sum_{i \in S} \left[\delta^{x_i} \frac{\tau_{adoption,it}}{1 + \tau_{adoption,it}} p_{x_i}^h + \sum_{t \geq x_i} \delta^t \frac{\tau_{usage,it}}{1 + \tau_{usage,it}} p_t \sum_{j \in G_i \cap S_t} Ed_{ij}(p_t, \phi_{it}, \phi_{jt}) \right]$$

Because there is a monotonic relationship between adoption date and utility, the lower and upper bound equilibria represent upper and lower bounds:

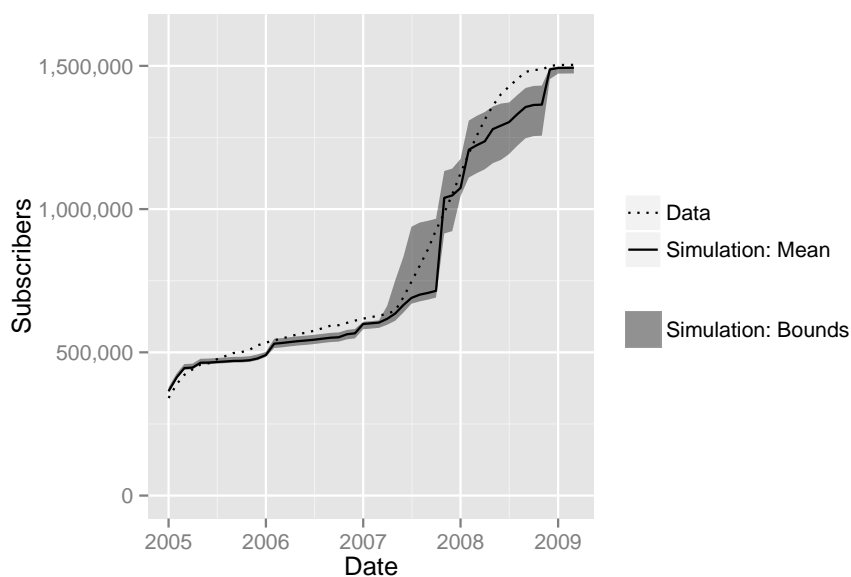
$$U_{calls}^{\underline{\Gamma}} \leq U_{calls} \leq U_{calls}^{\bar{\Gamma}}$$

If the usage tax is constant over time then firm revenue is also bounded by the upper and lower bound equilibria: $R_F^{\underline{\Gamma}} \leq R_F \leq R_F^{\bar{\Gamma}}$.

Because the net utility function omits idiosyncratic benefits, it does not match the utility each individual maximizes; there may be an equilibrium between $\underline{\Gamma}$ and $\bar{\Gamma}$ that has a net utility lying outside the bounds of $U_{net}^{\underline{\Gamma}}$ and $U_{net}^{\bar{\Gamma}}$. Similarly because handset prices are decreasing, government revenue may be a nonmonotonic function of adoption date, and there may be an equilibrium between $\underline{\Gamma}$ and $\bar{\Gamma}$ that generates government revenue outside the bounds of $R_G^{\underline{\Gamma}}$ and $R_G^{\bar{\Gamma}}$.

Baseline Simulation Results. I run the simulation on the same environment as the data to get a sense of the model's fit. As shown in Figure 4, the simulation

FIGURE 4. Simulation Fit



Fit metrics	Adoption month in data vs. adoption month under		
	Lower equilibrium	Mid	Upper equilibrium
Correlation	0.84	0.87	0.84
Mean deviation	5.17	2.62	-0.99
Mean absolute deviation	5.78	3.74	4.17
Median deviation	4	1	-2
Median absolute deviation	5	2	3

matches the general trend of the data. While adoption in the data grows continuously, the adoption path generated by the model has more discrete jumps, resulting from individuals coordinating on adoption dates, often at price changes.⁵² Under mean shocks the correlation between observed adoption month and simulated adoption month is 0.87, and the mean deviation is 2.62 months.

In the simulated equilibrium, I estimate the total benefit provided by the mobile phone system over the 4.5 years I observe to lie between \$347m and \$380m.⁵³ While several studies find welfare gains of mobile phones in developing economies in specific sectors (Jensen, 2007; Aker, 2010; Jack and Suri, 2014), to my knowledge these are

⁵²These jumps could be softened if individuals considered uncertainty about the future, if different subscribers faced different handset prices, or if subscribers had heterogeneous price sensitivities.

⁵³This is net of the hassle cost of obtaining coverage.

the first micro-identified estimates of the total welfare generated by a developing country mobile phone network. Of this total, between \$193m and \$211m accrues to the operator as revenue (an average of \$8-9 per subscriber per month).⁵⁴ The government collects \$78-84m as tax revenue on calls and handsets. (There is also a 30% tax on firm revenues after deductions; since I do not observe deductions I do not compute this explicitly, and as a result some additional portion of firm revenues should be transferred to the government.) Consumers obtain the remaining utility net of calling and hassle costs, \$75-85m (an average of \$3.20-3.62 per subscriber per month). Consumers face an additional cost, the cost of handset ownership, which I estimate to be between \$38-41m (an average of \$1.60-1.74 per subscriber per month), resulting in net utility U_{net} between \$37-44m (an average of \$1.60-1.88 per subscriber per month). Under these estimates the benefits of the phone system are split among the operator (62-63%), the government (25%), and consumers (13-14%).

Measuring Policy Impacts. For applications of this method in following sections, I am interested not simply in the levels of revenue and utility, but how revenue and utility change in response to policies. A natural measure of impact would be bounds on the changes in revenue and utility across the range of equilibria; however, this measure is computationally prohibitive because adoption decisions are interlinked. In many other empirical settings, errors can be sampled independently, but here i 's type η_i can potentially affect the decisions of the entire network. Even given the η_i 's restricted to lie within the realized bounds, the set of potential type vectors is large: $[\underline{\eta}_0, \bar{\eta}_0] \times [\underline{\eta}_1, \bar{\eta}_1] \times \cdots \times [\underline{\eta}_N, \bar{\eta}_N]$.

Instead, I measure policy impacts by reporting the change in the lowest and highest equilibria. These changes in the bounds need not coincide with bounds on the changes. In particular, the change in the upper bound equilibrium may be less than the change in the lower bound equilibrium, and there may be larger or smaller changes at type vectors within the bounds. When a policy change shifts the lower and upper bounds by similar amounts, I report one number to describe the approximate shift. When

⁵⁴This is comparable to statistics from the operator's annual reports: average revenue per user per month declined from \$19 in 2005 to \$7 in 2009, as calling prices were reduced and less talkative subscribers joined.

the lower and upper bounds shift by different amounts, I report both, and either note which bound has shifted or describe one as a high case and one as a low case.⁵⁵

8. APPLICATION: TARGETING ADOPTION SUBSIDIES

Adopting a network good benefits not only one’s contacts: by influencing their adoption, it also benefits others further away in the network. As a result, the adoption of network goods is likely to be inefficient: there may be nodes that would provide net social benefit who do not internalize enough private benefit to adopt.

One can imagine two scales for overcoming these inefficiencies:

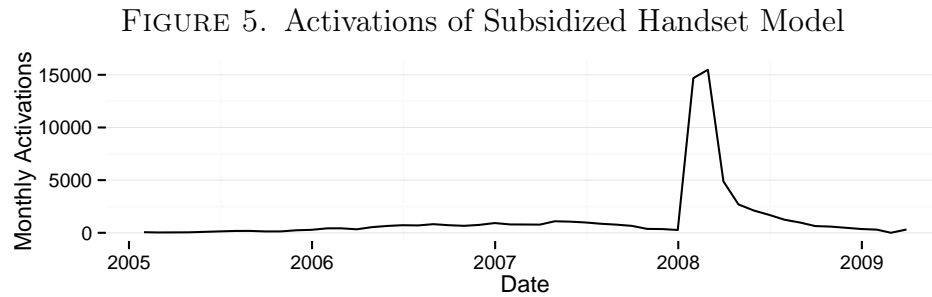
An individual node is aware of his local network structure, and may find it privately optimal to subsidize a neighboring node that otherwise would not adopt (say, buying a phone for a grandparent). However, if an inefficiency is dispersed beyond a handful of nodes it would be difficult for a region of the graph to coordinate to overcome it.

Firms and governments have objective functions that cover the graph more expansively, and may find it optimal to implement large scale subsidization or price discrimination programs that may improve efficiency (Katz and Shapiro, 1994). These programs are common: for example, Facebook currently subsidizes data usage in developing countries. However, global actors are constrained by information. While selecting an optimal policy may require perfect knowledge of the flows of benefits, they have only a rough image of the network structure and thus generally rely on intuition or simple theories to navigate what is a complex web of interconnected benefits.

In this section I demonstrate a method allowing global actors to use empirically measured network structure to evaluate—and potentially improve—targeting of adoption subsidies. I first evaluate a historical example, an adoption subsidy program implemented by the Rwandan government in 2008. I then describe how the method can be used to encourage the adoption of future network goods.

2008 Adoption Subsidy Program. To promote access to telecommunications, the Rwandan government in 2008 purchased 53,352 handsets (amounting to roughly 8% of the country’s stock of handsets at the time) and distributed them to individuals

⁵⁵Monte Carlo simulations in the Supplemental Appendix provide one test of this approach.

