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

DANIEL BJÖRKEGREN*

HARVARD UNIVERSITY

This paper uses transaction data from an African mobile phone system to estimate demand for mobile phones as a function of agents' social networks, coverage, and prices. Empirical work on network goods such as mobile phones has historically been limited because it is difficult to measure network effects and difficult to gather data on an entire network of users. This paper overcomes these issues using a new empirical approach and data on the adoption and subsequent usage of nearly the entire network of Rwandan mobile phone subscribers over 4.5 years of expansion. In my model, the utility of adoption is derived from usage. Each individual's set of contacts is revealed by who they call after adopting, and the value of each contact is revealed by the costly decision to place calls.

I use this model to simulate the effects of two policies to encourage network adoption. An adoption subsidy increased welfare by as much as 2%, with a substantial fraction of its impact due to spillovers on nonrecipients. A government coverage obligation had a small, positive effect on welfare, with most benefits accruing to individuals living outside the areas receiving coverage.

JEL Classification Codes: O33, L96, O180, L51

Keywords: network goods, infrastructure, information technology, social networks, big data

*E-mail: danbjork@fas.harvard.edu, Web: http://dan.bjorkegren.com

Revision December 4, 2013. 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, 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 access to data. This work was supported by the Stanford Institute for Economic Policy Research through the Shultz Fellowship in Economic Policy.

1. Introduction

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 by 4.25 billion—from 250 million to 4.5 billion (ITU, 2011b). 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. Although these goods can generate large efficiency gains (Jensen, 2007; Aker, 2010; Jack and Suri, 2011), their allocations are likely to be inefficient due to network effects. 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] In competitive markets, any single firm will likewise internalize only a small share of the benefits it generates. But if 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 some intuition about network effects (Rohlfs, 1974; Katz and Shapiro, 1986; Farrell and Saloner, 1985), there is little empirical work to guide policy choices.[2] Network effects are difficult to measure: 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. It has also been prohibitively costly to gather data on the adoption and usage decisions of everyone on a network. As a result, there remain open 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.

In this paper, I overcome previous limitations using a new empirical approach and 5.3 billion transaction records from Rwanda's dominant mobile phone operator, which held over 88% of the market.[3] I measure the adoption and usage decisions of over one million consumers, and estimate a structural model of demand for mobile phones. Based on calls, I construct the 125 million links that define Rwanda's graph of remote communication at

[1]An 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.

[2]Some examples of empirical work on network goods are Saloner and Shepard (1995), Goolsbee and Klenow (2002), and Tucker (2008).

[3]During this period, the Rwandan market was more concentrated than others, which is empirically convenient in that it is possible to measure nearly the entire network but could lead to concerns about generalizability. This concentration appears to be a result of regulatory policy rather than other features of the Rwandan setting. During this period, the Rwandan regulator restricted entry to two firms; it has since allocated more licenses and the dominant operator's market share has declined to 54% (RURA, 2013).

the time, and estimate how communication over each link changes in response to price and coverage.[4] I then use this structural model to simulate the effects of an adoption subsidy and a government coverage obligation.

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. I assume that the contacts that a subscriber would like to call are the contacts he eventually calls. I infer the utility of communicating with each contact from the calls placed, as a function of coverage and prices. An individual's value of being on the network at a given time is derived from the value of communicating with his contacts who are on the network at that time.[5] I overcome the traditional problem in estimating network effects in the presence of correlated traits or shocks by inferring the value of the network not merely from the presence of links with contacts, but from the costly calls placed across those links. These calls represent the actual source of value provided by the network. In the communication graph I derive, links represent actual flows of utility, not just rough proxies whose correlation with adoption is likely to pick up correlated attributes or shocks.[6]

Second, I model the decision to adopt a mobile phone. In choosing when to adopt, consumers weigh the increasing utility of communicating with the contacts on the network against the declining cost of purchasing a handset. Given this tradeoff, the date a consumer adopts reveals a second estimate of the value of the network: a consumer will adopt when the value from joining the network most exceeds the cost of a handset. To identify this second estimate, I use plausibly exogenous variation affecting the utility of usage and adoption. I use variation in contacts' adoption arising from a targeted government adoption subsidy, as well as variation in coverage induced by Rwanda's hilly topography, in a manner similar to Yanagizawa (2012).[7] The value of the network revealed by adoption is consistent with the value revealed by calling.

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 the full effect as it ripples through the network and across physical space.

[4]Using an algorithm analogous to triangulation, I infer each subscriber's desired calling locations from the locations of the cell towers used once coverage has expanded, and compute the coverage available at these locations back in time. I exploit the fact that over the period of interest coverage expands dramatically—from 60% to 95% of the country's land area. My location estimates are based mostly on cell towers used when coverage is near-complete. They thus reveal the set of desired locations conditional on near-complete coverage, assuming these locations are stable over time.

[5]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).

[6]Conveniently, calls are billed on the margin, by the second, so the duration spoken along a link reveals a lower bound on the caller's willingness to pay to communicate with the recipient.

[7]This second estimation strategy is similar to Tucker (2008), which identifies network effects for a video messaging system using a shock to its value from another use.

I use this approach to answer two sets of policy questions, by simulating how adoption would have proceeded if conditions were less favorable. 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. I first use phone data to determine how subsidized handsets were ultimately used, and then use 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—in particular, contacts of recipients account for more than 62% of the effect on revenue. Although the bounds are wide, the subsidy improved welfare, in a low case by \$191,108 (0.06%), and in a high case by \$5.6 million (2%).[8]

I also analyze the welfare implications of providing coverage to rural areas and the degree to which a network operator is able to internalize the value created. If a monopolist operator's ability to price discriminate is limited by technology or regulation, it can be optimal for a government to require the provision of coverage to areas that would otherwise be unprofitable. I find that in Rwanda, a government coverage obligation led to the building of a handful of rural towers that were unprofitable for the firm but welfare improving for the country. The impact was small, shifting bounds on welfare upward by at least \$179,381 (0.06%). The impact was also extremely dispersed: over 65% of the gain in consumer surplus accrued to individuals living outside the areas receiving coverage; some of these individuals called in to the covered areas and others were affected indirectly. Because of this dispersion, it would have been difficult for local communities to raise the funds to build the towers themselves.

That mobile phones have spread rapidly across varied environments suggests that failures by consumers and firms to internalize network effects in this case have not prevented widespread adoption. Yet because adoption is widespread—worldwide telecom spending was \$4.7 trillion in 2012—even small inefficiencies can have large welfare consequences (TIA, 2012). Also, it is difficult to gather data about the many useful network goods that have not diffused successfully. Analysis of mobile phone networks can guide the design of policies to achieve more efficient adoption of these goods.

Although the method I present uses network structure revealed by adoption and usage, it can also be applied to goods that have yet to be adopted. 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

[8]Due 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.

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 outline how the adoption of mobile phones can inform policy for mobile internet service.

This paper connects with several literatures:

Between $800-900 billion is spent annually on infrastructure in developing countries (Bhattacharya et al., 2012). Mobile phone networks represent a case where governments successfully leveraged private investment to provide infrastructure. But in order to guide investment towards socially optimal ends, governments need to know the structure of both private incentives and welfare. The analysis in this paper reveals the structure of both of these objects for an important component of infrastructure. In this manner, this paper connects to work analyzing how best to provide infrastructure (Kremer et al., 2011; Olken, 2007).

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 connects to a literature analyzing the rapid spread of information and communication technologies (ICTs) across developing countries (Aker and Mbiti, 2010). While many studies analyze the impacts of ICTs (Aker, 2010; Jensen, 2007), this paper analyzes the interplay between demand and service provision. In this regard it connects to Batzilis et al. (2010), which finds that the expansion of mobile phone coverage in Malawi is correlated with demand and cost variables.

A growing literature analyzes the impact of social networks on economic behavior (see Jackson, 2009). My paper is conceptually related to Banerjee et al. (2012), 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. This "big data" from developing countries overcomes some limitations of traditional sources of data (e.g., Zwane et al., 2011), and can also answer broader economic questions that could not be answered with equivalent data from a developed country. In developed economies, economic agents generally face many alternatives, leading to first order selection issues, since any given data source represents only a small slice of an agent's economic activities.[10] Within a developing economy, a single

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

[10]For 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.

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 rollout of mobile telephony across sub-Saharan Africa has been astoundingly successful, leading to both historical and forward-looking questions. For example, some services based on mobile phone platforms have seen great success (such as the M-PESA mobile money system in Kenya; see Jack and Suri, 2011), but success has not been easily replicated. We are still learning what makes network goods successful. This paper looks backwards, to better understand the provision of mobile phones, and forward, to guide the provision of new network goods.

The paper proceeds as follows. The next section describes the expansion of mobile phone networks worldwide and in Rwanda. Section 3 describes the data I use. Section 4 presents stylized facts about mobile phone usage in Rwanda. Section 5 introduces a model of phone adoption and subsequent usage. Section 6 describes the procedure I use to estimate the parameters of this model and the country's communication graph. Section 7 describes how the estimated model and communication graph can be used to simulate the effects of counterfactual policies. This simulation method is then used in two applications: Section 8 analyzes the effect of an adoption subsidy, and Section 9 analyzes operator incentives to provide service in rural areas. Section 10 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. While fixed line networks were often operated by monopolists, regulators began to liberalize mobile markets, and competition began to develop: by 2009, 67% of countries had three or more operators (GSMA, 2009). Handset prices were initially expensive, but gradually began to fall with reductions in component costs and economies of scale. The worldwide average selling price of a handset declined by more than 50% between 2002 and 2006 alone (Frost, 2009). This decline made phones accessible to poorer consumers, and operators adapted to this broader base of potential subscribers by expanding coverage beyond urban centers. From 2000 to 2008, the percentage of Africans covered by a mobile signal increased from 25% to 59% (ITU, 2009). Operators also reduced prices: across developing countries, mobile communication prices dropped by 22% between 2008 and 2010 (ITU, 2011a).[11] 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 during 2005-2009 provides in many ways an ideal setting to understand the spread of mobile phones in developing countries. Because the Rwandan regulator restricted entry,

[11]Based on a basket of 30 outgoing calls at peak, off-peak, and weekend rates, a combination of on-network and off-network calls, and 100 SMS.

the market during this period was extremely concentrated: the mobile operator whose data I use held above 88% of the market, so its records reveal nearly the entirety of the country's remote communication. While most available mobile phone data is from extremely short panels, 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. There are few alternatives to the mobile phone for remote communication: the fixed line network is small (with penetration below 0.4%), and mail service is insignificant.[12] There was significant 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: 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 Genocide and Civil War, the new Rwandan government attempted to spur 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, but in the early years the network remained accessible to only the elite. Coverage was sparse in rural areas, handsets were expensive (and not subsidized as with U.S. plans), and although most accounts were prepaid, network access fees were high.

In this context many programs were developed to expand access to the poor. To address the high cost of handsets, the operator in 2004 introduced payphones that ran on the mobile phone network and provided financing to entrepreneurs to become payphone operators (Tuvugane). To provide access in areas with poor coverage, between 2006 and 2007 the operator and the Grameen Foundation collaborated to outfit standard mobile phones with high powered antennas to be operated as payphones by rural shopkeepers (Village Phone or Tel'imbere; see Futch and McIntosh, 2009). A 2008 government program distributed heavily subsidized handsets to rural areas; this program is analyzed in detail in Section 8.

[12]The 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).

Two changes in the market influenced further rollout:

Global handset prices began to decline. 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, then less than one month's consumption. This decline made mobile phones accessible to broad segments of the population.

Regulatory changes induced in 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; after a fraud scandal and three significant changes in ownership, the company reached a maximum of 20% of market share for a brief period in 2010; in 2011 its license was revoked for failure to meet obligations, and the business was liquidated. In combination with providing a second license, the government attached minimum coverage obligations to the first operator's license. The combination of regulation and the competitive threat added to the incumbent operator's incentive to look beyond the urban elite.[13]

The dominant operator adapted, changing pricing structures to accommodate lower income users and expanding 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[14], and topping up a minimum of $4.53 when credit ran low.[15] By the middle of 2008, essentially all nonmarginal charges had been removed, talk time was billed by the second (so that the shortest call cost less than $0.01), and the minimum top up amount was reduced to $0.90. These adaptations were accompanied by an expansion of coverage, shown in Figure 1. 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%. 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. Figure 2 shows the trend of mobile and fixed line telephone subscriptions in Rwanda. Figure 3 shows the changes in prices, coverage, and network adoption.