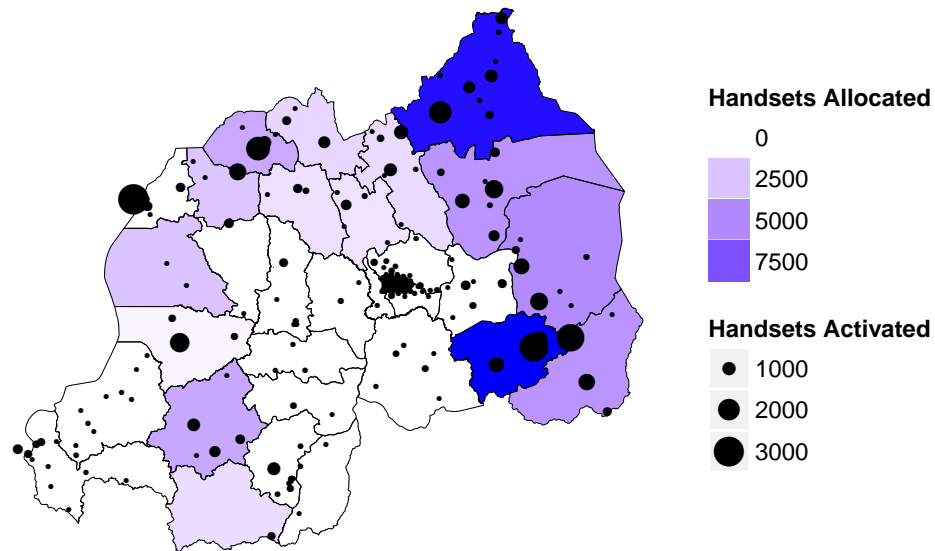


through local governments at a reduced price. Fifteen of 30 districts participated in the program. Handsets were generally allocated to rural districts with low baseline mobile phone adoption (in participating districts, 4% of households had mobile phones, versus 12% in nonparticipating districts). Allocations varied significantly: half of the districts were allocated no handsets; those allocated handsets received enough for between 1% and 15% of households. Each district handled its own distribution; generally, individuals came to the district office to voice interest.

The handsets were all the same model, the Motorola C113, which was chosen because it was low cost and had a long battery life. This particular model was otherwise rare in the country at the time, so I am able to identify beneficiaries based on receiving this model of handset during the dates of distribution. Figure 5 shows activations of this model over time. I consider an account as subsidized if it was activated during the first four months of 2008 and its mode handset was the subsidized model. There are 41,225 such accounts. That I observe fewer subsidized accounts than handsets allocated per government records could arise from subsidized handsets being activated later than April 2008, passed between accounts, used on the competing operator, or not being used within the country.⁵⁶ Beneficiaries were to pay a fraction of the full price of the handset (\$28) through monthly repayments of \$1.81, but few of these payments were made. I assume that each recipient made an average of 5 payments, so that the program represented a discount of \$18.94.

⁵⁶I consider handsets as subsidized only if activated during this period because during later months it is difficult to tell if activations are part of the subsidy program.

FIGURE 6. Handset Subsidy: Allocations and Activations



Allocations source: Banque Rwandaise de Développement. Motorola C113 handsets activated after January 2008 are considered distributed by the subsidy program.

Use of Subsidized Handsets. I observe the ultimate recipient of the handset.⁵⁷ Handsets appear to be used either where allocated or in urban areas. Figure 6 shows where handsets were allocated based on government records, and where these handsets were subsequently activated according to phone network records. There is a clear association between allocation district and location of activation, but also many handsets were activated in urban areas (the major clusters of activations in regions with no handsets allocated represent urban areas).

Recipients use handsets in a similar manner as nonrecipients who subscribed around the same time. One potential concern with a subsidy program is that goods may be allocated to consumers who do not value them. While I cannot conclude much about the initial recipient, the ultimate recipients of subsidized handsets use their phones less than individuals who subscribed earlier, but on par with individuals who

⁵⁷If there was an exchange I do not observe the associated transfer. In any exchange, I assume that the subsidy amount is passed through to the ultimate recipient.

purchased phones around the same time, in terms of calls, durations, and total number of contacts.⁵⁸

An optimal subsidy program to overcome dispersed network externalities would target individuals with particular network structure, but recipients' network structure is similar to others who subscribed around the same time. One would want to target those who provide benefits to others who have yet to subscribe, who would not subscribe in absence of the target's adoption. One metric of these benefits is the eventual duration spoken with contacts that have yet to subscribe, which is very similar for recipients and nonrecipients (35% for subsidy recipients, 33% for all subscribing in the same months).⁵⁹ I also compute the clustering coefficient (the fraction of a node's neighbors who are themselves connected, which is 0.082 for subsidy recipients and 0.081 for all subscribing in the same months).

The results are suggestive of a program that increased the supply of handsets, with handsets ultimately being used by relatively typical users. However, the ultimate impact on network adoption depends on the interaction of the recipients' adoption decision with the network of benefit flows, which I analyze using the simulation method developed in this paper.

Simulated Impact of Adoption Subsidy. I simulate how equilibrium adoption would change if the subsidy were not provided, using three assumptions:

- *All eligible individuals took up the subsidy.* Given the decentralized nature of the implemented subsidy program, it is difficult to determine the entire set of individuals who were eligible. Since the subsidy was very attractive, I assume that all eligible individuals took up the subsidy and that it was valid only in the month they adopted.
- *Recipients did not delay adoption in order to receive a subsidy.*
- *Recipients preferred taking the subsidy at the point of adoption to purchasing any time in the following 4 years.*⁶⁰

⁵⁸See the Supplemental Appendix for a more detailed description of the subsidy targeting.

⁵⁹Subsidy recipients represent 13% of those subscribing in these months.

⁶⁰For more details on these last two assumptions, see the simulation details in the Supplemental Appendix.

These restrictions on the set of recipients and their preferences allow me to compute the effect of the subsidy. Results are shown in Table 6. I compute the baseline simulation (“with subsidy” in the table), as well as two simulations where the subsidy has been removed. The first captures only the immediate effect of removing the subsidy: I allow each recipient to reoptimize their decision individually, without allowing those changes to ripple through the network (“no subsidy, only proximal effect of removal”).⁶¹ The second is the equilibrium that results after all nodes have adjusted their decisions (“no subsidy, proximal and ripple effects”). The first column shows the results for all nodes; subsequent columns show results for subsidy recipients and nonrecipients (most of whom are connected to at least one subsidy recipient).

Because the decision to purchase a subsidized good only loosely reveals how much the recipient values it, the bounds I obtain are wide. The upper bound presents an optimistic scenario: targeted individuals would have delayed adoption by an average of only 1.8 months in the absence of the subsidy. The lower bound presents a more pessimistic scenario: targeted individuals would have delayed adoption by an average of approximately 2 years. These bounds could be made tighter by either gathering more information or making more assumptions about the price sensitivity of subsidy recipients.

As described in Section 7, I measure the impact of the subsidy by reporting changes in the bounds on revenue and consumer surplus around the set of equilibria, rather than bounds of the changes. In this application the lower bound equilibrium shifts more than the upper bound equilibrium because the targeted individuals change their decision more in the pessimistic scenario.

I find:

The subsidy improved welfare. Factoring in the net present cost of the subsidy of \$569,741, it shifted the bounds on net welfare upward \$4,869,961 (lower equilibrium) and \$301,980 (upper equilibrium).⁶²

⁶¹It would be more natural to simulate the direct impact of providing rather than removing the subsidy, but this is difficult for technical reasons due to the way η_i is backed out for subsidized nodes. See Supplemental Appendix for a discussion.

⁶²These results consider the portion of the subsidy allocated only to the 41,225 individuals I can clearly identify as recipients. The subsidy for the other 12,127 handsets would have represented an additional net present cost of \$159,130. In the most extreme case where this value was destroyed through misallocation, this cost should be subtracted from the welfare gains.

TABLE 6. Impact of Adoption Subsidy Program

	All nodes	Recipients	Nonrecipients
Number	1,503,670	41,225	1,462,445
Adoption Time (mean)			
... with subsidy	[25.38, 22.16]	[37.38, 37.38]	[25.04, 21.73]
... no subsidy, only proximal effect of removal	[25.96, 22.21]	[58.61, 39.09]	[25.04, 21.73]
... no subsidy, proximal and ripple effects	[26.12, 22.22]	[60.97, 39.14]	[25.14, 21.74]
Total Impact of Subsidy	-0.74, -0.06	-23.59, -1.76	-0.10, -0.01
... proximal effect of removal	-0.58, -0.05	-21.23, -1.72	0.00, 0.00
... additional ripple effect	-0.16, -0.01	-2.36, -0.04	-0.10, -0.01
Revenue (total)			
... with subsidy	[193.30, 210.76]	[1.03, 1.11]	[192.26, 209.65]
... no subsidy, only proximal effect of removal	[191.63, 210.66]	[0.27, 1.07]	[191.36, 209.59]
... no subsidy, proximal and ripple effects	[191.08, 210.61]	[0.21, 1.07]	[190.86, 209.54]
Total Impact of Subsidy	2.22, 0.15	0.82, 0.04	1.40, 0.11
... proximal effect of removal	1.67, 0.10	0.77, 0.04	0.90, 0.06
... additional ripple effect	0.55, 0.05	0.05, 0.00	0.50, 0.05
Consumer Surplus (total)			
... with subsidy	[37.37, 43.84]	[0.94, 1.04]	[36.43, 42.80]
... no subsidy, only proximal effect of removal	[35.87, 43.24]	[0.21, 0.48]	[35.65, 42.75]
... no subsidy, proximal and ripple effects	[35.37, 43.19]	[0.18, 0.48]	[35.19, 42.71]
Total Impact of Subsidy	2.00, 0.65	0.76, 0.55	1.23, 0.09
... proximal effect of removal	1.50, 0.60	0.73, 0.55	0.77, 0.05
... additional ripple effect	0.50, 0.05	0.03, 0.00	0.46, 0.04
Government Revenue (total)			
... with subsidy	[78.28, 84.41]	[0.38, 0.40]	[77.90, 84.01]
... no subsidy, only proximal effect of removal	[77.88, 84.93]	[0.24, 0.93]	[77.64, 83.99]
... no subsidy, proximal and ripple effects	[77.63, 84.91]	[0.20, 0.93]	[77.43, 83.97]
Total Impact of Subsidy	0.65, -0.50	0.18, -0.53	0.47, 0.04
... proximal effect of removal	0.40, -0.52	0.13, -0.53	0.26, 0.02
... additional ripple effect	0.26, 0.02	0.05, 0.00	0.21, 0.02

Results in each cell reported for the lower bound and upper bound estimate of the equilibrium. Impacts represent the difference in these bounds. I hold fixed the adoption decision of 6 subsidized nodes that have crossed bounds for η_i (for details see Supplemental Appendix). Utility and revenue reported in 2005 U.S. Dollars, discounted at a rate of 0.9 annually. Consumer surplus includes the surplus utility each individual receives from the call model through May 2009, minus the cost of holding a handset from the time of adoption until May 2009.

A substantial fraction of calling benefits accrued to nonrecipients. Recipients' utility increased by \$764,796 (lower) or \$554,982 (upper), from the combination of increased calling and the direct value of the discount. Nonrecipients only received utility from increased calling, but obtained 62% of all benefits in the lower equilibrium and 14% in the upper equilibrium.

It may have been profitable for the operator to finance the subsidy itself. If in absence of the subsidy, the targeted individuals would have substantially delayed adoption, it would have been profitable for the firm to subsidize their adoption itself. If the firm had financed the subsidy, the bounds on its profits would shift upward by \$1,652,102 in the lower equilibrium, but downward by \$418,513 in the upper equilibrium. Similarly, the tax revenue generated from the policy may have been enough to pay for it; on net the government earned \$650,868 in the lower equilibrium, and spent \$496,072 in the upper equilibrium.

Nonrecipients accounted for over 63% of the increase in revenue. In this setting, the operator has a near monopoly, and so would be able to internalize revenue generated by nearly all nodes on the network. Financing a subsidy would have been much less attractive for a competitor that was less able to internalize revenue from the rest of the network (for example, if interconnection fees were capped).

Most of the effect is a proximal effect of the subsidy. Ripple effects account for 25% (lower equilibrium) and 36% (upper equilibrium) of the effect on revenue and 26% (lower) and 10% (upper) of the effect on consumer surplus.

In the Supplemental Appendix, as a robustness check I analyze the program under the assumption that the surplus from calls also accrues to receivers ($w = 1$); results are qualitatively similar: the bounds on welfare increase by \$6,496,114 (lower) or \$67,075 (upper).

Overall, the impact of the subsidy on network adoption is consistent with what might have been expected from the descriptive evidence: it induces targeted individuals to subscribe earlier, and has a moderate impact on those further away in the network. That a substantial fraction of the effect of the subsidy spills over to contacts of the recipients suggests that subsidies for network goods should be thought of as benefitting neighborhoods of the graph rather than just individuals.

In this method, benefits are revealed only for regions that have adopted; but to guide adoption of a new good, a policymaker would most like to know the benefits to adoption in ‘dark’ regions of the network that have yet to adopt. An extension of this method can predict benefits in regions of the network that have yet to adopt a new good, using network structure revealed by a good that has already diffused: for example, mobile calling behavior can guide mobile internet policy in developing countries.⁶³ Since future mobile internet adopters are likely to be current mobile phone subscribers, their behavior is captured in operator data. Features from this data can predict unobserved characteristics; for example, Bjorkegren and Grissen (2015) finds that behavior revealed in mobile phone usage predicts loan repayment. A predictive mapping between network properties observed in call data and mobile internet usage can be estimated for early adopters, and then used to predict how mobile internet would be used in regions of the network that have yet to adopt. These predicted benefits can be analyzed directly, or can be used with the simulation method developed in this paper to evaluate policies to guide adoption.⁶⁴

9. APPLICATION: THE PROVISION OF SERVICE TO RURAL AREAS

Due to difficulties internalizing network effects, network good industries tend to be highly concentrated. This tendency towards concentration is strengthened when a good relies on high fixed costs or scarce resources, such as electromagnetic spectrum in the case of mobile phones. Because concentration would likely lead to inefficient provision in absence of regulation, network good industries are often regulated.

For communication services, a key question for regulators is whether—and if so, how—to ensure service to poor and remote communities. This remains an active question: over 1 billion people are not covered by a network providing mobile internet (3G/4G), and it is expected that even for basic voice service it will not be profitable for the private sector to serve 2-5% of the world’s population (GSMA, 2006). A

⁶³Mobile phones may be the most convenient modality to deliver internet service to poor and remote areas. In Rwanda in 2012, only 0.7% of households own a computer with an internet connection, while 19% of individuals aged 15 and over own a mobile phone capable of browsing the internet. (RIA, 2012)

⁶⁴See Supplemental Appendix for an outline of how this could be employed.

wide variety of policy instruments are currently in use to encourage rural service provision, including tax-and-transfer schemes, service obligations, and universal service funds that collect a fraction of operator revenues to spend on government-led projects (GSMA, 2013).

Whether and how to ensure service to remote areas depends crucially on both the shape of private benefits that would accrue to a network operator, and the benefits to consumers. Both are difficult to measure due to spillovers induced by geographical interconnectedness and network effects. In this section, I use the simulation method developed in this paper to measure both of these objects. Specifically, I estimate the effects of an expansion in rural service in Rwanda induced by the introduction of an expanded coverage requirement. I then demonstrate how results from this model can be used to predict impacts in other areas that have yet to receive coverage, by perturbing them to match population density.

Background. A social planner would expand coverage until the point where building any marginal set of towers would not improve welfare. Firms may stop building before reaching this point. A firm is likely to internalize only a fraction of the benefits of expanding the network: in any industry, price discrimination is limited; in network good industries it is often further limited by regulation. And if the market is competitive, benefits from expansion ripple into competitors' networks. While in theory a firm could set interconnection fees to capture the benefits provided to the border of other networks, these fees are often regulated to be near cost. In cases where it is not profit maximizing to expand the network to the social optimal, it may be optimal for a government to regulate the provision of coverage. In this section I evaluate the social and private benefits from a marginal expansion of coverage in rural areas.

In this context, rural areas are less lucrative because of lower demand and higher costs. Rural demand tends to be lower due to lower incomes and population densities; in Rwanda during this period, mean monthly revenue from an urban tower is nearly twice that of a rural tower. Costs are also higher because infrastructure is lower quality. The total annualized cost of owning and operating a tower in Rwanda is

\$51,000 per year, plus \$29,584 for towers that are far from the electric grid that must be powered by generators.⁶⁵

Impact of Rural Expansion in Rwanda. In Rwanda, the regulator required a rollout plan culminating in near-complete coverage. Although license obligations are spelled out in the legal code (Rwanda, 2008), they are likely to have been anticipated by the operator and formed in the course of ongoing discussions, so I do not attempt to evaluate the direct impact of specific obligations.

Ideally, I would compare the revenue and consumer surplus generated under the actual rollout to that generated by the rollout that maximizes profits in absence of regulation. It is computationally infeasible to determine this profit maximizing rollout. Instead, I simulate a suggestive counterfactual: a counterfactual where the operator trims back rollout, and does not build marginal, unprofitable towers. Computing this counterfactual involves three steps. First, I rank each tower by how desirable it is for the firm to build, using a proxy for desirability. Second, I trim the least desirable towers from the rollout plan, and simulate the adoption equilibria that would result. I compute these equilibria for trimming a small number of rural towers (6%) and a large number (12%). This process identifies a set of towers that the firm could have trimmed that would have led to an increase in profits, if it were unconstrained by coverage obligations. My estimate for the effect of the regulation is then given by the welfare difference between the baseline and the counterfactual where these unprofitable towers are not built.

I rank constructed towers by the empirical revenue of the transactions that were transmitted through them ('direct baseline revenue'). The distribution of monthly revenue by tower is shown in Figure 7a. This provides a rough gauge of the desirability of a tower, but does not capture the causal impact of building a tower on revenue: it omits substitution between towers and the effect of coverage on adoption. I determine the causal impact using my simulation method.

⁶⁵Costs based on financial data provided by operators to the Rwandan Utilities Regulatory Agency (RURA, 2011). Building a tower costs approximately \$130,000; I consider the total cost of ownership to operate a tower, which includes operating expenses, annualized depreciation, and a 15% cost of capital. Calculated depreciation assumes lifespans of 15 years for towers, 8 years for electric grid access, and 4 years for generators.

I trim back the rollout by removing successive low revenue towers until the marginal revenue generated is approximately zero.⁶⁶ Figure 7 shows results from two simulations: removing the last 6% of rural towers, and the last 12% of rural towers. Direct baseline revenue of these towers are highlighted in Figure 7a, and full simulation results are presented in 7b. This exercise suggests that the rural towers constructed between 2005 and January 2009 with direct baseline revenue in the lowest 6-12% were profitable, but those in the last 6% were not profitable. Thus, in absence of the coverage obligation I would expect the operator to not build the last 6% of rural towers. I consider the effect on revenue and welfare of building these towers during the sample period.⁶⁷ The building and operation of these towers represented a net present cost of \$496,660 in 2005. Two of these low revenue towers cover border crossing points, for which there was an explicit coverage requirement.