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

[14]Cost of the first minute under the per minute peak rate. "Per minute" calls were billed by the first minute and then each half minute thereafter.

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

3. Data

This project uses several data sources:

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 88% of the market. 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.[16] There are 5.3 billion total transactions during this time period. For each transaction, the data reports:[17]

- Two anonymous identifiers for sender and receiver, corresponding to the phone number and handset
- Handset models
- Date and time stamp
- Duration (for voice calls)
- Cell towers used at the start and end of the transaction
- The incurred charge (for transactions before August 2008)

Cell tower locations: The cell tower identifiers can be linked to geographical coordinates provided by the operator. 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 Appendix C.

Individual locations: I infer each subscriber's set of geographical locations using a version of Isaacman et al. (2011)'s 'important places' algorithm that I have modified to improve performance in rural areas. The procedure is detailed in Appendix D.

Coverage maps: I create raw coverage maps by computing the region visible by the set of towers live in each month, using 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 smoothed coverage maps which represent the coverage available within a short walk of a given location, which are used to compute the quality of coverage available to an individual. I also compute two instruments for coverage, incidental coverage from the placement of the electric grid and the slope of the surface. For details, see Appendix A.

[16]There are 2,092,477 accounts ever referenced in the data, but many do not appear to represent active accounts. For the analysis, I omit the 528,737 accounts that have made fewer than 10 outgoing calls, and 38,679 further accounts for which the time spanned between the first and last observed transaction is less than 90 days (some of these are short term visitors to the country). This results in a sample of 1.5m accounts.
[17]Two months of data are missing from the call records: May 2005 and February 2009; and one month is missing from the billing records: October 2006.

Handset prices: I create a monthly handset price series for 160 models in Rwanda using data from three sources: historical versions of the operator's website, operator sales records, and sales records from an independent phone store in Kigali. I compile these prices into a monthly handset price index $p_t^{handset}$, 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 a handset of comparable quality. For details, see Appendix E.3.

Operator billing policies: Details on the operator's historical billing policies are obtained from several sources, including archived versions of the operator's web site, reports from the government regulator, and news articles. The resulting billing model was checked against billing records and adjusted until it fit.

Household surveys: I use several nationally representative household surveys to provide background information: DHS and government surveys (EICV) from 2005 and 2010, and Research ICT Africa's 2007 survey about technology usage (Stork and Stork, 2008).

4. Patterns of mobile phone use

The use of mobile phones in developing societies has adapted to a few key features: incomes are low, other communication methods are costly, and handsets represent a large investment.

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. Missed calls are used to communicate simple information. Despite a popular perception that the poor use phones only for business, calls are primarily social: 92% of subscribers report that the main purpose of the last 10 calls was social, and 90% report that most calls are to family and friends, according to a representative household survey (Stork and Stork 2008). Most calls cover a short distance: roughly 70% of calls are between towers closer than 5 km.

The average usage profile is described in Table 2.

The primary unit of observation is an account, which corresponds to a phone number. Although accounts are prepaid and not explicitly linked to individuals, in the Rwandan context it is natural to think of an account as an individual: there is one major operator and there is a disincentive to switch phone numbers.[18] In the rest of the paper I will refer to accounts as individuals or nodes.

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

[18]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. There are an average of 1.03 accounts per user (Gillwald and Stork, 2008).

for remote communication, which I refer to as the communication graph.[19] Given that most calls are social calls to family and friends, I model communication as arising from the desire to keep in touch with a fixed set of contacts.[20] I assume that I observe 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.[21]

The prepaid billing structure is empirically convenient in that the calling party always pays on the margin for a call,[22] 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. 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

[19]This graph may not correspond with the social network that would be revealed through survey methods used in the literature: a husband and wife may communicate intensely face to face and have no need to call each other. Any network used in empirical work is a projection; this project analyzes demand for calling, for which the call graph is exactly the projection of interest. 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).

[20]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 coincides with adoption and would be predicted by the individuals.

[21]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. I do not model this. This omission is less problematic than in other settings for two reasons. First, the data is a relatively long panel (I use 4.5 years of data), so there is time for many information shocks to be realized. Second, like many developing countries, Rwanda has little in the way of formal emergency response. Emergency calls are much more likely to be directed to close contacts, for whom I'd likely observe realized calls.

[22]Most calls are billed by the second after February 2006, and by the first minute and each subsequent half minute before.

a subscriber would reduce calls to existing contacts, resulting 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 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 decreasing the calling price by \$0.005 per minute. For comparison, the median number of contacts is 61, and the final peak calling rate is \$0.23 per minute.[23]

Low substitutability is reasonable in this setting: penetration is still low, so a substantial fraction of expansion represents new households adopting phones rather than households with phones purchasing additional phones. Also, 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. Links would likely be more substitutable in a mature network, or with relationships that are more transactional rather than social. To simplify the model, I assume the utility obtained from a contact is independent of the state of other contacts on the network. I will model the underlying desire to communicate along each link as arising from a stationary distribution.

Adoption. Joining the network entails opening an account and investing in a handset. It is easy and cheap to open a mobile phone account. A prepaid account can be opened with any of the operator's agents, which are present in even small towns, by buying a Subscriber Identity Module (SIM card) for roughly \$1. In order to use the account, the SIM card must be placed into a handset. Any transactions placed using the handset are billed to the account associated with the SIM card that is currently inserted, and the handset will receive transactions sent to the associated account.

Handsets are expensive, representing 3.5 months of the average person's consumption in 2005. In many Western countries, operators subsidize handsets for consumers and either lock the devices to be used solely on their network, or lock users into a postpaid account contract. In the Rwandan context and many other African contexts, it would be costly to manage postpaid contracts at scale; instead, almost all accounts are prepaid and handsets were offered at retail prices. Most handsets are mainstream, imported models, and the handset market is competitive; I treat handset prices as exogenous and do not model the handset market.

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 plan ahead when considering adoption: when asked in 2007, 88.9% of individuals without phones planned to purchase a phone in the future. I model adoption as a dynamic decision, where individuals incorporate expectations of future improvements into the adoption decision.

[23]Complementarity and substitutability could coexist in different parts of the network, in which case this test and my estimates would identify an average of the two effects.

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.

I focus on voice calls and do not explicitly model the utility from nonvoice transactions such as SMS and missed calls. To the extent these transactions are important, my call utility estimates serve as a proxy for total communication utility. Though important in other contexts, in Rwanda text messaging or SMS represents less than 13% of revenue and 16% of transactions.[24] 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. SMS was quite costly ($0.10 per message) and prices did not change. Only calls that are answered incur a charge; subscribers 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 problems or inability to connect), I do not explicitly model missed calls. Since I have no information about foreign subscribers, I also do not model the small fraction of international calls.

Given the high cost of handsets, sharing phones is common: 55% of phone owners report they allow others to use their handset regularly (Stork and Stork, 2008). That subscribers are willing to spend significant sums of money to buy a personal handset suggests that the hassle cost of borrowing is high; thus, although borrowing is common, the actual volume of calls due to borrowing is likely to be low. Modeling phone sharing would require making assumptions about the set of borrowers for each handset, the allocation of utility between owner and borrower, and the hassle cost of borrowing a phone to place a call. Since it would be difficult to defend these assumptions, I omit the possibility of handset sharing.

An individual may learn about the benefits of using a phone from observing the usage of those around him; I do not model this. For more discussion about these simplifications, see Appendix F.

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 in the environment that improved the utility of communicating across a given link, specifically the expansion of coverage and reduction of calling prices.

Let G be the communication graph (social network). Each individual i has a fixed set of contacts $G_i \subset G$, where a directed link $ij \in G$ indicates that i has a potential desire to call j over the mobile phone network. Let S_t be the subset of nodes subscribing in month t.

[24] Revenue statistic based on the period of data where charges are reported, which covers 2005-August 2008.

Calling Decision. At each period t, 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. The shock is drawn from a link-specific distribution, $\epsilon_{ijt} \overset{iid}{\sim} F_{ij}$ that will be specified later.[25] Given the shock, i chooses a total duration $d \geq 0$ for that month, solving:

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

where $v(d, \epsilon)$ represents the benefit of making calls of a total duration of d 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.[26]

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

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

where β_{call} represents call price sensitivity, and $h(\phi_{it}, \phi_{jt})$ represents the hassle cost given the caller and receiver's level of 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.[27] I parameterize the hassle cost linearly:[28]

$$h(\phi_{it}, \phi_{jt}) = \beta_{coverage.from} \phi_{it} + \beta_{coverage.to} \phi_{jt} + \beta_{coverage.interaction} \phi_{it} \phi_{jt}$$

[25] I model communication shocks as independent across network links and time. It is clear that in the true data generating process these shocks will be correlated over links and time. However, I am primarily concerned with the durations of calls and not their timing; this distribution ends up being collapsed into its expectation for the adoption decision.

[26] There 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 10 properties that a reasonable functional form of utility from telephone calls should satisfy (see Appendix G), which leads to the selected form.

[27] Using an algorithm analogous to triangulation, I identify the set of most used locations for each individual. Around each location, I compute the fraction of area receiving coverage using a two-dimensional Gaussian kernel with radius 2.25 km. I then compute a weighted average of this fraction over the individual's locations, weighting each location by the number of days calls were placed from that location. For more details see Appendices A and D.

[28] Rwanda 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.

Given this functional form, calling prices, and coverage of both sender and receiver affect both the frequency and duration of calls. In month t, i calls j if his desire to communicate is strong enough:

$$\epsilon_{ijt} > \underline{\epsilon}_{ijt} := \frac{\alpha}{1 - \beta_{call} p_t - h(\phi_{it}, \phi_{jt})}$$

and the length of the call increases with the desire to communicate:

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

If the desire to communicate is not strong enough, i will not place a call ($d_{ijt} = 0$).

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

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

where to simplify I write $\boldsymbol{\phi}_t$ to represent 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:

$$u_{it} = \sum_{j \in G_i \cap S_t} Eu_{ijt}(p_t, \boldsymbol{\phi}_t) + w \cdot Eu_{jit}(p_t, \boldsymbol{\phi}_t) + \eta_i$$

where u_{ijt} represents calls from i to j (which i pays for), u_{jit} represents calls from j to i (which j pays for), and $w \in \{0, 1\}$ specifies whether recipients value incoming calls.[29] η_i represents an idiosyncratic benefit from being on the network that is observed by the individual but not to the econometrician, with $E\eta_i = 0$. Each month that i is not on the network he receives utility zero.

Individual i chooses when to adopt by weighing the discounted stream of these benefits against the price of a handset, which is represented by the price index $p_t^{handset}$.[30] Then, i considers the utility of adopting at time τ to be:

$$U_i^{\tau} = \sum_{t=\tau}^{\infty} \delta^t Eu_{it}(p_t, \boldsymbol{\phi}_t) - \delta^{\tau} \beta^{handset} p_{\tau}^{handset}$$

where $\beta^{handset}$ is the price sensitivity parameter for purchasing a handset, which may differ from the sensitivity for marginal calling prices in the presence of credit constraints or

[29] I assume individuals have no private information about forthcoming call shocks, and measure the utility of being on the network as an expected utility rather than realized utility. I could compute realized utilities u_{ijt} for months where both i and j are subscribers, but not for counterfactual periods, which are needed for estimation.

[30] The price index uses the prices of 160 models in Rwanda, weighted by sales, accounting for quality changes. See Appendix E.3 for details.

if individuals obtain value from nonvoice transactions.[31] The adoption decision represents an optimal stopping problem, which requires assumptions on expectations that will be described in the following section. There are likely multiple equilibria in the adoption decisions of the network; my estimation procedure uses necessary conditions for equilibrium and will be valid in the presence of multiple equilibria. The next section describes how I estimate the parameters of this model.

6. Estimation

Individuals choose when to adopt a mobile phone, and if they adopt, how to use the phone. In the usage decision, I use data on phone calls to estimate the country's latent communication graph and how usage responds to prices and coverage. I estimate responsiveness to prices and coverage using time series variation in both quantities. I resolve both an endogeneity problem and a selection problem using an analogue of fixed effects at the link level, and estimate parameters using maximum likelihood.

In the adoption decision, individuals weigh the price of a handset against the discounted stream of calling benefits it provides. Calling benefits are computed from the model of usage. I identify individuals' responsiveness to handset prices using plausibly exogenous variation affecting the utility of usage: I use variation in coverage induced by Rwanda's hilly topography as well as variation in contacts' adoption arising from a targeted government adoption subsidy. I assume that individuals' expectations of future utility deviates from perfect foresight only by a mean zero error, which allows me to estimate the adoption decision using a simple moment inequality approach.

Identification. In the product use decision, the operator changes prices and quality (coverage) in response to demand. Here, the most significant change in demand is the changing composition of marginal consumers: over time, less talkative and more price sensitive individuals adopt, leading the operator to steadily reduce prices and improve coverage. Since all of this variation is in the time series, a naive estimate would find elasticities with signs in the wrong direction: average durations decrease over time despite improvements in coverage and reductions in price. In many settings this selection problem would naturally be addressed by using characteristics such as income or occupation to get at the underlying differences between individuals. Since I do not observe a standard set of individual characteristics, and—more fundamentally—individual characteristics would be too aggregated to be useful since I am model at the level of links, I address this problem in a manner analogous to using fixed effects.[32] I estimate the parameters of call shock distribution for each link (using

[31] I assume that once purchased, a handset lasts forever; depreciation would be picked up in $\beta^{handset}$.

[32] A primary purpose of the call utility model is to compare the utility of different links *within* an individual. We might know that i is a rich, 45 year old male banker in the capital and j is a 40 year old female school teacher living in a different city, but their calling behavior and resulting utility would depend greatly on not just characteristics of i and j, but characteristics of the relationship ij: whether they are siblings, colleagues,

link-specific means and node-specific variances), so that the response of durations to prices and coverage is estimated using within-link variation.[33]

In the adoption decision, I estimate how adoption responds to handset prices relative to the utility of communicating with contacts. I use variation in the time of adoption induced by plausibly exogenous shifters of the utility of being on the network. I exploit variation in the cost of providing coverage to different areas due to Rwanda's hilly geography, in a manner analogous to Yanagizawa (2012), as well as in the contacts on the network induced by a government adoption subsidy program.

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. Since factors associated with close proximity to the electric grid could violate the exclusion restriction (these areas tend to be more urban), I use only variation in this instrument for individuals who were at least 5 km from the electric grid.[34] I also use a more general instrument based on topography: the slope of the landscape. 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 an adoption subsidy program. In the first four months of 2008, the Rwandan government distributed 53,352 heavily subsidized handsets in rural areas, then roughly 8% of the stock of handsets in the country. By inducing handset recipients to

in a lending relationship, or in a romantic relationship. Even within those categories, huge variances in the natures and intensities of relationships would filter through to duration and utility estimates.

[33]There is also an issue with selection based on changes: it is likely that different individuals have different responsiveness to coverage. Those that are highly sensitive will wait until there is sufficient coverage in their area. But I observe calling changes only for people who had already subscribed. This would tend to bias estimates of coverage responsiveness downward. This could be addressed by estimating the response to coverage in the adoption decision. I have not done this because it adds substantial complexity.

[34]The 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 leads to scattered patches of coverage throughout the country; see Appendix B for maps and more details.

adopt earlier, the program increased the utility that their contacts would obtain from joining the network. I exploit variation in the fraction of an individual's contacts who receive subsidies, assuming that the contacts of recipients 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 discuss these instruments more in Appendix B and present the evidence on the exclusion restriction.

The main cost of adoption is the price of a handset. The set of available handsets is driven by the global market; Rwanda is a small market, handsets are mainstream, imported models, and local price trends are consistent with global trends. In contrast to markets like the U.S., handsets are sold at retail price, so there is less scope for adjusting prices. Nonmarginal fees associated with using the network did change over this time, but these were small relative to the price of a handset.[35]

Estimation Procedure

Calling Decision. 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 $1-q_i$, which corresponds to not calling regardless of the cost.[36] The parameter q_i thus controls the amount of censoring independent of cost. (The utility function parameter α controls the amount of censoring that depends on cost.)

The calling decision has 9 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 (q_i, σ_i), and I assume that the shape and sensitivity parameters are common to all links (γ, α, β_{call}, $\beta_{coverage.from}$, $\beta_{coverage.to}$, $\beta_{coverage.interaction}$). 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:[37]

$$\epsilon\left(d\,|p_t, \phi_{it}, \phi_{jt}\right) = \frac{d^{\gamma-1} + \alpha}{1 - \beta_{call} p_t - h(\phi_{it}, \phi_{jt})}$$

[35]Before June 2007, subscribers needed to add roughly \$4.53 (2500 RwF) 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 (500 RwF) 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.

[36]On average, there is a call across a given link only 12% of months. If I used only a common 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.

[37]p_t represents the per-second price of a call. Prior to February 2006, calls were billed by the first minute and each subsequent half minute; after, subscribers could opt in to per second billing. Modeling the per-minute charges would add significant complexity, so instead I impose an equivalent per second price based on an a basket of calls for all accounts before the introduction of per second billing. I assume the average call is 30 seconds.

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, so that a month without a call has likelihood:[38]

$$F_{ij}\left[\epsilon\left(1\,|p_t,\phi_{it},\phi_{jt}\right)\right]$$

If the call shock is large enough, i will place a call; the higher ϵ, the longer the call duration. The likelihood of an observed call of duration d_{ijt} from i to j in month t is:

$$F_{ij}\left[\epsilon\left(d+1\,|p_t,\phi_{it},\phi_{jt}\right)\right] - F_{ij}\left[\epsilon\left(d\,|p_t,\phi_{it},\phi_{jt}\right)\right]$$

I estimate the 6 common parameters and as well as the 127.6 million distribution parameters defining the communication graph. To make estimation tractable, I perform two steps. First, I jointly estimate common and distribution parameters for a random subset of 1,500 nodes and their full set of 92,386 links (representing a total of 2.5 million link-month observations). Then, I impose the common parameters estimated in the first step to estimate the remaining distribution parameters for the full sample of 1,525,061 nodes and their 124.6 million links (representing a total of 4 billion link-month observations. The median number of observations per node is 637 and per link is 45, leading to a median of 41 observations per parameter[39]). The individual likelihoods are separable conditional on the common parameters, so this latter step is computationally much less demanding than performing a full joint estimation.

These parameter estimates allow me to compute the expected duration along the relationship ij conditional on calling prices and coverage:

$$Ed_{ij}(p_t,\boldsymbol{\phi}_t) = \int_{\underline{\epsilon}_{ijt}}^{\infty} d(\epsilon,p_t,\boldsymbol{\phi}_t)\cdot dF_{ij}(\epsilon)$$

as well as expected utility:

$$Eu_{ijt}(p_t,\boldsymbol{\phi}_t) = \int_{\underline{\epsilon}_{ijt}}^{\infty}\left[d(\epsilon,p_t,\boldsymbol{\phi}_t)\cdot\Big(1-\beta_{call}p_t - h(\phi_{it},\phi_{jt}) - \frac{\alpha}{\epsilon}\Big) - \frac{1}{\epsilon}\frac{d(\epsilon,p_t,\boldsymbol{\phi}_t)^{\gamma}}{\gamma}\right]dF_{ij}(\epsilon)$$

[38]Note that chosen durations can only be integers, so there is a range of epsilons that map to the same observed duration d; since utility is concave, the duration will be rounded down.

[39]Note that the 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. This lower tail of links could lead to an incidental parameter problem and affect the consistency of estimates: the 25th percentile of observations per parameter is 21 and the 1st percentile is 6. One way to get around this problem is to estimate the communication graph only for the smaller subgraph that subscribed by a certain date, so there is sufficient data—say, individuals that subscribed by 2008—and then simulate counterfactuals of adoption only up to that point. This approach would work well to simulate counterfactual policies affecting early adoption, but the counterfactuals I run in this paper primarily impact adoption in the later parts of the data. See Table 4 for the quantiles of the observations per link, node, and observation.

where $d(\epsilon, p_t, \boldsymbol{\phi}_t) = [\epsilon\,(1 - \beta_{call} p_t - h(\phi_{it}, \phi_{jt})) - \alpha]^{\frac{1}{\gamma - 1}}$.

I compute these expectations for all relationships ij, using the common time path of calling prices, and the paths of coverage specific to caller i and receiver j. The integrals are evaluated using Monte Carlo draws.[40] The utility of owning a handset in a given period, u_{it}, is derived from these relationships.[41]

Adoption Decision. In choosing when to adopt, individuals weigh the price of handsets against the discounted stream of benefits from being on the network. These benefits will continue after the data ends, since a purchased handset will last beyond the end of the data. Because the network continues to grow, these benefits are nonstationary. I estimate the adoption decision using moment inequalities, which allow me to difference out these future utility streams.

For the following exposition, I assume that individuals have perfect foresight and make adoption decisions independently; I will then loosen these assumptions. Then i forecasts utility correctly and will adopt at the time $\tau_i = \arg\max_\tau U_i^\tau$. If time were modeled as continuous, the optimum would be obtained from the first order condition $\left.\frac{\partial U_i^\tau}{\partial \tau}\right|_{\tau_i} = 0$. I compute the discrete time analogue using differences.

I observe each individual's month of adoption, τ_i, and consider the utility he would have received had he adopted a different month. At time τ_i, i faced the decision of buying a handset and obtaining utility $U_i^{\tau_i}$, or postponing adoption by K months for utility $U_i^{\tau_i+K}$. Since he adopted at τ_i, revealed preference implies $U_i^{\tau_i} \geq U_i^{\tau_i+K}$. The utility of being on the network during the following K months must have exceeded the value of the drop in handset prices:[42]

$$\sum_{k=0}^{K-1} \delta^k u_{i\tau_i+k}(p_{\tau_i+k}, \boldsymbol{\phi}_{\tau_i+k}) \geq \beta^{handset}(p_{\tau_i}^{handset} - \delta^K p_{\tau_i+K}^{handset})$$

Similarly, i could have purchased a handset K months earlier. At time $\tau_i - K$, i chose to postpone adoption to obtain expected utility $U_i^{\tau_i}$ instead of buying a handset and getting utility $U_i^{\tau_i-K}$. This implies $U_i^{\tau_i} \geq U_i^{\tau_i-K}$. Because i chose to postpone adoption by K

[40] Note that both integrals are nonlinear functions of estimated parameters, so uncertainty in parameter estimates could bias the estimates of these expectations. One way to account for this error is to integrate also over the estimated sampling distribution; I am working on calculations with this correction.

[41] In the adoption decision, I assume that individuals forecasts of these usage utilities are correct on average. For a graph of the quantiles of these utilities including all policy changes, see Appendix H.

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

months at $\tau_i - K$, the utility from the K months prior to purchase must have been worth less than the drop in handset prices:

$$\sum_{k=1}^{K} \delta^{K-k} u_{i\tau_i-k}(p_{\tau_i-k}, \phi_{\tau_i-k}) \leq \beta^{handset}(p^{handset}_{\tau_i-K} - \delta^K p^{handset}_{\tau_i})$$

In the exposition I assumed that individuals made independent adoption decisions, but individuals may actually coordinate adoption: for example, a husband and wife may buy phones at the same time. If two individuals tightly coordinate adoption to purchase a handset in the same period, the bounds I derive are simply wider than the bounds that would be obtained if the coordination pattern were accounted for.[43]

These results follow if individuals have perfect foresight over the future path of prices, their own and their contacts' coverage, and their contacts' adoption dates. In reality, individuals are likely to have uncertainty about future paths of utility. The model is robust to a particular deviation from perfect foresight: if individuals hold beliefs over future paths of utility that differ from the true expectation by an error that is mean zero across individuals, this forecast error will be absorbed into η_i.