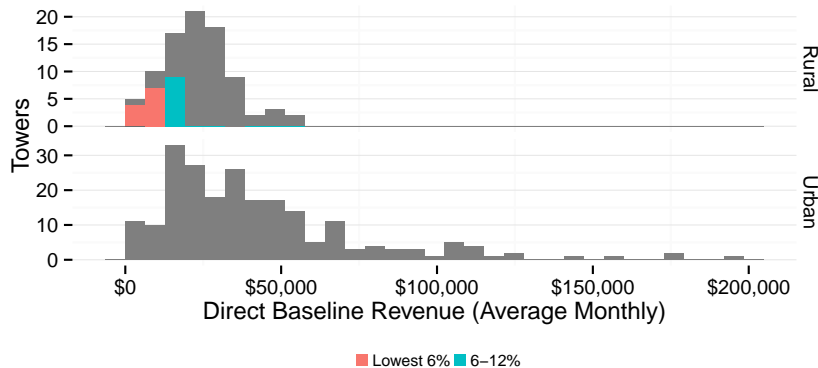
I first compute the progression of coverage omitting this set of rural towers: Figure 7c shows the towers omitted in this counterfactual rollout. I compute each individual's time series of coverage and the resulting link utilities and durations. I then simulate the new equilibrium given this counterfactual progression of coverage, allowing individuals to reoptimize their adoption decisions until an equilibrium is reached. Table 7 presents the results for adoption months, revenue, and consumer surplus.

The change in coverage has an immediate effect on calls: lower coverage increases the hassle cost of placing a call, reducing durations and the utility from calling. Consumers who obtain less utility from calling may also change their adoption decision, which can cause even consumers who were not directly affected by the change in coverage to change their adoption decisions. In the rows of Table 7, I present the baseline

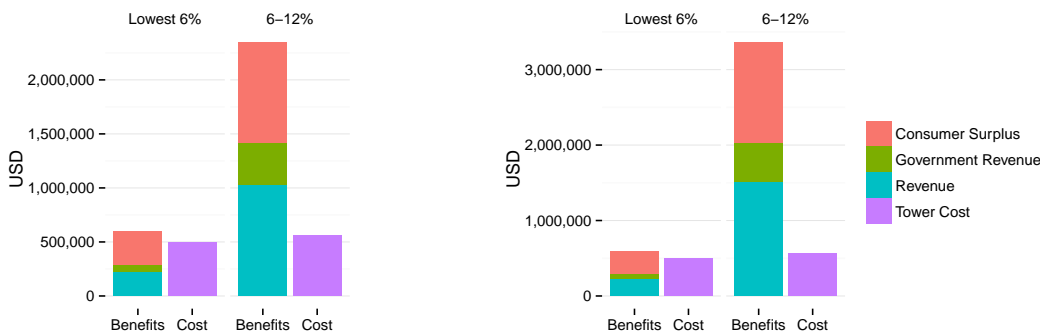
⁶⁶I consider marginal groups of towers rather than individual towers because (i) there is noise in the estimation of coverage which is smoothed out when considering groups, and (ii) it is computationally costly to evaluate the removal of individual towers.

⁶⁷Although the operator would not know the revenue generated by each tower ex-ante, they would likely have precise estimates given their experience in other contexts and given that these towers expand coverage on the margins. Note that based on the data I have I cannot compute revenues and consumer surplus beyond May 2009. The full impact of tower construction on the path of revenues could be more positive if demand is dynamic, or there is a first mover advantage in building out towers in advance of the third operator license being allocated. The full effect on profits could be more negative if demand does not increase in the affected areas and the unprofitable towers continue to lose money in the future.

FIGURE 7. Rural Expansion
(a) Baseline Tower Revenue



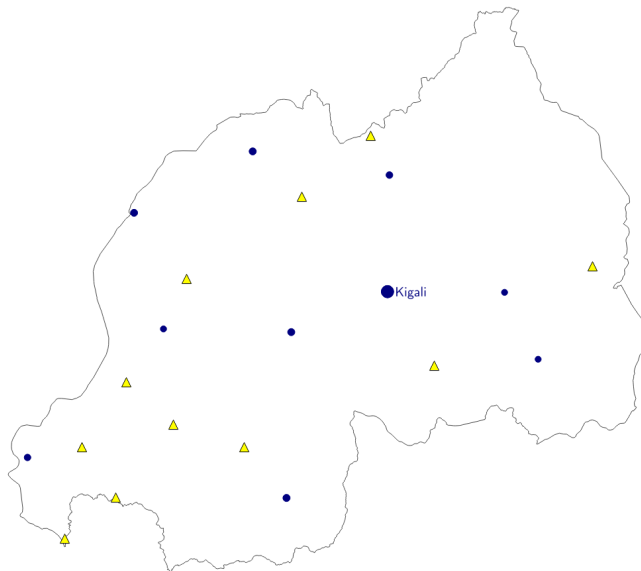
(b) Simulated Costs and Benefits of Building Low Revenue Towers



Low Equilibrium

High Equilibrium

(c) Unprofitable Towers (Lowest 6% Baseline Revenue)



Revenue includes domestic voice calls originating at that tower, billed by the average basket of prepaid rates, averaged over all months the tower was operational. For the counterfactual, I drop the 11 lowest revenue rural towers built during the data. In the map, dropped towers are denoted by triangles and cities are denoted by circles.

TABLE 7. Impact of Rural Service Expansion

Number	All nodes		Nodes by change in coverage	
	> 0.5%pt coverage change	≤ 0.5%pt coverage change		
Adoption Time (mean)				
... with expansion	month	[25.38, 22.16]	[29.40, 25.79]	[24.74, 21.59]
... no expansion, only immediate effect on calls	month	[25.38, 22.16]	[29.40, 25.79]	[24.74, 21.59]
... no expansion, full impact including adoption	month	[25.40, 22.17]	[29.46, 25.83]	[24.75, 21.59]
Total Impact of Expansion	month	-0.02, -0.01	-0.06, -0.04	-0.01, -0.01
Revenue (total)				
... with expansion	million \$	[193.30, 210.76]	[13.23, 15.09]	[180.07, 195.66]
... no expansion, only immediate effect on calls	million \$	[193.15, 210.60]	[13.16, 15.02]	[179.99, 195.57]
... no expansion, full impact including adoption	million \$	[193.08, 210.53]	[13.14, 15.00]	[179.94, 195.53]
Total Impact of Expansion	million \$	0.22, 0.22	0.09, 0.09	0.13, 0.14
... immediate effect on calls	million \$	0.15, 0.16	0.06, 0.07	0.08, 0.09
... added effect through adoption	million \$	0.07, 0.07	0.02, 0.02	0.05, 0.05
Consumer Surplus (total)				
... with expansion	million \$	[37.37, 43.84]	[1.81, 2.49]	[35.56, 41.35]
... no expansion, only immediate effect on calls	million \$	[37.12, 43.59]	[1.69, 2.37]	[35.43, 41.21]
... no expansion, full impact including adoption	million \$	[37.06, 43.55]	[1.67, 2.36]	[35.39, 41.18]
Total Impact of Expansion	million \$	0.31, 0.29	0.13, 0.13	0.17, 0.16
... immediate effect on calls	million \$	0.25, 0.25	0.12, 0.12	0.13, 0.13
... added effect through adoption	million \$	0.06, 0.04	0.02, 0.01	0.04, 0.03
Government Revenue (total)				
... with expansion	million \$	[78.28, 84.41]	[6.98, 7.70]	[71.29, 76.71]
... no expansion, only immediate effect on calls	million \$	[78.24, 84.36]	[6.97, 7.68]	[71.27, 76.68]
... no expansion, full impact including adoption	million \$	[78.21, 84.34]	[6.95, 7.67]	[71.26, 76.67]
Total Impact of Expansion	million \$	0.07, 0.07	0.03, 0.03	0.04, 0.04
... immediate effect on calls	million \$	0.04, 0.05	0.02, 0.02	0.02, 0.03
... added effect through adoption	million \$	0.03, 0.02	0.01, 0.01	0.02, 0.02

Results in each cell reported for the lower bound and upper bound estimate of the equilibrium. Impacts represent the difference in these bounds. I hold fixed the adoption of the 41,225 subsidized nodes (for details see Supplemental Appendix). Utility and revenue reported in 2005 U.S. Dollars, discounted at a rate of 0.9 annually. Consumer surplus includes the surplus utility each individual receives from the call model through May 2009, minus the cost of holding a handset from the time of adoption until May 2009. When splitting the results by change in coverage I use coverage in January 2009.

simulation with the expansion, and two counterfactual simulations, one showing only the immediate impact on calling, and one incorporating the full impact.

The first column of Table 7 presents results for all nodes and the following two columns break down the effect, on individuals whose coverage was substantially affected and on those whose coverage was minimally affected.⁶⁸ The expansion moves the former's adoption forward by an average of 0.06 months in the lower equilibrium and 0.04 months in the higher equilibrium, and the latter's adoption forward by 0.01 months in either equilibrium. I find:

By construction, the building the expansion was unprofitable for the operator. Building the towers shifted bounds on the operator's profits downward by roughly \$274,000.

Rural expansion slightly improved welfare. Building the lowest revenue towers shifts bounds on welfare upward by approximately \$31,789 (lower equilibrium) and \$22,271 (upper equilibrium).