I form these two relations into moment inequalities (Chernozhukov et al., 2007; Pakes, 2010) to estimate the sensitivity to handset prices $\beta^{handset}$:

$$\begin{aligned} E\left[Z_{mi}\left(\sum_{k=0}^{K-1}\delta^k\left(\sum_{j\in G_i\cap S_{\tau+k}} Eu_{ij\tau+k} + w\cdot Eu_{ji\tau+k}\right) - \beta^{handset}(p^h_\tau - \delta^K p^h_{\tau+K}) + \frac{1-\delta^K}{1-\delta}\eta_i\right)\right] &\geq 0 \\ E\left[Z_{mi}\left(\sum_{k=1}^{K}\delta^{K-k}\left(\sum_{j\in G_i\cap S_{\tau-k}} Eu_{ij\tau-k} + w\cdot Eu_{ji\tau-k}\right) - \beta^{handset}(p^h_{\tau-K} - \delta^K p^h_{\tau}) + \frac{1-\delta^K}{1-\delta}\eta_i\right)\right] &\leq 0 \end{aligned}$$

for a set of instruments Z. I include $Z_{0i} = 1$, based on the restriction $E[\eta_i] = 0$. I also 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). The assumption required for these instruments is that they impact the adoption decision but are orthogonal to the unobserved benefit of being on the network η_i: $E[\eta_i|Z_i] = 0$. 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

[43]The bounds I estimate hold fixed j's decision in computing i's utility from adopting at different times. Consider if i and j coordinate to always adopt in the same month. If i adopts earlier, j will adopt earlier, and the true utility of adopting earlier is greater than I estimate. The true inequality will then imply the inequality I measure: I estimate a looser bound. If i adopts later, j will adopt later, but the utility starting when they adopt will continue to be the same as if they did not coordinate.

communication network and the quality of handset model purchased (see Appendix B). I include the instruments, squared terms, and interactions.[44]

To balance precision with smoothing, I select $K = 2$ months. 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. Given this choice of K, I estimate $\beta^{handset}$ using changes in utility and prices during the two months prior to and two months following adoption. In the two months leading up to adoption, the median consumer gains 3 contacts and the price of a handset declines by \$0.90. The median consumer has 37 contacts when they adopt. In the two month following adoption, the median gains 3 more contacts and the price of a handset declines by \$0.94.[45]

During months extra fees were charged, I incorporate the fee schedule.[46] I fix the discount factor $\delta = 0.9916 \sim (0.9)^{1/12}$. For computation details see Appendix J.

Results. Parameter estimates are reported in Table 4.

The estimate for the cost-dependent censoring parameter α is above zero, suggesting that the cost of placing a call affects the extensive margin of whether to call in a given month.

The coverage sensitivity estimates suggest that the most important term in the hassle cost is the interaction of sender and receiver's coverage ($\beta_{coverage.interaction}$), followed by the receiver's coverage ($\beta_{coverage.to}$). This is consistent with it being more of a hassle to place a call to an area with poor coverage (one must attempt many times in order to reach the receiver) than from an area with poor coverage (one must simply walk to an area with good coverage), and for it being a large hassle to call when both parties have poor coverage.

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

[44]While in theory $\beta^{handset}$ could be set identified, given computational constraints in the simulation, I would be unable to bound equilibria given a set estimate. Including these instruments allows me to recover a point estimate.

[45]I 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, if I were to include subsidy recipients in estimating $\beta^{handset}$, I would obtain extremely wide bounds. Instead, I estimate $\beta^{handset}$ on all individuals who did not receive a subsidy and who subscribed after the first 2 months of the data. This relies on the assumption that the parameter $\beta^{handset}$ is the same for subsidy recipients and nonrecipients. As I show in Section 8, subsidy recipients appear similar to nonrecipients in terms of phone usage and social network attributes.

[46]Before April 2005 there was a monthly fee. Before June 2007, subscribers had to top up their balance with a minimum of \$4.53 every 30 days to keep their account active. If the subscriber incurred charges less than this amount, the leftover balance would accumulate. This would be a binding restriction for most subscribers, as most spend less than this amount. I model the hassle of having to accumulate balance in this way as an extra fee of half the extra top up amount, based on the average duration talked with the contacts on the network during that month, billed at the average basket of rates (50% peak, 46% off-peak and 4% discount). Ideally, I'd use the optimal durations computed for that month; this is a simplification for computational reasons. I do not explicitly model another common way to respond to the restriction, which is to cycle in and out of account activity. The hassle cost I model here accounts for this cycling in a simplified way.

estimates resulting from these two methods can be compared at the point of adoption. In the adoption decision, I find that if the median consumer had adopted two months earlier he would have paid \$0.90 more for a handset, suggesting the expected benefit of being on the network those two months prior to adoption was less than \$0.90 (or about half a day's consumption, given an average nominal daily consumption of \$1.67 for individuals in phone owning households in 2010). Had he adopted two months later than he did, he could have saved \$0.94 on purchasing a handset, suggesting the expected benefit of the two months following adoption was more than \$0.94. Based on the estimates of the call utility model, I find that if the median consumer had adopted two months earlier, he would have obtained an additional \$0.64 in expected call utility; had he delayed adoption by two months, he would have given up \$0.87 in expected call utility. (These estimates assume that recipients do not value incoming calls ($w = 0$); if recipients value incoming calls as much as outgoing calls ($w = 1$), the call utility roughly double counts the surplus from calls. I proceed assuming $w = 0$.[47]) That the call model appears to slightly underweight utility could arise from the omission of SMS and missed calls; as a rough check, if I inflate the call utility estimates by the ratio of revenue from both SMS and calls to revenue from just calls (SMS represents roughly 13% of revenue), they fit within the bounds suggested by the adoption decision: \$0.73 < \$0.90 and \$1.00 > \$0.94.

From these two decisions I estimate one parameter for price sensitivity for calls ($\beta_{call} = 0.20$) and one parameter for the price sensitivity in the adoption decision ($\beta^{handset}$). While in theory handset price sensitivity could be set identified, I obtain a point estimate, $\beta^{handset} = 0.14$. The estimate is lower than the estimate from the calling decision, consistent with the reduced form comparison of tradeoffs around the adoption date, and with the interpretation of the utility captured by the call model representing a proxy for the utility from all communication across a link. The parameter $\beta^{handset}$ scales up the utility of usage uniformly to account for the value of the transactions I have not explicitly modeled: SMS, missed calls, and foreign calls.[48]

The parameters of the communication graph are shown in the second panel of Table 4. Although I estimate many parameters (127.6 million), the number of observations is large (4 billion). Most parameters are estimated off of a sufficient number of observations for standard errors to be reasonable in size.[49]

I interpret the parameters of the call model by presenting comparative statics in Table 5. I set the shock variance σ_i and cost-independent censoring parameter q_i to their medians,

[47] When incoming calls are valued the same as outgoing calls ($w = 1$), I find the expected call utility of the two months prior to adoption is \$1.68 and two months following is \$2.24.

[48] This scaling relies on the assumption that the utility from nonvoice communication represents the same fraction of total utility on each link. The correlation between a link's call duration and call attempts is 0.58 and the correlation between a node's total duration and total SMS is 0.53 (I cannot compute the correlation for SMS at a link level because I cannot pair sender and receiver).

[49] For a short discussion of the incidental parameter problem, see footnote 39.

and show expected outcomes for the 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 cost. It is informative to interpret a few metrics under these conditions for the median link (the center column). Calls across the median link are infrequent: the probability of making in a given month is 0.22. Durations are short: conditional on making a call, the expected duration for that month is 43.2 seconds. The expected monthly cost of communicating across the median link is \$0.04, which corresponds to 0.08% of the average monthly per capita consumption in a phone owning household in 2010, and the link provides an expected utility of \$0.07. 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 half: 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.

Model Fit. The call model has two goals: to uncover from observed durations and costs the underlying conditional distributions, and to translate these durations and costs into utilities. Since the data cannot directly distinguish between the shape of the utility function and the distribution of shocks, I have narrowed the choice of utility function using theoretical restrictions and then selected a distribution of shocks that matches the data well.

One test of this separation is the fit of the calling distribution. The fit of the duration distribution is shown in Figure 4. The fit is helped by the large number of parameters estimated, but the choice of functional form is still important, as is evident from the predicted distribution's slight systematic deviations from the data.

7. Simulation of Network Good Adoption

In this section I outline 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.[50]

Simulation Method. There is some initial set of adopters S_0 whose decisions are unaffected by the change in the environment. Other individuals decide on an adoption month $\tau_i \in \left[1, ..., \bar{T}\right]$. For simplicity in exposition, I assume that each adoption decision is made independently, without coordination; I discuss robustness to coordination later. Because

[50]For simulation results to be exact, the strict assumption is that the counterfactual utility that would be provided to nodes adopting after May 2009 in the observed data never exceeds that provided in the actual data. If it does, the results will underestimate adoption.

counterfactuals may induce individuals to delay adoption, I set the end date for the simulation three years beyond the limits of the calling data, using aggregate adoption statistics to scale utility to account for expansion in the network after my data ends.[51]

I define an equilibrium as a function of a vector of types $\boldsymbol{\eta} = [\eta_i]$, where an individual's type represents their idiosyncratic benefit of being on the network.

An equilibrium $\Gamma(\boldsymbol{\eta})$ is defined by adoption times $\boldsymbol{\tau} = [\tau_i]_{i \in S}$ satisfying:

(1) Fixed initial adopters: the adoption date for an initial adopter $i \in S_0$ is $\tau_i = 0$
(2) Individual rationality: the adoption date of each other individual $i \in S \backslash S_0$ is optimal given type η_i and others' adoption dates:

$$\tau_i = \arg\max_t U_i^t(\eta_i, \tau_{-i})$$

To identify an equilibrium given $\boldsymbol{\eta}$, I use an iterated best response algorithm:

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

$$\tau_i^1 = \arg\max_t \delta^t U_i^t(\eta_i, \boldsymbol{\tau}_{-i}^0)$$

(3) Iterate, using the path from the previous step $\boldsymbol{\tau}^k$ to form the next:

$$\tau_i^{k+1} = \arg\max_t \delta^t U_i^t(\eta_i, \boldsymbol{\tau}_{-i}^k)$$

(4) Stop when the equilibrium converges: $\tau_i^{k+1} = \tau_i^k$ for all i[52]

The equilibrium identified depends on the candidate adoption path $\boldsymbol{\tau}^0$ as well as the vector of types $\boldsymbol{\eta}$.

For baseline simulations, I set the candidate adoption path to the observed adoption path so that at the first step of the algorithm individuals expect the observed equilibrium. By doing this I am likely to recover the equilibrium closest to the observed path.[53] The observed path is also my best guess of individuals' expectations for adoption for counterfactuals that

[51]See Appendix I 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.)

[52]With the aim of speeding convergence, in practice at each step k I use the path defined by τ_j^k for individuals j that have reoptimized in this step and τ_j^{k-1} for individuals who have not yet reoptimized in this step, in the same manner as the Gauss-Seidel method. 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.

[53]For counterfactual simulations, to speed up computation I set the candidate adoption path to the corresponding baseline equilibrium adoption path.

do not greatly shift adoption. I do not attempt to recover all possible equilibria,[54] but since there is uncertainty in $\boldsymbol{\eta}$ will recover a set of equilibria.

Ideally, I would compute a sample of the equilibria arising from many draws from a distribution of $\boldsymbol{\eta}$, but doing so is computationally prohibitive because demand is interlinked. In other settings, each η_i could be sampled independently; however, because adoption decisions are interlinked, i's particular draw of η_i can potentially affect the decisions of the entire network. I could sample the over one million η_i's jointly and compute the resulting equilibria, but it would be computationally prohibitive to sample a sufficient number of these draws.[55] Instead, I back out bounds for each individual's realized type based on the adoption inequalities and the estimate of $\beta^{handset}$. I find $\underline{\eta}_i \leq \eta_i \leq \bar{\eta}_i$, where:[56]

$$\underline{\eta} = \quad -\frac{1-\delta}{1-\delta^K}\left[\sum_{k=0}^{K-1}\delta^k E\left(\sum_{j\in G_i\cap S_{\tau_i+k}} u_{ij\tau_i+k} + w\cdot Eu_{ji\tau+k}\right) - \beta^{handset}(p^h_\tau - \delta^K p^h_{\tau+K})\right]$$

$$\bar{\eta}_i = \quad -\frac{1-\delta}{1-\delta^K}\left[\sum_{k=1}^{K}\delta^{K-k} E\left(\sum_{j\in G_i\cap S_{\tau_i-k}} u_{ij\tau_i-k} + w\cdot Eu_{ji\tau-k}\right) - \beta^{handset}(p^h_{\tau-K} - \delta^K p^h_\tau)\right]$$

Because each individual's type is backed out as a set rather than a point, the set of types $\{\boldsymbol{\eta}|\underline{\eta}_i \leq \eta_i \leq \bar{\eta}_i\}$ may trace out a set of equilibria rather than a single equilibrium. I derive bounds for this 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 τ_i: a higher type η_i weakly decreases i's optimal adoption date. Second, note that the underlying game has strategic complements: a decrease in i's adoption date τ_i weakly decreases j's optimal adoption date.[57] Thus the equilibrium $\Gamma(\underline{\boldsymbol{\eta}})$, where each individual's η_i is set to its lower bound, represents the lower bound of the identified set of equilibria, and the corresponding equilibrium $\Gamma(\bar{\boldsymbol{\eta}})$, where each individual's η_i is set to its upper bound, represents the upper bound. These filter through to provide bounds on the adoption date for each individual, $[\underline{\tau}_i, \bar{\tau}_i]$ within the identified set of equilibria. I compute a third equilibrium $\Gamma\left(\frac{\underline{\eta}_i+\bar{\eta}_i}{2}\right)$, which

[54]Because the game is supermodular it is possible to recover bounds on the set of equilibria by starting the algorithm with the lowest candidate adoption path (all individuals delay adoption until the end of the data) and the highest candidate adoption path (all individuals adopt in the initial period).