The revenue generated by the expansion was dispersed. Over 59% of revenue comes from individuals whose personal coverage was not substantially affected.

The benefits were too low and dispersed for consumers to finance tower construction themselves. A nationwide consumer group would not realize enough benefits to finance the tower construction themselves: it would reduce bounds on overall consumer surplus by \$188,899 and \$201,932. A related question is whether citizens would be willing to raise local taxes to finance local infrastructure improvements. However, 55% of the consumer surplus from tower construction accrues to individuals whose personal coverage was not substantially affected. If the most affected citizens banded together to raise money for the towers, they would incur a huge utility loss: bounds on their consumer surplus would have declined by \$363,269 and \$362,755; this despite generating substantial benefits both for consumers in other locations, the operator, and the government.

In the Supplemental Appendix, as a robustness check I analyze the expansion under the assumption that the surplus from calls also accrues to receivers ($w = 1$). I find

⁶⁸I define an individual as affected if their coverage changes by more than 0.5 percentage points in the counterfactual, as of January 2009.

that the expansion still reduced bounds on profits, by \$245,365 (lower equilibrium) and \$273,415 (upper equilibrium), and the rollout was socially beneficial, increasing bounds on welfare by \$677,187 (lower) and \$548,432 (upper).

While broader rollout appears to be driven largely by private incentives, I find that a Rwandan coverage obligation led to unprofitable but welfare improving tower construction, because the operator was unable to capture a sufficient amount of the value it generated. An operator that was able to price in a more sophisticated manner (for example, by charging nonlinear or location-specific prices) may be able to internalize sufficient value to serve these marginal areas. If pricing is not sufficiently flexible, it may be optimal for governments to encourage service in these areas.

Impact by Population Density. Areas with low population density are more costly to serve because more towers are needed to cover the same number of consumers. Rwanda’s population density is high at 416 people per square kilometer: it is denser than Rhode Island, Belgium, or Israel. I compute a simple perturbation of the results by scaling the country’s population density. Intuitively, the exercise is to keep Rwanda’s geographic size fixed, but scale the population. When the population is scaled down, a given tower will cost the same and cover the same geographical area, but serve fewer potential subscribers. Instead of scaling down the number of people discretely, for equilibrium Γ , I simply scale down the revenues R^Γ and consumer surplus U_{net}^Γ , holding fixed the costs C . If the population density were scaled by a factor ρ , the predicted impact on revenues and consumer surplus would be:

$$\Delta \tilde{R}_F^\Gamma = \rho \Delta R_F^\Gamma - C$$

$$\Delta \tilde{R}_G^\Gamma = \rho \Delta R_G^\Gamma$$

$$\Delta \tilde{U}_{net}^\Gamma = \rho \Delta U_{net}^\Gamma$$

where ΔX^Γ is the impact of the tower construction on X in equilibrium Γ .⁶⁹

For high population densities, with $\rho > 2.25$, it would have been both socially and privately optimal in both bounds to expand the network, so an intervention to encourage coverage would be inframarginal. For low population densities, with $\rho < 0.94$, it is both unprofitable and welfare reducing to expand the network in

⁶⁹As an approximation, I scale R and U_{net} linearly with population density; the true relationship is likely nonlinear.

both bounds, so that a coverage obligation would reduce welfare. However, in the range $0.96 < \rho < 2.22$ expanding the network would be socially optimal but not be profitable, in both bounds, so an intervention to provide coverage would improve welfare.⁷⁰

10. APPLICATION: OPTIMAL TELECOM TAXATION IN DEVELOPING COUNTRIES

Generating public revenue is a perennial challenge for developing countries, where collection costs are first order. Many governments are confined to a small set of feasible instruments which can be distortionary (Gordon and Li, 2009). However, even in countries with very little other capacity to collect revenue, telecom represents a thriving sector operated by a few formal firms that can easily be taxed. Further, due to concentration arising from network effects the industry is likely to collect rents.

Developing country governments recognize this convenient source of revenue: the mobile industry contributed an average of 7% of government revenue in sub-Saharan Africa as early as 2007 (GSMA, 2012).⁷¹ In addition to standard taxes, governments charge spectrum license fees and specific taxes on telecom equipment, mobile handsets, and airtime. While it is clear that this emerging sector provides a public finance opportunity for poor countries, it is unclear how to best exploit this opportunity. There is a widespread concern that countries may continue to tax telecom heavily in the short term at the expense of long term growth. The former Director of ICT at the World Bank, Mohsen Khalil, voices this concern: “the indirect benefits to the economy of having affordable access to telecommunications services far outweigh any short-term benefit to the budget.”

Two key tax policies affecting adoption are the handset tax and tax on usage (together with import duties on telecom equipment, these represented 66% of tax revenue from telecom in sub-Saharan African countries in 2006 (GSMA, 2012)⁷²). Consumers faced an average adoption tax of 31% and usage tax of 20% (respectively, 48% and 23% in Rwanda) (GSMA, 2012).⁷³

⁷⁰Values of ρ not covered by these three cases have different effects on the lower and upper equilibria.

⁷¹For a sample of 19 countries from which data is available.

⁷²For the 15 countries from which data is available.

⁷³Including VAT, handset import duties, and additional airtime taxes. For a sample of 16 countries from which data is available in 2007.

If telecoms could enforce contracts, they could alter how taxes are exposed to consumers by subsidizing handset purchases. When contract enforcement is costly and consumers use prepaid plans, firms have less flexibility in how taxes are exposed to consumers. Indeed many telecoms argue that high handset taxes slowed the adoption of feature phones and are currently inhibiting the adoption of smartphones and thus the internet. Some argue that the adoption boost driven by lowered handset taxes would result in higher long term tax revenue, and some countries have tried eliminating handset taxes, including Kenya (eliminated 2009, reinstated 2013), Senegal (eliminated 2009), and Rwanda (eliminated 2010). But overall there is little evidence guiding how and how much to tax telecom. In particular, the choice of instruments interacts with network effects: if heavy users provide the most network benefits, a government may want to tax them less. And, the optimal policy is likely to vary over time as the network becomes more attractive to consumers with different characteristics.

This application simulates adoption under a variety of tax policies. I present estimates that are conservative in two respects. First, simulations that make adoption more attractive will tend to underestimate adoption. The previous two applications made adoption less attractive, so that the consumers who adopt in a counterfactual will necessarily be a subset of those who adopted in the data. In this application I also consider counterfactuals which make adoption more attractive, which may have induced individuals outside my data to adopt sooner. Because in the data adoption became more attractive with time, my simulations will cover two periods: a first period where the utility of adopting does not exceed that represented in the data, and a second period in which it may. My simulations will provide the correct estimates for the first period, and will underestimate adoption for the second period. Second, I will make conservative assumptions about tax passthrough. Handset taxes are likely to pass through to consumers, as handsets are offered by a competitive market. I also evaluate results under the two extremes of complete and no passthrough of airtime taxes.

Simulation results are presented in Table 8. The first row presents the baseline tax regime (48% handset tax and 23% airtime tax). The following rows present the

effects of altering the handset tax, first at the original level of usage tax (23%) and then at a higher usage tax (30%). For selected counterfactuals I present the immediate effect of the change in tax policy on individuals, without allowing changes to ripple through the network (“proximal effect”) as well as the full effect after all nodes have reoptimized (“proximal and ripple”). The columns present the revenue accruing to the telecom and government, and the net surplus accruing to consumers in the upper and lower equilibria. I find:

Handset taxes had a substantial cost to consumer surplus and telecom revenues. Relative to a counterfactual with no adoption taxes, handset taxes raised \$17m in the lower equilibrium (\$19m in the upper equilibrium) at the cost of \$27m (\$23m) in consumer surplus and \$15m (\$9m) in telecom revenue. This corresponds with an average welfare cost of \$2.56 (\$1.62) for each dollar of government revenue raised. This is a higher cost than estimates of marginal cost of public funds from the literature, of 1.21 for sub-Saharan Africa and 1.37 for Rwanda (Auriol and Warlters, 2012), suggesting it would be preferable to use alternative instruments to raise these revenues. Since in this model telecoms earn no revenue from handset sales, the entire effect on telecom revenue is driven indirectly, by reduced usage.

Neglecting network effects would substantially underestimate the effects of adoption taxes on telecom revenue, and overestimate their effect on tax revenue. Network ripple effects account for up to 45% of the effect of handset taxation on telecom revenues (45% in the lower equilibrium which is more sensitive to coordination, and 31% in the upper equilibrium). Additionally, ripple effects generate additional government revenue and consumer surplus that would be neglected by a model the only considered individual responses (\$0.8m or \$2.0m in government revenue and \$0.9m or \$4.9m in consumer surplus). Because a naïve estimate would omit these effects it would suggest the average cost of raising a dollar of government revenue from handset taxes would be much lower—\$1.64 in the lower equilibrium and \$1.37 in the upper. Under these estimates handset taxes would have looked nearly as attractive as other tax instruments as reported by Auriol and Warlters (2012).

The government could have increased tax revenue and consumer surplus by shifting from adoption to usage taxes, though this would reduce firm revenue.

Particularly, I consider eliminating the handset tax in 2006 and raising the usage tax to 30%. If the increase in usage tax were not passed through to consumers, social surplus would increase by 8.5% (4%); if completely passed through, social surplus would increase by 1.9% (decrease by 0.2%). Complete passthrough is unlikely because conditional on the tax level, the firm earns higher revenues by not passing the increase through. The welfare cost of increasing the usage tax is similar to that of the original handset tax, \$2.28 (\$1.81) per dollar of government revenue. The Rwandan government ultimately did lower handset taxes and raise usage taxes in 2010; these estimates suggest it may have been optimal to shift these taxes as early as 2006.