[55]Given 12 hours to compute one equilibrium, it would take one instance one year to compute the equilibria resulting from roughly 730 values of $\boldsymbol{\eta}$. Multiple instances could be run simultaneously, but each requires roughly 115 GB of memory. This form of sampling would eventually be feasible given sufficient access to a powerful enough computing cluster; in a future version of this paper I hope to describe various empirical approaches to fit various computational resources.

[56]I back out these bounds differently for subsidy recipients; see Appendix I for details. Also note that the error structure cannot rationalize all adoption decisions: there are some observed decisions (roughly 20%) for which the inequalities of η_i cross. In these cases, I assume η_i is the mean of the two bounds. A second error could be added to rationalize these decisions, but this would have to be a random effect from a distribution. This would make simulation intractable: since adoption decisions are interlinked, individual i's draw affects the adoption decisions of the rest of the network.

[57]This follows from the lattice structure of $\boldsymbol{\tau}$ and because $U^{\tau_i}(\eta_i, \boldsymbol{\tau}_{-i})$ has increasing differences in τ_i and τ_j / is supermodular in $\boldsymbol{\tau}$; see Topkis (1978) and Milgrom and Shannon (1994).

is my best guess of the equilibrium that would be observed, by setting $\eta_i = \frac{\underline{\eta}_i+\bar{\eta}_i}{2}$, the mean of the low and high bounds for each individual.[58]

The state space is large: there are $\bar{T}^{|S\setminus S_0|}$ or on the order of $89^{1{,}000{,}000}$ possible states, but the algorithm identifies an equilibrium in about 12 hours.[59]

Revenue and Utility. For each equilibrium I compute the net present value of revenue and utility, as of January 2005. The revenue from equilibrium Γ is computed by summing the price times the expected duration across each link:

$$R^{\Gamma} = \sum_{i \in S} \sum_{t \geq \tau_i} \delta^t p_t \cdot \sum_{j \in G_i \cap S_t} E d_{ijt}(p_t, \phi_{it}, \phi_{jt})$$

Total utility from calls is computed analogously:

$$U^{\Gamma}_{calls} = \sum_{i \in S} \sum_{t \geq \tau_i} \delta^t \sum_{j \in G_i \cap S_t} E u_{ijt}(p_t, \phi_{it}, \phi_{jt}) + w \cdot E u_{jit}(p_t, \phi_{jt}, \phi_{it})$$

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

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 time τ_i and then sells it back at the end of the data at the prevailing price. This yields the following cost of handset ownership:

$$C^{\Gamma}_{handsets} = \sum_{i \in S} \left[\delta^{\tau_i} p^h_{i\tau_i} - \delta^{\bar{T}^{data}} p^h_{i\bar{T}^{data}} \right]$$

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

$$U^{\Gamma}_{net} = \frac{1}{\beta^{handset}} U^{\Gamma}_{calls} - C^{\Gamma}_{handsets}$$

where I convert the utility from calling into dollars using the handset price sensitivity $\beta^{handset}$, which will measure the total communication utility. In welfare calculations I omit the idiosyncratic benefit term 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. I also omit any potential profits earned in the handset market.

[58] I handle the simulation of adoption subsidy recipients slightly differently from nonrecipients due to the sharp discontinuity in cost generated by the subsidy; see Appendix I for more detail.

[59] Several factors contribute to the algorithm's performance. 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. There are thick regions of indifference: changes in the network only affect i if they affect i's direct neighbors.

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

$$\begin{aligned} R^{\Gamma(\underline{\boldsymbol{\eta}})} \leq \quad & R \leq \quad R^{\Gamma(\bar{\boldsymbol{\eta}})} \\ U_{calls}^{\Gamma(\underline{\boldsymbol{\eta}})} \leq \quad & U_{calls} \leq \quad U_{calls}^{\Gamma(\bar{\boldsymbol{\eta}})} \end{aligned}$$

Because the net utility function omits idiosyncratic benefits, it does not match the utility each individual maximizes, and there may be an equilibrium between $\Gamma(\underline{\boldsymbol{\eta}})$ and $\Gamma(\bar{\boldsymbol{\eta}})$ that has a net utility lying outside the bounds of $U_{net}^{\Gamma(\underline{\boldsymbol{\eta}})}$ and $U_{net}^{\Gamma(\bar{\boldsymbol{\eta}})}$.

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 5, the simulation matches the general trend of the data. While adoption in the data grows more continuously, the adoption path generated by the model has more discrete jumps, resulting from individuals settling on adoption dates at price changes. These jumps would be softened if either there was some uncertainty about the future, if subscribers faced different handset prices, or if subscribers had heterogenous price sensitivities. Under mean shocks the correlation between observed adoption month and simulated adoption month is 0.87, and the mean deviation is 2.82 months. In the simulated equilibrium, I estimate the net present value of revenue to lie between \$215m and \$235m (an average of \$10-11 per subscriber per month, or 3-7% of household consumption[60]). This is comparable with statistics from the operator's annual reports: the operator reports that the 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;[61] these numbers suggest a total discounted revenue for the period of \$302m. This total revenue will include revenue from SMS, international calls, and special plans (including corporate lines and mobile payphones), so should be larger than the model's estimate.

I estimate the total net present value of utility from calls U_{calls} to be between \$75-91m (an average of \$3-4 per subscriber per month, or 1.0-2.4% of household consumption), net of calling and hassle costs. I estimate the cost of handset ownership to be between \$21-26m (an average of \$1 per subscriber per month, or 0.3-0.6% of household consumption), resulting in net utility U_{net} between \$54-65m (an average of \$2-3 per subscriber per month, or 0.6-1.8% of household consumption).

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 a change in the environment. The most natural measure of impact would

[60]The large spread arises because the distribution of subscriber income changes over time as poorer households subscribe. The average nominal monthly consumption in households with mobile phones was \$472.59 in 2005 and \$249.71 in 2010. The average subscribing household in 2010 had 1.52 mobile phones. (See Table 1.)

[61]The operator reports an ARPU of \$19 in 2005, \$17 in 2006, \$12 in 2007, \$11 in 2008, and \$7 in 2009.

be bounds on the change in revenue and utility; however, this measure is computationally prohibitive for the same reasons it is difficult to compute equilibria for a sample of types $\boldsymbol{\eta}$—adoption decisions are interlinked. Even given the $\eta_i' s$ restricted to lie within the realized bounds, the set of potential type vectors is large: $[\underline{\eta}_0, \bar{\eta}_0] \times [\underline{\eta}_1, \bar{\eta}_1] \times \cdots \times [\underline{\eta}_N, \bar{\eta}_N]$, and it would be computationally prohibitive to sample a sufficient number of draws.

Instead, I measure policy impacts by reporting changes in the bounds on revenue and consumer surplus. It is important to note that these do not necessarily represent bounds on the potential changes. In particular, the change in the upper bound equilibrium may be less than the change in the lower bound equilibrium. When a policy change causes the lower and upper bounds to shift by similar amounts, I may report one number to describe the approximate shift. When 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.

Robustness.

Coordinated adoption. As described in the estimation section, if two individuals tightly coordinate adoption the estimated bounds I obtain are simply wider than would be obtained if the coordination pattern were accounted for. The simulation method will faithfully represent coordinated decisions present in the data; for counterfactuals it will discover coordinated changes if the coordinated outcome could be obtained by successive individually rational deviations.

Handset sharing. In the presence of handset sharing, this model would attribute the surplus utility from shared calls to the account owner. It would differ in two respects from a model that incorporated handset sharing: there are no sharing costs, and the call shock distributions are assumed to be independent.[62] It would be possible to make the model more flexible, but since I do not observe sharing patterns this would require assumptions on the set of individuals who obtain utility, how utility is divided between them, and the costs of sharing.

Utility from incoming calls. The results I present here assume that the surplus from a call accrues to the caller ($w = 0$); as a robustness check I also compute results under the assumption that both caller and receiver obtain the same surplus ($w = 1$). These results are similar to the baseline results; see Appendix K for details.

[62] I have assumed call shocks are independently and identically distributed, but if a link represents shared use of the phone, the distribution of calls on that link would change if the person sharing the phone purchased a handset.

Homophily: correlated traits or shocks. In many social network analyses, it is difficult to disentangle the effects from peers across social network links from nonpeer effects that are correlated among network neighbors. A translation of this concern into this setting would be that correlated shocks or attributes among neighbors might bias estimation of the network effect. I get around this problem by exploiting the fact that I observe the actual behavior of interest across links: the communication I observe represents the utility derived from the network.[63]

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 grandmother a mobile phone). Detailed knowledge of the local network structure makes it possible to overcome local inefficiencies. 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 result in improved efficiency (Katz and Shapiro, 1994). These programs are common: for example, Facebook currently subsidizes data usage in developing countries to boost adoption (BBC, 2010). However, global actors are constrained by information. While selecting an optimal policy would 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 eventually 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.

[63]Correlated traits or shocks could affect either the call model or the adoption model. If connected individuals call similar amounts in the cross section, the parameters of the call shock distributions I estimate would faithfully represent these correlations. If connected individuals call in response to correlated shocks in time, the estimates will average these out over time. Because I use plausibly exogenous variation affecting the timing of adoption to identify the handset price sensitivity parameter $\beta^{handset}$, correlated traits or shocks that affect adoption would not bias this estimate and instead be absorbed into the idiosyncratic benefit η_i.

2008 Adoption Subsidy Program. The Rwanda Utilities Regulatory Authority collects 2% of all operator revenues into a Universal Access Fund, to be used for initiatives to accelerate the use of Information and Communication Technologies. One of the initiatives implemented using revenue from this fund was a targeted adoption subsidy program in 2008. The government purchased 53,352 handsets (amounting to roughly 8% of the country's stock of handsets at the time) and distributed them to individuals through local governments at a reduced price.

Fifteen of 30 districts participated in the program. Generally, individuals came to the district office to voice interest in the program, and the local government allocated handsets to interested parties. Each district handled its own distribution and thus the allocation methods differed by district.

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.

The full price of the handsets was $28. Beneficiaries were to pay a fraction of this price 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.

I evaluate the effects of this subsidy program on network adoption. I begin with descriptive evidence about the allocation and use of the subsidized handsets.

Allocation of Subsidized Handsets. The characteristics of the districts that were allocated handsets are shown in Table 6. Handsets were generally allocated to rural districts with low baseline mobile phone adoption. Allocations varied significantly: half of the districts were allocated no handsets; those allocated handsets received enough for between 1% and 15% of households.

As a first step, I analyze the impact of the program using national household survey data collected by the government in 2005 and 2010. I aggregate to the district level. A regression of the change in number of households owning phones on the number of handsets allocated provides correlational evidence of the impact of the program.[64] For a good with no network or learning effects, I would expect one good allocated to result in weakly less than one good owned in a follow up survey, where the decline is due to depreciation, exchange across borders, and the fact that some decisions are inframarginal. If it spurs network effects, the allocation of one good would cause others to adopt, and could result in more than one good owned in a follow up survey.

[64]The earlier survey does not ask about the number of handsets owned within a household, so I only look at the fraction of households owning at least one handset. If the subsidized handsets were distributed to households with existing handsets, this would underestimate the association.

Results are presented in Table 7. Allocating handsets to an additional percentage point of households in a district in 2008 is associated with an increase of adoption between 2.00 and 3.39 percentage points between 2005 and 2010. Districts allocated handsets had lower initial levels of adoption, and there is evidence that adoption grew more slowly in areas with lower initial adoption: controlling for the 2005 number of households with phones, or restricting to districts that received some handsets from the program tends to raise estimates.

These results are suggestive of a moderate impact on broader network adoption, but are descriptive, not causal. Allocations were not random and may have been targeted towards districts that otherwise would have differential adoption trends. Another issue is that network effects need not remain constrained within a district; spillovers across district borders would bias estimates downward. In the next subsection I analyze usage data, which suggests that this is the case: although all of the handsets were allocated to rural areas, many of the handsets were used in urban centers.

Use of Subsidized Handsets. We can learn more about the impact of the program by analyzing subsequent usage data from phone records.

Subsidized handsets are identifiable in the phone data. Figure 6 shows the number of activations of the particular model of handset distributed under the subsidy: it was not a common model, and the subsidy program represents a large spike. Most studies of distribution programs will observe the initial recipient, but in the phone data I observe the ultimate recipient of the handset.[65]

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 handsets than were allocated per government records could arise from subsidized handsets being activated later than April 2008,[66] allocated to the competing operator, or not being used at all.

Handsets appear to be used either where allocated or in urban areas. Figure 7 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). This latter point is noteworthy, because no handsets were allocated to these areas.

To understand further, Figure 8 shows the inferred locations of all accounts, and of accounts affiliated with subsidized handsets. The spatial distribution of subsidized accounts is very similar to that of initial activations, suggesting that these are also the locations

[65]If 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.

[66]I 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.

where handsets were ultimately used. These results suggest that the program's effects are not confined to beneficiary districts, so that the reduced form approach taken in the earlier section may not yield a valid estimate of its impacts.

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 purchased phones around the same time, in terms of calls, durations, and total number of contacts (see Table 8).

Recipients' network structure is similar to others who subscribed around the same time. An optimal subsidy program to overcome dispersed network externalities would target individuals who provide benefits to others who have yet to subscribe, who would not subscribe in absence of the target's adoption.[67] For mobile phones, the most direct benefit results from phone calls, which can be measured based on usage. I compute one metric of these benefits: the eventual duration spoken with contacts that have yet to subscribe. I also compute the clustering coefficient (the fraction of a node's neighbors who are themselves connected). By all these metrics, recipients' network structure looks similar to others who subscribed around the same time.

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 structure of the network of benefit flows. This is difficult to analyze in reduced form. In the next section, I use the simulation method to evaluate the impact of the program.

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

- *Subsidy recipients represent the full set of eligible individuals.* 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.* This would hold if the subsidy were unanticipated, or if recipients anticipated the subsidy but expected it to be less generous that it was.
- *Recipients preferred taking the subsidy at the point of adoption to purchasing any time in the following 4 years.*

For more details on these last two assumptions, see Appendix I.

[67] Of course, in a dynamic setting, individuals will anticipate others' adoption, so an anticipated subsidy could impact adoption prior to the subsidy.

Under these assumptions I can compute the effect of the subsidy. Results are shown in Table 9. 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").[68] 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 different parts of the network: the subsidy recipients, the contacts of recipients, and nodes that are not connected to recipients.

The bounds I obtain are wide because it is difficult to know when a subsidy recipient would have adopted in absence of the subsidy. The upper bound presents an optimistic scenario: targeted individuals would have delayed adoption by an average of only 2.05 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 2.12 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, 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. The net present cost of the subsidy was $569,741, but it shifted the bounds on welfare upward by $5,628,126 (lower equilibrium) and $760,849 (upper equilibrium), resulting in an increase in the bounds on net welfare of $5,058,385 (lower equilibrium) and $191,108 (upper equilibrium).[69]

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 should shift upward by $2,110,828 in the lower equilibrium, but downward by $442,372 in the upper equilibrium.

Most of the effect is a proximal effect of the subsidy. Ripple effects account for 33% (lower equilibrium) and 27% (upper equilibrium) of the effect on revenue and 30% (lower) and 5% (upper) of the effect on consumer surplus.

[68]It 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 $\underline{\eta}_i$ is backed out for subsidized nodes. See Appendix I for a discussion.

[69]These 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 would be subtracted from the welfare gains.

The subsidy provides substantial benefits to the contacts of recipients:

- **A significant fraction of calling benefits accrued to contacts of subsidized nodes.** Recipients' utility increased by $1,254,189 (lower) or $557,459 (upper), from the combination of increased calling and the direct value of the discount. Contacts of recipients received utility only from increased calling, but obtained 53% of all benefits in the lower equilibrium and 11% in the upper equilibrium.
- **More than 62% of the increase in revenue comes from contacts of recipients.**

Most spillovers accrue to the contacts of subsidy recipients, which is sensible given that they receive a direct benefit from being able to call the subsidy recipient.

In Appendix K as a robustness check I analyze the program under the assumption that the surplus from calls is evenly split between caller and receiver ($w = 1$); results are qualitatively similar but a bit smaller: the bounds on welfare increase by $1,135,650 (lower bound) and $9,599 (upper bound).

Overall, the impact of the subsidy on network adoption is consistent with what might have been expected from the reduced form 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 not as targeting individuals, but rather as targeting neighborhoods of the graph. In the next section I outline how the simulation method can be used to improve the targeting of subsidies for future network goods by exploiting network structure.

Improving Mobile Internet Adoption. While failures to internalize network effects have not prevented widespread adoption of mobile phones, they may prevent widespread adoption of affiliated services such as mobile money and mobile internet.

If a government or operator knew how the benefits from adoption of these goods would be distributed across the network, it could target localized inefficiencies and improve welfare. But while a policymaker would most like to know the benefits to adoption in 'dark' regions of the network that have yet to adopt, benefits are revealed only for regions that have adopted.

I propose a method to 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. Here I outline how mobile phone use can inform mobile internet policy.

Mobile phones are likely to be the most convenient modality to deliver internet service in Africa.[70] This suggests that the model of adoption developed in this paper could be

[70] In many poor countries, traditional computers are rare, apart from in centralized institutions such as schools, offices and internet cafes. In Rwanda, only 2% of households own a working computer; and only 0.7% own a computer with an internet connection, while 19% of individuals aged 15 and over own a mobile phone capable of browsing the internet. Mobile phones are a primary channel for accessing the internet: of individuals using the internet in last 12 months, 71% have used it from a mobile phone, compared with

extended to internet access. Since future mobile internet adopters are likely to be current mobile phone subscribers, features of their calling network will be revealed by mobile phone use. A large component of internet use is social—e-mail, forums, or social networking between subscribers—suggesting that social network measures derived from mobile phone usage may be predictive of the utility derived from internet access.[71]

In Appendix L, I provide a brief outline of a procedure to predict the benefits for mobile internet usage. I first outline a model of mobile internet adoption. I then describe how to estimate a mapping between network properties observed in call data and mobile internet usage, using either surveys or the combined internet and calling behavior of early adopters. This mapping can then be used to predict mobile internet usage for regions of the network that have yet to adopt. The simulation method developed in this paper can be used on this predicted benefit network to evaluate policies to encourage adoption.

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 is still an active question for basic voice services: while currently 90% of the world's population has mobile phone coverage and expansion is expected to continue, it is expected that 2-5% of the world's population will not be profitable to serve by the private sector (GSMA, 2006). It is also important for newer services that use the mobile network such as mobile internet: only 45% of the world's population has mobile broadband coverage (3G) (ITU, 2011b). A wide variety of policy instruments are currently in use to encourage rural service provision, from tax-and-transfer schemes like those used to support rural telephony in the U.S., to service obligations, to 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 social benefits to

52% at work, 50% at an internet cafe, and 31% at a place of education. Mobile phones are cheaper than traditional computers and can be powered by batteries, which is especially important in a context where few households have electricity. (RIA, 2012)

[71]Uses of the web can broadly be divided into accessing content and interacting with other subscribers through social services such as e-mail, social networking, and forums. The latter social uses are important: Facebook is the most popular website in Rwanda (Alexa, 2013), and among Rwandans aged 15 and over who use the internet, 88% have signed up for a social network and 97% have an email address.

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 determine the full 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 aggregates such as population density.

Background. Rural areas in this context are less lucrative due to lower demand (incomes and population densities are lower) and higher costs (infrastructure is lower quality[72]). I focus on the case that service is provided by a monopolist operator, which is the case in 11% of countries (GSMA, 2009), but the logic follows similarly for provision under oligopoly. Under certain conditions it is socially optimal to mandate that a regulated monopolist provide service to remote areas. Assume there is a monopolist operator who currently serves an urban market. It has the option of expanding into a rural area that is more price sensitive, at an expansion cost F. This decision is illustrated in Figure 9. If the operator is allowed to set separate prices, it can treat the rural market independently, and weigh the profits resulting from a profit-maximizing price against the expansion cost, as shown in panel (a) of Figure 9. However, many regulators forbid operators from pricing in a discriminatory manner. If instead the operator is required to offer a uniform price, its decision to expand will also be affected by the urban market. It will weigh the additional revenue from the rural market against the loss in revenue from setting a lower price in the urban market, as shown in panel (b) of Figure 9, and is less likely to find expansion profitable. Even in cases where pricing is not restricted by regulation, a monopolist will not internalize all of the surplus it generates. Thus for some demand curves and expansion costs, it would be socially optimal to mandate the provision of service.

Building a tower costs approximately \$130,000. The Rwandan regulator estimates the total annualized cost of owning and operating a tower as \$51,000 per year, plus \$29,584 for towers that are far from the electric grid and must be powered by generators.[73] Rural towers also tend to generate less revenue; mean monthly revenue from an urban tower is nearly twice that of an rural tower.

In the absence of regulation, an operator would build out towers to the point where any set of marginal towers would not be profitable. Coverage regulation can induce an operator

[72]Building a remote tower may require building an access road, and if it is far from the electric grid, operating a diesel generator.

[73]As part of developing infrastructure sharing guidelines in 2011, the Rwandan Utilities Regulatory Agency analyzed the costs associated with tower construction and operation after requesting financial data from operators (RURA, 2011). I use their figures of total cost of ownership to operate a tower, which sum 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. Per the recommendation of the operator's engineer, I assume a tower height of 35m.

to build beyond this point. The impact of the regulation depends on the net cost to the operator to build these marginal towers and the resulting benefit to consumers.

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,[74] but it is straightforward to simulate a suggestive counterfactual: a counterfactual where the operator trims back rollout, and does not build marginal, low revenue towers. I compute the effect on revenue and welfare of building and operating the 10 rural towers earning the lowest monthly revenue that were constructed between 2005 and January 2009.[75] These towers represented 3% of total towers, and their building and operation represented a net present cost of $467,186 in 2005. The distribution of monthly revenue by tower is shown in Figure 10; the omitted towers are highlighted.[76] Two of these low revenue towers cover border crossing points, for which there was an explicit coverage requirement.

While historical revenues provide a rough gauge of the revenue generated by a tower, they do not capture the causal impact on revenue: they omit substitution between towers and the effect of coverage on adoption. I determine the causal impact using my simulation method.

I compute the progression of coverage omitting these 10 rural towers: Figure 11 shows the regions that lose coverage in this counterfactual rollout. I then 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, and allowing each individual to reoptimize their adoption decision until an equilibrium is reached. Table 10 presents the results for adoption months, revenue, and consumer surplus. As described in Section 7, for computational reasons I measure the impact of rural expansion by reporting changes in the

[74]Computing an equilibrium under an alternate tower rollout plan requires estimating the resulting coverage maps for each month, computing each individual's new coverage at each month ($1.5m \times 53$ values), computing the resulting utility on each link for each month ($125m \times 53$ values), and simulating equilibrium adoption. This process takes roughly one week for one alternate rollout plan.

[75]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 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. I can only compute the difference in revenue between constructing the rural towers according to the original rollout plan and delaying construction until after the data ends.