A shift from adoption to usage taxes would more than double the consumer surplus accruing to light users. I explore the distributional implications of shifting from adoption to usage taxes in Table 9. I present the baseline tax regime and two alternate tax policies that eliminate handset taxes at the beginning of 2006. I show revenues and consumer surplus for the entire sample, and then for a subsample of heavy users (above the 90th quantile of average daily duration) and lighter users (below the 90th quantile). Under the baseline tax regime, although the top 10% of users account for roughly 58% of telecom revenues and 72-75% of consumer surplus, they account for only 45% of government revenue. Since all users must pay the fixed cost of a handset to join the network regardless of usage, light users end up paying a substantial portion of tax. If instead handset taxes are eliminated in 2006, light users pay less tax and obtain nearly triple the consumer surplus. If this lost government revenue is earned back by raising usage taxes to 30%, the consumer surplus accruing to heavy users is roughly the same as baseline, but that accruing to light users increases by more than 117%. Since potential adopters who are not in my data are likely to be light users, these omitted users would likely benefit from the policy as well in which case these represent underestimates of the full impact of a change in policy.

One factor driving the success of mobile phones among the poor is that usage charges are primarily marginal; these simulations suggest that governments can encourage adoption of these technologies by the poor by taxing on the margin of usage rather than adoption. One feature that drives this result is that at this point in time

the variation in taxable usage was much larger than the variation in taxable handset purchases. The result could be reversed for technologies like smartphones if variation in taxable usage were reduced (e.g., if there are substantial quantity discounts to data usage) or variation in handset purchases increases (e.g., if there is a wider range of qualities available, and users upgrade at different frequencies). However, even in this case it is possible to design a tax policy that encourages adoption. Purchasing a handset only generates network effects if it enables the purchaser to use a new network service. This is the case during the time period I study, as most handsets are purchased by new subscribers who obtain voice service for the first time. This is also the case as developing countries transition to smartphones: when a voice subscriber upgrades from a feature phone to a smartphone they may obtain internet service for the first time. Thus in the presence of public finance concerns, governments may want to differentially tax handset purchases that enable new network services. Such a policy could be implemented in two steps. Products could be categorized by network functionality (e.g., allows voice service, also allows internet service). Individuals' first purchase of a product within each category could be taxed at a lower rate, and subsequent purchases could be taxed at a higher rate. It would be feasible to track this information given that many developing countries now require subscribers to link their identity to their mobile phone account through SIM registration.⁷⁴

11. CONCLUSION

This paper introduces a new method for estimating and simulating the adoption of network goods. I overcome measurement issues that have limited empirical work on network goods using a tractable framework and rich data on the adoption and usage of nearly an entire network of mobile phone users.

I demonstrate this method with three applications. I evaluate a rural adoption subsidy and find that it improved net welfare, and that a large fraction of its impact results from its effects on nonrecipients. These spillovers suggest that adoption subsidies for network goods should be thought of not as targeting individual nodes,

⁷⁴National priorities can determine which network functionalities should be prioritized. While voice service and mobile internet service are clear categories, countries would have to decide whether to similarly prioritize features such as Near Field Communication which allows the use of mobile wallet services like Google Wallet or Apple Pay.

TABLE 8. Optimal Telecom Taxation

Tax Regime	Effect	Revenue (\$m)	Government	Consumer Surplus (\$m)	Underestimated After Date	
Usage Handset	Baseline	[193.30, 210.76]	[78.28, 84.41]	[37.37, 43.84]	-	
23%	48%					
23%	48% until 2007, then 0%	Proximal	[198.85, 214.22]	[65.00, 69.63]	[55.25, 61.33]	2007.1
		Proximal + Ripple	[205.21, 216.46]	[66.89, 70.28]	[60.20, 62.28]	
23%	48% until 2006, then 0%	Proximal	[199.94, 214.51]	[64.00, 68.36]	[57.41, 65.38]	2006.11
		Proximal + Ripple	[206.55, 216.96]	[65.98, 69.09]	[62.66, 66.63]	
23%	0%	Proximal	[201.42, 216.81]	[59.60, 64.20]	[59.85, 65.51]	2006.11
		Proximal + Ripple	[208.36, 219.55]	[61.67, 65.02]	[64.77, 66.41]	
Complete passthrough of usage tax						
30%	48%	Proximal + Ripple	[173.17, 183.34]	[94.57, 99.59]	[31.20, 35.10]	-
30%	48% until 2007, then 0%	Proximal + Ripple	[179.84, 192.16]	[82.62, 87.94]	[49.98, 55.29]	2007.1
30%	48% until 2006, then 0%	Proximal + Ripple	[181.12, 192.99]	[81.90, 86.99]	[51.86, 58.25]	2006.11
30%	0%	Proximal + Ripple	[181.77, 194.33]	[77.33, 82.72]	[55.63, 60.40]	2006.11
No passthrough of usage tax						
30%	48%	Proximal + Ripple	[175.73, 191.60]	[95.85, 103.57]	[37.37, 43.84]	-
30%	48% until 2007, then 0%	Proximal + Ripple	[186.66, 196.78]	[85.55, 89.96]	[60.20, 62.28]	2007.1
30%	48% until 2006, then 0%	Proximal + Ripple	[187.77, 197.24]	[84.76, 88.81]	[62.66, 66.63]	2006.11
30%	0%	Proximal + Ripple	[189.42, 199.59]	[80.61, 84.98]	[64.77, 66.41]	2006.11

Results in each cell reported for the lower bound and upper bound estimate of the equilibrium. I hold fixed the adoption of the 41,225 subsidized nodes (for details see Supplemental Appendix). Utility and revenue reported in 2005 U.S. Dollars, discounted at a rate of 0.9 annually. Consumer surplus includes the surplus utility each individual receives from the call model through May 2009, minus the cost of holding a handset from the time of adoption until May 2009.

TABLE 9. Telecom Taxation - Distribution

Tax Regime	Usage Handset	Sample Split	Revenue (\$m)	Government	Consumer Surplus (\$m)	Underestimated After Date
Usage Handset			Telecom			
23% 48%		All	[193.30, 210.76]	[78.28, 84.41]	[37.37, 43.84]	-
		Above Q90 usage	[114.08, 121.54]	[35.36, 37.62]	[28.20, 31.53]	
		Below Q90 usage	[79.22, 89.22]	[42.91, 46.78]	[9.17, 12.31]	
23%	48% until 2006, then 0%	All	[206.55, 216.96]	[65.98, 69.09]	[62.66, 66.63]	2006.11
		Above Q90 usage	[118.86, 123.71]	[35.87, 37.32]	[33.15, 35.41]	
		Below Q90 usage	[87.70, 93.25]	[30.10, 31.76]	[29.51, 31.22]	
Complete passthrough of usage tax 30%	48% until 2006, then 0%	All	[181.12, 192.99]	[81.90, 86.99]	[51.86, 58.25]	2006.11
		Above Q90 usage	[105.34, 110.23]	[45.51, 47.61]	[28.57, 31.55]	
		Below Q90 usage	[75.78, 82.77]	[36.39, 39.38]	[23.28, 26.70]	
No passthrough of usage tax 30%	48% until 2006, then 0%	All	[187.77, 197.24]	[84.76, 88.81]	[62.66, 66.63]	2006.11
		Above Q90 usage	[108.05, 112.46]	[46.68, 48.57]	[33.15, 35.41]	
		Below Q90 usage	[79.73, 84.77]	[38.07, 40.24]	[29.51, 31.22]	

Results in each cell reported for the lower bound and upper bound estimate of the equilibrium. Sample split by quantile of average daily usage. I hold fixed the adoption of the 41,225 subsidized nodes (for details see Supplemental Appendix). Utility and revenue reported in 2005 U.S. Dollars, discounted at a rate of 0.9 annually. Consumer surplus includes the surplus utility each individual receives from the call model through May 2009, minus the cost of holding a handset from the time of adoption until May 2009.

but as targeting neighborhoods of the graph. I also analyze the expansion of the mobile phone network into rural areas. I find that while most of the expansion of the network appears to be driven by private incentives, an obligation to provide coverage in rural areas led to the building of a handful of otherwise unprofitable towers that improved welfare. Finally, I simulate the effect of different tax policies and find that a shift from adoption to usage taxes would have increased consumer surplus as well as government revenues. Future work can extend the method developed in this paper to guide the design of other network good policies.