[76]A few towers with low revenue were built before 2005. Since adoption at that point had already internalized the presence of these towers, I leave these towers be.

bounds on revenue and consumer surplus, rather than bounds of the changes. I sometimes find that the lower equilibrium shifts more than the upper equilibrium.

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 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 10 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.[77] As would be expected, the expansion affects the adoption of nodes whose coverage was substantially affected more than those minimally affected, moving the former's adoption forward by an average of 0.16 months in the lower equilibrium and 0.07 months in the higher equilibrium, and the latter's adoption forward by 0.02 months in either equilibrium. However, because there are so many more minimally affected nodes, the bulk of the total effect in all dimensions accrues to individuals whose coverage was only minimally affected.

I find:

Rural expansion improved welfare. Building the 10 lowest revenue towers shifts bounds on welfare upward by \$243,032 (lower equilibrium) and \$179,381 (upper equilibrium).

Private benefits were too dispersed for rollout in the absence of intervention:

- **The rollout was unprofitable for the operator.** Building the towers shifted bounds on the operator's profits downward by \$96,659 (lower equilibrium) and \$140,443 (upper equilibrium). In many cases in economics, competition brings provision closer to the social optimum, but due to network effects the effect of competition on rollout would be ambiguous. In a more competitive setting, each operator would own less of the network and may internalize less of the benefits of an expansion, both because the interconnection fees that can be charged for calls connecting between networks are generally regulated to be near cost, and because a fraction of the benefits would ripple into the competitor's network. The benefits generated by the expansion were quite dispersed, suggesting they would likely ripple across networks: over 76% of the increase in revenue generated by the building of the new towers 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

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

> benefits to finance the tower construction themselves: it would reduce bounds on overall consumer surplus by \$127,496 and \$147,362. A related question is whether citizens would be willing to raise local taxes to finance local infrastructure improvements. However, over 65% 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 \$349,460 and \$358,767; this despite generating substantial benefits both for consumers in other locations (increasing their bounds on consumer surplus by \$221,965 and \$211,406) as well as the operator (increasing bounds on profit by \$370,527 and \$326,743).

If I inflate the revenue estimates to account for revenue from SMS,[78] I find that the expansion still reduced bounds on profits, but less—by \$42,542 and \$92,722—and the rollout was more socially beneficial, increasing bounds on welfare by \$297,148 and \$227,103.

In Appendix K, as a robustness check I analyze the expansion under the assumption that the surplus from calls is evenly split between caller and receiver ($w = 1$). I find that the expansion still reduced bounds on profits, by \$65,417 (lower equilibrium) and \$140,948 (upper equilibrium), and the rollout was socially beneficial, increasing bounds on welfare by slightly less: \$216,918 (lower equilibrium) and \$94,364 (upper equilibrium).

These results suggest that the benefits from constructing a mobile phone network are highly dispersed, over different actors as well as across space. I find that in this case, a Rwandan government obligation to provide service to rural areas induced the building of marginal towers that increased bounds on consumer surplus by 0.6% (lower equilibrium) and 0.5% (upper equilibrium). Overall, the impact of the regulation is small and the broader network rollout appears to be driven largely by private incentives. This is consistent with the worldwide expansion of mobile phone networks across diverse settings, as well as Batzilis et al. (2010), which finds that in the presence of coverage obligations, market factors are predictive of the rollout of the mobile phone network in Malawi.

Impact by Population Density. Population density is an important factor in the profitability of providing mobile phone coverage: areas with lower 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. The demand traced out in rural regions of Rwanda can provide insight about the demand in uncovered regions in other countries. In this section I compute the profitability of providing coverage as a function of population density.

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.

[78] SMS represented 13% of revenue.

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, I simply scale down the revenues and consumer surplus, holding fixed the operating costs.

From the previous section I obtain the net present cost C of building and operating the 10 lowest revenue rural towers and the revenue R^{Γ} and consumer surplus U^{Γ}_{net} in equilibrium Γ. If the population density were scaled by a factor λ, the predicted impact on revenue and consumer surplus would be:

$$\Delta\tilde{R}^{\Gamma} = \lambda\Delta R^{\Gamma} - C$$
$$\Delta\tilde{U}^{\Gamma}_{net} = \lambda\Delta U^{\Gamma}_{net}$$

where ΔX^{Γ} is the impact of the tower construction on X in equilibrium Γ.[79]

For high population densities, with $\lambda > 1.43$, it is 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 $\lambda < 0.66$, it is both unprofitable and welfare reducing to expand the network in both bounds, so that a coverage obligation would reduce welfare. However, there is a range $0.72 < \lambda < 1.26$ where expanding the network would be socially optimal but not be profitable, in both bounds.[80] In this range, an intervention to provide coverage would improve welfare. (This range corresponds to scaling the Rwandan average population density to lie between 301 and 525 people per square kilometer.[81])

Results suggest that the coverage obligation led to unprofitable but welfare improving tower construction in Rwanda, 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 location-specific prices) may be able to internalize sufficient value. If there are restrictions on pricing, it may be optimal for governments to encourage service in populated but marginal areas that otherwise would not receive service.

The provision of infrastructure by a private firm is subject to the constraint that the firm still finds it profitable to provide service. Regulators thus face a constraint on the obligations they can impose in service of social welfare. The type of analysis presented in this paper can guide these choices.

---

[79] As an approximation, I scale R and U_{net} linearly with population density. The actual relationship would be more complex: the quantities are a function of the number and weights of the links in the network, not just the number of nodes. It is not obvious how the number of links would grow with the number of nodes (if each node spent a fixed budget on calling it could grow linearly, but if each node increases calling proportionally with the number other nodes, it could grow as the square, for example).

[80] Values of λ not covered by these three cases have different effects on the lower and upper bound equilibria.

[81] This range does not represent a rule of thumb for other contexts, as it is specific to the geography around these 10 particular towers and represents the average density of the country, not the affected area.

The model estimated in Rwanda can also be extended to more richly model incentives to provide service in other settings. For example, the locations of individuals can be redistributed to match aggregate population densities, the links between regions and cities can be predicted using a gravity model, and coverage can be adjusted based on geography.

10. 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 rich new data on the adoption and usage of nearly an entire network of mobile phone users.

I turn this method towards two applications. It can be optimal to subsidize adoption when individuals do not internalize the benefits their adoption provides to the rest of the network. I find that a rural adoption subsidy program costing \$569,741 improved net welfare in a low case by \$191,108, or 0.06%, and in a high case by \$5.6 million, or 2%. I find that a large fraction of its impact accrues to nonrecipients: in particular, contacts of recipients account for more than 62% of the effect on revenue. These spillovers suggest that adoption subsidies for network goods should be thought of not as targeting individual nodes, but as targeting neighborhoods of the graph. Future work can use network structure revealed by mobile phone use to better target adoption subsidies for other goods such as mobile internet.

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, shifting bounds upward by at least \$179,381, or 0.06%. Future work can guide the design of other regulations for network goods.

This paper shows an inward application of the tremendous stock of data generated by mobile phone networks, to understand the economics of mobile phones themselves. There are more questions to answer in this direction, which can answer longstanding academic questions about network goods and information technologies, as well as form a compelling business case for potential data sharing partners.

However, an outward line of research has the potential for broader social impact. Information technologies can act as social sensors, illuminating human behavior that was once unobservable or unquantifiable. These data sets are not just stocks but flows, which open the possibility of performing analyses that guide policymaking in real time, or which underlie new products and services, such as credit scores derived from behavioral signatures in mobile phone usage.

Services that are provided by formal institutions in developed countries, such as insurance and public goods, are often provided in developing countries informally by social networks (Chandrasekhar et al., 2013). The study of these networks can improve our understanding of the process of development. Mobile phone data provides an inexpensive, objective, and rich way to measure social networks, and thus can uncover new insights about how social networks mediate these basic services. As examples of this line of work, Blumenstock et al. (2011) exploits rich measurement on the mobile phone network to determine whether transfers made in response to a natural disaster are part of a reciprocal risk sharing arrangement. In separate, ongoing work, I am investigating how learning propagates over social networks, noting that as an individual learns to use a mobile phone, rich data is recorded on each transaction. I exploit this fact to analyze how a new, discounted mobile phone plan spread through the social network, tracing how individuals learn from their own billing experience and the experiences of their contacts. Future work can layer rich measurement with targeted experiments to answer a broader set of questions.

Privacy remains an ongoing concern with the use of large, passively collected data sets, and is especially important to consider in developing societies that may not have as many safeguards for individuals. Usage traces such as the data in this paper are now being collected passively in many sectors of the economy by private entities, and are already being used for internal purposes. The gathering and analysis of this big data holds risks that are currently quite uncertain (e.g., Narayanan and Shmatikov, 2008; de Montjoye et al., 2013), but tremendous potential for improving the wellbeing of societies. Vigorous public research in both directions is needed, so that societies can mitigate the risks, and absorb the benefits.

References

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.

Alexa (2013): "Top Sites in Rwanda," *Website*.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2012): "The Diffusion of Microfinance," Working Paper 17743, National Bureau of Economic Research.

Batzilis, D., T. Dinkelman, E. Oster, R. Thornton, and D. Zanera (2010): "New cellular networks in Malawi: Correlates of service rollout and network performance," *National Bureau of Economic Research Working Paper Series*, No. 16616.

BBC (2010): "Facebook launches 'Zero' for mobile," *British Broadcasting Corporation*.

Bhattacharya, A., M. Romani, and N. Stern (2012): "Infrastructure for development: meeting the challenge," in *Centre for Climate Change Economics and Policy*, vol. 15.

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.

Blumenstock, J. (2012): "Inferring patterns of internal migration from mobile phone call records: evidence from Rwanda," *Information Technology for Development*, 18, 107–125.

Blumenstock, J., N. Eagle, and M. Fafchamps (2011): "Risk Sharing Over the Mobile Phone Network: Evidence from Rwanda," *Working Paper*.

Chandrasekhar, A., C. Kinnan, and H. Larreguy (2013): "Can Networks Substitute for Contracts? Evidence from a Lab Experiment in the Field," *Working Paper*.

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243–1284.

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.

de Montjoye, Y.-A., C. A. Hidalgo, M. Verleysen, and V. D. Blondel (2013): "Unique in the Crowd: The privacy bounds of human mobility," *Scientific Reports*, 3.

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. Kobrick, 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*.

Frost (2009): "Strategic Analysis of the Sub Saharan African Mobile Handsets Market," Tech. rep., Frost and Sullivan.

Futch, M. D. and C. T. McIntosh (2009): "Tracking the Introduction of the Village Phone Product in Rwanda," *Information Technologies & International Development*, 5, pp. 54–81.

Gillwald, A. and C. Stork (2008): "Towards Evidence-Based ICT Policy and Regulation: ICT access and usage in Africa," Tech. rep., Research ICT Africa.

González, M. C., C. A. Hidalgo, and A.-L. Barabási (2008): "Understanding individual human mobility patterns," *Nature*, 453, 779–782.

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.

Griliches, Z. (1957): "Hybrid Corn: An Exploration in the Economics of Technological Change," *Econometrica*, 25, 501.

GSMA (2006): "Universal Access: How Mobile Can Bring Communications to All," Tech. rep., GSM Association.

——— (2009): "Mobile For Development Intelligence," Tech. rep., GSM Association.

——— (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.

Isaacman, S., R. Becker, R. Cáceres, S. Kobourov, J. Rowland, and A. Varshavsky (2010): "A tale of two cities," in *Proceedings of the Eleventh Workshop on Mobile Computing Systems*, New York, NY, USA: ACM, HotMobile '10, 19–24.

ITU (2009): "Information Society Statistical Profiles: Africa," Tech. rep., International Telecommunication Union.

——— (2011a): "Measuring the Information Society," Tech. rep., International Telecommunication Union.

——— (2011b): "World telecommunication/ICT indicators database," Tech. rep., International Telecommunication Union.

Jack, W. and T. Suri (2011): "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution," *Working Paper*.

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.

Katz, M. L. and C. Shapiro (1986): "Technology Adoption in the Presence of Network Externalities," *Journal of Political Economy*, 94, 822–841.

——— (1994): "Systems Competition and Network Effects," *The Journal of Economic Perspectives*, 8, 93–115.

Kremer, M., J. Leino, E. Miguel, and A. P. Zwane (2011): "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions," *The Quarterly Journal of Economics*, 126, 145–205.