REFERENCES

- ACKERBERG, D. A. AND G. GOWRISANKARAN (2006): “Quantifying equilibrium network externalities in the ACH banking industry,” *The Rand Journal of Economics*, 37, 738–761.
- AKER, J. C. (2010): “Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger,” *American Economic Journal: Applied Economics*, 2, 46–59.
- AKER, J. C. AND I. M. MBITI (2010): “Mobile Phones and Economic Development in Africa,” *Journal of Economic Perspectives*, 24, 207–232.
- AURIOL, E. AND M. WARLTERS (2012): “The marginal cost of public funds and tax reform in Africa,” *Journal of Development Economics*, 97, 58–72.
- BANERJEE, A., A. G. CHANDRASEKHAR, E. DUFLO, AND M. O. JACKSON (2013): “The Diffusion of Microfinance,” *Science*, 341, 1236498.
- BIRKE, D. AND G. P. SWANN (2010): “Network effects, network structure and consumer interaction in mobile telecommunications in Europe and Asia,” *Journal of Economic Behavior & Organization*, 76, 153–167.
- BJORKEGREN, D. AND D. GRISSIN (2015): “Behavior Revealed in Mobile Phone Usage Predicts Loan Repayment,” *Working Paper*.
- BLUMENSTOCK, J. (2012): “Inferring patterns of internal migration from mobile phone call records: evidence from Rwanda,” *Information Technology for Development*, 18, 107–125.
- BRYNJOLFSSON, E. AND C. F. KEMERER (1996): “Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market,” *Management Science*, 42, 1627–1647.
- COMIN, D. AND B. HOBIJN (2010): “An Exploration of Technology Diffusion,” *American Economic Review*, 100, 2031–2059.
- CONLEY, T. G. AND C. R. UDRY (2010): “Learning about a New Technology: Pineapple in Ghana,” *American Economic Review*, 100, 35–69.
- DONNER, J. (2007): “The Rules of Beeping: Exchanging Messages Via Intentional "Missed Calls" on Mobile Phones,” *Journal of Computer-Mediated Communication*, 13.
- FARR, T. G., P. A. ROSEN, E. CARO, R. CRIPPEN, R. DUREN, S. HENSLEY, M. KOBRIK, M. PALLER, E. RODRIGUEZ, L. ROTH, D. SEAL, S. SHAFFER, J. SHIMADA, J. UMLAND, M. WERNER, M. OSKIN, D. BURBANK, AND D. ALSDORF (2007): “The Shuttle Radar Topography Mission,” *Reviews of Geophysics*, 45.
- FARRELL, J. AND G. SALONER (1985): “Standardization, Compatibility, and Innovation,” *The RAND Journal of Economics*, 16, 70–83.

- 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*.
- GILLWALD, A. AND C. STORK (2008): "Towards Evidence-Based ICT Policy and Regulation: ICT access and usage in Africa," Tech. rep., Research ICT Africa.
- GOOLSBEE, A. AND P. J. KLENOW (2002): "Evidence on Learning and Network Externalities in the Diffusion of Home Computers," *Journal of Law and Economics*, 45, 317–343.
- GORDON, R. AND W. LI (2009): "Tax structures in developing countries: Many puzzles and a possible explanation," *Journal of Public Economics*, 93, 855–866.
- GOWRISANKARAN, G., M. RYSMAN, AND M. PARK (2010): "Measuring Network Effects in a Dynamic Environment," SSRN Scholarly Paper ID 1647037, Social Science Research Network, Rochester, NY.
- GRILICHES, Z. (1957): "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 25, 501–522.
- GSMA (2006): "Universal Access: How Mobile Can Bring Communications to All," Tech. rep., GSM Association.
- (2012): "Taxation and the Growth of Mobile Services in Sub-Saharan Africa," .
- (2013): "Universal Service Fund Study," Tech. rep., GSM Association.
- ISAACMAN, S., R. BECKER, R. CACERES, S. KOBOUROV, M. MARTONOSI, J. ROWLAND, AND A. VARSHAVSKY (2011): "Identifying Important Places in People's Lives from Cellular Network Data," in *Pervasive Computing*, ed. by K. Lyons, J. Hightower, and E. Huang, Springer, vol. 6696 of *Lecture Notes in Computer Science*, 133–151.
- ITU (2011): "World telecommunication/ICT indicators database," Tech. rep., International Telecommunication Union.
- JACK, W. AND T. SURI (2014): "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution," *American Economic Review*, 104, 183–223.
- JACKSON, M. O. (2009): "Networks and Economic Behavior," *Annual Review of Economics*, 1, 489–511.
- JARVIS, A., H. REUTER, A. NELSON, AND E. GUEVARA (2008): "Hole-filled seamless SRTM data V4," *CIAT*.
- JENSEN, R. (2007): "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector," *Quarterly Journal of Economics*, 122, 879–924.
- JIA, P. (2008): "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 76, 1263–1316.
- KATHURIA, R. AND M. UPPAL (2009): "An econometric analysis of the impact of mobile," *Vodafone policy paper series*, 9, 5–20.
- KATZ, M. L. AND C. SHAPIRO (1985): "Network Externalities, Competition, and Compatibility," *The American Economic Review*, 75, 424–440.
- (1994): "Systems Competition and Network Effects," *The Journal of Economic Perspectives*, 8, 93–115.
- LEE, R. S. (2013): "Vertical Integration and Exclusivity in Platform and Two-Sided Markets," *American Economic Review*, 103, 2960–3000.
- MILGROM, P. AND C. SHANNON (1994): "Monotone Comparative Statics," *Econometrica*, 62, 157–180.
- OHASHI, H. (2003): "The Role of Network Effects in the US VCR Market, 1978-1986," *Journal of Economics & Management Strategy*, 12, 447–494.

- PAKES, A. (2010): “Alternative Models for Moment Inequalities,” *Econometrica*, 78, 1783–1822.
- RIA (2012): “Household Survey,” *Research ICT Africa*.
- ROHLFS, J. (1974): “A Theory of Interdependent Demand for a Communications Service,” *The Bell Journal of Economics and Management Science*, 5, 16–37.
- RURA (2011): “Guidelines for Siting and Sharing of Telecommunication Base Station Infrastructure,” *Republic of Rwanda*.
- (2013): “Active Mobile Phone Subscribers,” *Rwanda Utilities Regulatory Authority*.
- RWANDA (2008): “Official Gazette of the Republic of Rwanda,” 47.
- RYAN, S. P. AND C. TUCKER (2012): “Heterogeneity and the dynamics of technology adoption,” *Quantitative Marketing and Economics*, 10, 63–109.
- SALONER, G. AND A. SHEPARD (1995): “Adoption of Technologies with Network Effects: An Empirical Examination of the Adoption of Automated Teller Machines,” *The RAND Journal of Economics*, 26, 479–501.
- STORK, C. AND M. STORK (2008): “ICT Household Survey Methodology & Fieldwork,” Tech. Rep. 1, Research ICT Africa.
- TOPKIS, D. M. (1978): “Minimizing a Submodular Function on a Lattice,” *Operations Research*, 26, 305–321.
- TUCKER, C. (2008): “Identifying Formal and Informal Influence in Technology Adoption with Network Externalities,” *Management Science*, 54, 2024–2038.
- WAVERMAN, L., M. MESCHI, AND M. FUSS (2005): “The impact of telecoms on economic growth in developing countries,” *Vodafone policy paper series*, 2, 10–24.
- WDI (2013): “World Development Indicators,” *World Bank*.
- YANAGIZAWA-DROTT, D. (2014): “Propaganda and Conflict: Evidence from the Rwandan Genocide,” *The Quarterly Journal of Economics*, 129, 1947–1994.
- ZWANE, A. P., J. ZINMAN, E. VAN DUSEN, W. PARIENTE, C. NULL, E. MIGUEL, M. KREMER, D. S. KARLAN, R. HORNBECK, X. GINE, 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 A. ASSUMPTIONS AND SIMPLIFICATIONS

Handset Sharing. Given the high cost of handsets, sharing is common. 55% of Rwandan phone owners report they allow others to use their handset regularly (Stork and Stork, 2008). An individual may open an account but use it with others’ handsets, by inserting their SIM card, but this practice is rare.⁷⁵ It is more common that a person borrows another’s handset and account.⁷⁶ The model assumes that each node

⁷⁵This allows them their own phone number and balance, but it is difficult to receive calls. A representative survey found that fewer than 1% of individuals in 2007 owned SIMs without handsets (Stork and Stork, 2008), and within the phone data on average there are actually 3% more handsets than accounts active in a given month.

⁷⁶This pattern would include the use of payphones that run on the mobile network, which I omit from this analysis.

in the network represents a unitary entity such as an individual, firm, or household. I assume that this entity weighs the communication benefits accruing to the node against the cost of adoption, and that the communication graph is fixed over time. If multiple people use a particular phone, then the node whose demand I estimate will represent their aggregate demand. The model will correctly account for this demand if the composition of people using a particular phone is fixed over time and the adoption decision takes into account the utility of all users (for example, if the owner internalized the benefits of other users' calls through side payments). If the composition of people using a particular phone changes in response to adoption (say, if a couple initially shares a phone but later obtains separate phones and splits its communication), then the communication graph I estimate will be similar to a weighted average of the underlying networks. In that case, during simulations the nodes will not account for changes in usage when borrowers obtain their own phones, nor coordination of adoption times between the nodes. Modeling changes in phone sharing would require making assumptions about the set of borrowers for each handset over time, the allocation of utility between owner and borrower, and the hassle cost of borrowing a phone to place a call. All of these assumptions would be difficult to defend.

Call Graph. Since the decision to communicate over the phone depends on whether it is possible to communicate in person, the measured call graph is conditioned on individuals' geographic locations. If there were internal migration, these locations would change over time, making it difficult to interpret the measured graph. Permanent internal migration is low in Rwanda over this time period (Blumenstock, 2012).

Adopting a phone may transform an individual's social network—they may keep in touch with friends or family living further away, for example. I uncover the communication graph after any transformation associated with adoption: the graph conditional on phone ownership. The inference in this paper remains valid as long as any such transformation is anticipated and coincides with adoption.

One of the benefits of owning a phone is the option value of being able to place calls, which is valued even if the option is not realized. An extreme example would

be a phone purchased solely for emergency use, which provides expected utility even though it may never be used. Since the utility computed in this model relies on realized calls, it necessarily underweights option value for unrealized calls. It would be possible to include utility from nodes that are on the network but for which no calls have been realized, but this would require a careful decision about which nodes provide option value and which do not. This omission is less problematic than in other settings: the panel is relatively long (4.5 years), so there is time for many communication shocks to be realized. Also, like many developing countries, Rwanda has little in the way of formal emergency response; emergency calls are likely to be directed to close contacts, for whom I'd likely at least observe a link, if not the full utility that link provides.

SMS and Missed Calls. I do not explicitly model utility from SMS and missed calls. If different relationships use different modes of communication, this omission will underweight the importance of SMS and missed-call relationships in the adoption decision. The data suggests that the different modes pick up slightly different relationships: the correlation between a node's total calls and total SMS is 0.53⁷⁷, and the correlation between calls and call attempts within a link is 0.58.

The omission of nonvoice communication could also affect the estimation of parameters based on changes; for example, if subscribers substitute between missed calls and calls as the price or coverage changes. The price for sending an SMS is constant and relatively high throughout the period (\$0.10, the same as a call of 24 seconds under the lowest peak price), and there appears to be little substitution between communication modes as calling prices change. There may be substitution between SMS and calls as coverage improves.

Other Omissions. I omit the cost of charging a phone (the four most popular handsets have more than two weeks of battery life on standby). Accounts must be topped up with a minimum denomination of credit (the minimum was \$0.90 by the middle of the data); I treat these charges as continuous rather than lumpy.

⁷⁷There are a small number of users who use SMS heavily; to prevent these users from skewing the statistic, I compute the correlation omitting the top 1% of SMS users.

APPENDIX B. FUNCTIONAL FORM OF CALLING UTILITY

The form of calling utility determines the utility of being on the network, and thus the adoption decision, and it determines call durations, and thus operator revenue. Given a call shock ϵ_{ijt} , I seek a function describing the utility i obtains from calling j , of the form:

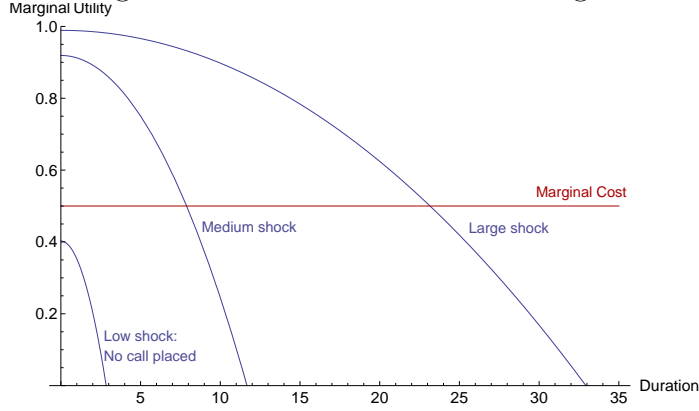
$$u_{ij}(d, \epsilon_{ijt}) = v(d, \epsilon_{ijt}) - cd$$

where d is the number of seconds called and c represents a per second cost. A first restriction on the model is that cost enters linearly, so that the duration choice is separable across contacts. I seek a function satisfying the following properties derived from theory and intuition:

- (1) *No utility from no call*: zero duration yields zero utility: $v(0, \epsilon) = 0$
- (2) *Diminishing returns to duration*: $v(d, \epsilon)$ is concave in d
- (3) *For some values of ϵ and c , a call is placed*. The optimal duration yields nonnegative utility: $v(d^*, \epsilon) \geq 0$ where d^* solves $\frac{\partial v}{\partial d}(d^*, \epsilon) = c$ or is zero.
- (4) *Even if calls were free, you wouldn't talk forever*: there is bounded demand under zero cost: $\frac{\partial v}{\partial d}(\bar{d}, \epsilon) = 0$ for some \bar{d} .
- (5) *Changing the cost of a call affects the extensive decision to call*: this requires that marginal utility be finite at zero: $\frac{\partial v}{\partial d}(0, \epsilon) < \infty$
- (6) *Changing the marginal cost of a call affects longer calls more than short*: $\frac{\partial^2 d^*}{\partial c \partial \epsilon} < 0$
- (7) *The amount of information maps to duration*: given an observed duration \bar{d} , there is a one to one mapping to underlying parameter ϵ , $\epsilon(\bar{d})$, which has an analytic solution that is efficient to compute.
- (8) *Relationships with higher information flows provide more utility*: the optimized utility is increasing in the optimal duration: $\frac{\partial}{\partial d}v(\bar{d}, \epsilon(\bar{d})) > 0$

I use the specification: $v_{ij}(d, \epsilon) = d - \frac{1}{\epsilon} \left[\frac{d^\gamma}{\gamma} + \alpha d \right]$, which satisfies these properties. This form results in a marginal benefit of calling as depicted in Figure 8.

FIGURE 8. Marginal Benefits and Costs of Calling Based on Shock



APPENDIX C. ESTIMATION OF ADOPTION MODEL

The adoption decision implies a set of inequalities on the utility of adopting at different times. To obtain a separate estimate of the price coefficient β_{price} , I form these relations into moment inequalities (see, for example, Pakes, 2010):

$$E \left[Z_{mi} \left(\sum_{k=0}^{K-1} \delta^k \left(\sum_{j \in G_i \cap S_{x+k}} E u_{ijx+k} + w \cdot E u_{jix+k} \right) - \beta_{price} (p_x^h - \delta^K p_{x+K}^h) + (1 - \delta^K) \eta_i \right) \right] \geq 0$$

$$E \left[Z_{mi} \left(\sum_{k=1}^K \delta^{K-k} \left(\sum_{j \in G_i \cap S_{x-k}} E u_{ijx-k} + w \cdot E u_{jix-k} \right) - \beta_{price} (p_{x-K}^h - \delta^K p_x^h) + (1 - \delta^K) \eta_i \right) \right] \leq 0$$

for a set of instruments Z . As in the body of the paper I select $K = 2$. A lower K results in tighter bounds, while a higher K would better smooth any time-varying shocks that could cause an individual to shift their adoption date, like an income shock. (In the two months leading up to adoption, the median consumer gains 3 contacts and the price of a handset declines by \$0.94. The median consumer has 34 contacts when they adopt. In the two month following adoption, the median gains 2 more contacts and the price of a handset declines by \$0.94.⁷⁸)

⁷⁸I identify 41,225 individuals who received subsidized handsets from the government. Because the time-limited subsidy made it extremely desirable for these individuals to adopt when they did, including subsidy recipients leads to extremely wide bounds. Instead, I estimate β_{price} for all individuals who did not receive a subsidy and who subscribed after the first 2 months of the data.

Instruments. I include instruments Z_{mi} that shift the cost of providing service (including geographic slope, and incidental coverage from the presence of electric lines of both the individual and the average of his contacts) and the benefit of joining (the fraction of contacts who received subsidized handsets in the government's 2008 subsidy program). Instruments need to be orthogonal to the unobserved benefit of adopting a phone ($E[\eta_i|Z_{mi}] = 0$); η_i includes forecast error and deviation from the utility predicted by the call model.

Hills block the propagation of cellular signal. Because Rwanda's topography is extremely hilly, the coverage provided by a given cell tower is highly irregular, leading to scattered patches of coverage. These scattered patches can be seen in the coverage maps shown in Figure 1. The interaction of topography and existing infrastructure creates large cost differentials in providing coverage to adjacent areas that are otherwise similar. For example, imagine two villages on either side of a hill far from the electric grid. Since it is much cheaper to operate towers connected to the grid, the village on the side of the hill that faces the electric grid is likely to receive coverage earlier. Although a village very close to the grid is likely to differ in unobservable ways from a village further from the grid, this effect is likely to attenuate quickly with distance from the grid, while cell towers have a range of up to 35 km. Thus, I create an instrument for the coverage provided in remote areas using incidental coverage based on the location of the electric grid: the coverage that would result from building towers along the full network of power lines. These areas of the country had a higher ex-ante probability of receiving coverage because of the interaction of their geographic features with the existing electric grid.⁷⁹ Since factors associated with close proximity to the electric grid could violate the exclusion restriction (these areas tend to be more developed), I use only variation in this instrument for individuals who were at least 5 km from the electric grid.⁸⁰ I also use a more general instrument

⁷⁹Since this instrument makes use of the location of the electric grid and not the locations of towers, endogenous placement is less of a concern than in Yanagizawa-Drott (2014). Endogenous placement would be a concern if desired tower locations influenced the location of the electric grid. Ideally I would use the preexisting electric grid prior to the rollout of the phone network, but I only have the electric grid as of 2008.

⁸⁰The precise exclusion restriction is that individuals in locations further than 5 km from the electric grid that would receive coverage had a line of towers been built along the whole of the electric grid do not in unobservable ways value the network more than those who would not. The instrument

based on topography: the slope of the terrain. I use both variation in an individual's coverage instrument as well as variation in the coverage instrument of their contacts.

I also exploit variation induced by the rural adoption subsidy program. By inducing handset recipients to adopt earlier, the program increased the utility that their contacts would obtain from joining the network. I use variation in the fraction of an individual's contacts who receive subsidies, making the assumption that individuals with more recipients among their contacts do not obtain unobservably different utility from being on the network. I analyze this program in more detail in Section 8 and find that recipients themselves do not appear substantially different from nonrecipients.

I run suggestive tests and find that these instruments have low correlation with observables that could suggest different unobserved benefits of being on the network, including the structure of an individual's communication network and the quality of handset model purchased (see Supplemental Appendix).

leads to scattered patches of coverage throughout the country; see Supplemental Appendix for maps and more details.