Milgrom, P. and C. Shannon (1994): "Monotone Comparative Statics," *Econometrica*, 62, 157–180.

Mulliner, C. (2013): "Public Research TAC Database," *Website*.

Narayanan, A. and V. Shmatikov (2008): "Robust De-anonymization of Large Sparse Datasets," in *IEEE Symposium on Security and Privacy, 2008. SP 2008*, 111–125.

Olken, B. A. (2007): "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy*, 115, 200–249.

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.

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.

TIA (2012): "ICT Market Review & Forecast," Tech. rep., Telecommunications Industry Association.

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.

WDI (2013): "World Development Indicators," *World Bank*.

Yanagizawa, D. (2012): "Propaganda and Conflict: Theory and Evidence From the Rwandan Genocide," *Working Paper*.

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

Appendix A. Measurement

This paper reports most prices in United States Dollars, but the prices charged to consumers are in Rwandan Francs (RwF). The exchange rate was relatively stable over the period of data (1.2005-5.2009), ranging between 543 and 570 RwF to the dollar.[82] I use the mean exchange rate of 552 RwF to the dollar.

From the data outlined in Section 3, I measure the following:

Account openings and closings. I infer an account as opened the date that the first transaction is made from it. Account are not explicitly closed; prepaid accounts that are not topped up regularly are disabled by the operator but can be used again when next topped up. Some accounts cycle through periods of being disabled but many are used again later; for this reason I ignore the possibility of account closure.

Communication graph (social network). Let the set of nodes S be the set of active accounts subscribing before 31 May 2009.[83] I define a directed graph over these nodes: there is a link from i to j if i has called j at least twice. Since the caller pays for a call, keeping track of the directionality is important. This realized network necessarily represents a subgraph of the country's full communication graph in two respects. First, it does not represent all nodes: consumers who subscribe after 31 May 2009, and consumers who subscribed to the minority operator are not included. For this reason I narrow the analysis to consider the behavior of the subset of individuals subscribing before this time. Second, within this subset of nodes, the realized edges are a subset of the edges of the underlying graph. Different edges would have been observed given different stochastic draws. This problem is partially mitigated by the length of the panel (up to 4.5 years).

Individual location. I observe the cell towers used to transmit each call, which is the only information I observe about each individual's location. To identify a location for each individual, I have modified Isaacman et al. (2011)'s 'important places' algorithm to improve performance in rural areas. The algorithm generates a set of important locations for each individual, $L_i = \{((x_{il}, y_{il}), d_{il})\}_l$, where (x_{il}, y_{il}) represents the geographical coordinates of i's lth location, and d_{il} represents the number of days that the user made transactions from that location. The identified locations have a slight bias towards the location of existing cell towers. This would be problematic for individual coverage estimation (it would tend to locate individuals in higher coverage areas than they actually are) but is mitigated by the fact that most of the transactions occur in the later years of the data when there is near universal coverage.

Raw coverage. I predict the coverage of mobile phone service at each location and time using tower locations and a elevation map.

[82] Average of selling and buying price, National Bank of Rwanda.

[83] I define an account as active if it made at least 10 outgoing calls, and the span of time between the first and last observed transaction is at least 90 days.

Tower coordinates (latitude and longitude) for most towers were provided by the operator. For towers whose locations are missing from these records, I infer the location using a procedure detailed in Appendix C. I infer the date each tower becomes operational by the date the first transaction that flows through it; I assume that once built, towers are never taken offline. Elevation data is from NASA Shuttle Radar Topographic Mission (SRTM) data, at 90m resolution; I use the version of the data from Jarvis et al. (2008), which has been processed to fill in data gaps.

If I had more information on the towers (specific equipment, tilt, antenna design), it would be possible to precisely predict coverage with commercial packages (the same as used by operators for coverage planning). As an approximation I predict coverage based on uninterrupted visibility, using the viewshed tool in ArcGIS. Based on the recommendations of the operator's network planner, I assume the antenna on each tower is located 35m above the ground, all antennas are omnidirectional, and that the signal has a maximum range of 15km.[84] I threshold the resulting image so that it indicates whether each location has coverage from at least one tower. This provides a raw coverage map for each month, which is my best estimate of the network availability at each location.

Individual coverage. The raw coverage maps indicate the coverage available at a given set of coordinates (x, y) during month t. However, phones are mobile, which has two implications: first, given that an individual is in a given location, there is a radius within which they are likely to make a call. In order to account for this, I compute a smoothed coverage map, so that $\phi_t(x, y)$ represents an average of the raw coverage available near (x, y), with the average weighted by a two-dimensional Gaussian kernel with standard deviation 25 pixels, or roughly 2.25km. Second, an individual may make calls from several locations, such as a village and the capital. To account for this, I compute an average of the coverage at each individual's important locations weighted by the days spent at each location: $\phi_{it} = \frac{\sum_l \phi_t(x_{il}, y_{il}) \cdot d_{il}}{\sum_l d_{il}}$.

Appendix B. Instruments

Here I describe the construction of the three instruments used to identify sensitivity to handset prices in the adoption decision, and present evidence on validity.

Slope instrument. I compute the slope of land for each coordinate (x, y) using ArcGIS, which is correlated with coverage but likely to satisfy the exclusion restriction.

Incidental coverage instrument. Towers are powered by electric lines or with generators. It is much cheaper to operate towers on the electric grid, and as a result the proximity to an electric grid is an important determinant of tower placement. However, while proximity to the grid directly affects the location of the towers themselves, given Rwanda's hilliness it is not the best measure of the resulting coverage. Instead I compute an incidental coverage

[84]Although the maximum technical range of a GSM tower is 35km, the range in practical use is smaller.

map: the coverage that would result from building towers along the full network of power lines.[85] These areas of the country had a higher ex-ante probability of receiving coverage because of the interaction between their geographic features and the existing electric grid. One obvious concern is that areas closer to power lines will have higher incidental coverage, and these areas may differ for other reasons that violate the exclusion restriction (for one, households are more likely to have electricity). For this reason, I use variation in incidental coverage only for locations further than 5km from the grid. For locations within 5km of the grid, I set the instrument's value to the mean value of incidental coverage outside the buffer region. The resulting instrument picks up incidental coverage based on geographical idiosyncrasies, such as whether they are on a hillside facing towards or away from a power line, for households further than 5km from the electric grid. See FIgure B.1 for a visual of the construction of the instrument.

Fraction of contacts receiving subsidized handsets. As detailed in Section 8, the Rwandan government allocated subsidized handsets to rural areas in the first few months of 2008. 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, which was otherwise rare in the country at the time. There are 41,225 such accounts. Then, for every individual, I compute the fraction of contacts that received subsidized handsets. Imagine two individuals who have yet to subscribe, who do not themselves receive a subsidy. The subsidy represents a shock that induces a fraction of their contacts to join. The one that has a higher fraction of contacts affected by the subsidy will receive a larger shock to the utility of being on the network. For individuals that had subscribed before the subsidy, the effect is ambiguous, because a higher fraction of contacts who are subsidized also implies a higher fraction of contacts who wait to join the network. For this reason I use variation in this instrument only for individuals subscribing after the beginning of the subsidy period in January 2008.

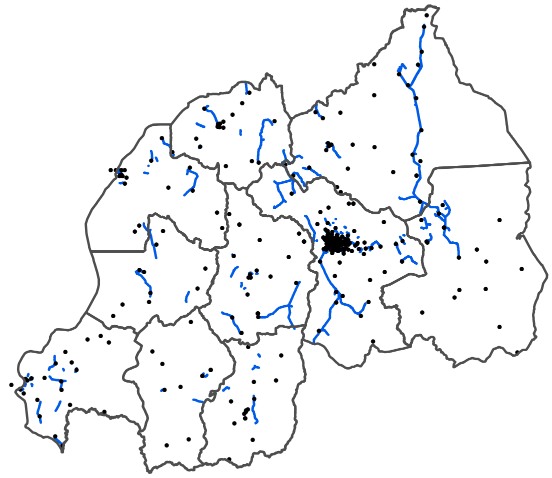
Tests. In order for the instruments to be valid, they must induce variation in the utility at adoption but be uncorrelated with the unobserved idiosyncratic benefit of being on the network (η_i—the exclusion restriction). Note that I observe a lot—every individual use of the phone. The idiosyncratic benefit would pick up differences in individuals' average valuations for calling, differences in the utility of owning a handset independent of the calling decision (such as SMS), or forecast errors in joining the network.

The two coverage instruments are correlated with coverage: my measure of slope is negatively correlated with coverage (-0.19 in 2005 and -0.13 in 2009), and incidental coverage is positively correlated with coverage, especially later in the data when more of the rural

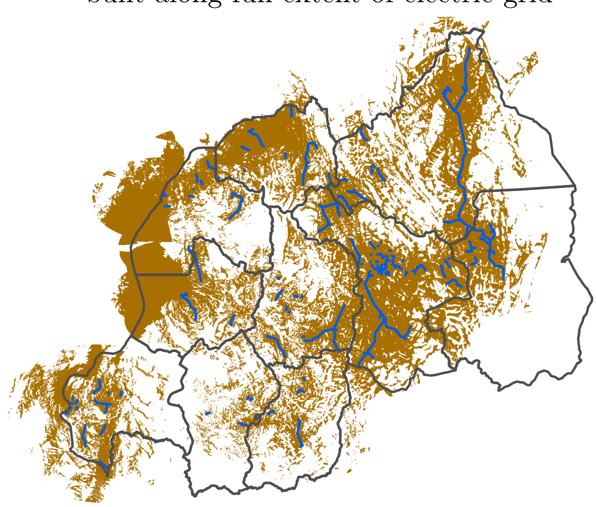
[85] I use a GIS layer of the electric grid as of 2008 provided by Rwanda's Energy, Water and Sanitation Authority.

Figure B.1. Incidental Coverage from Electric Grid

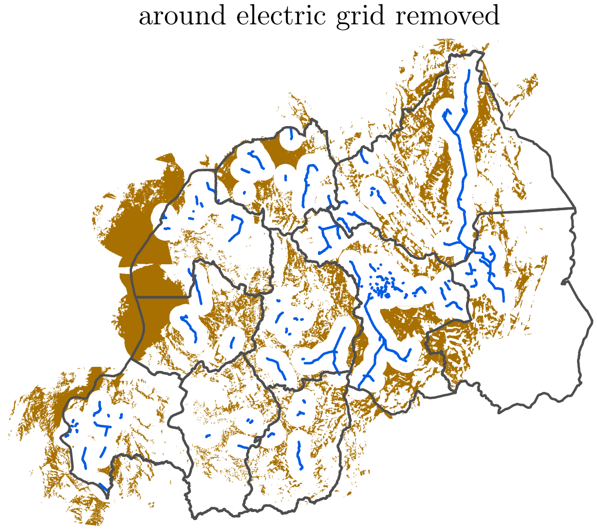
(a) Locations of electric grid and towers, January 2009



(b) Areas that would receive coverage if towers were built along full extent of electric grid



(c) Incidental coverage instrument, with 5km buffer around electric grid removed



network has been rolled out (0.08 in 2005 and 0.49 in 2009). The fraction of contacts subsidized is positively correlated with the total number of contacts subscribing over the months of the subsidy (0.14).

In Table B.1, I present correlations that measure mechanisms that I assume are excluded. Since I do not have standard characteristics for subscribers, I derive metrics from transaction data to describe channels that should be excluded. The first three columns represent correlations for the three instruments. As a comparison test, I include two more columns representing correlations with coverage at the beginning and end of the data; coverage itself is likely to fail the exclusion restriction because the operator is more likely to build towers in locations where individuals receive more idiosyncratic benefit from the network.

First, I consider measures of network structure. Individuals with different network structure may receive different benefits of being on the network: a trader with many dispersed contacts may receive a different utility than a mother in a rural area communicating with a few, well-connected family members. I present results for the number of contacts (degree) as well as for the clustering coefficient (the fraction of a node's neighbors who are themselves connected), both measured using the final network revealed through the end of the data, by which point coverage had expanded. I find that both measures are most correlated with coverage in January 2009 (0.11 and -0.15): individuals with higher coverage tend to have larger, more dispersed networks. The correlations with the instruments are much lower - the largest magnitudes come from slopes' correlation with contacts (-0.04) and clustering coefficient (0.04).

Last, I consider the quality of handset used, which is likely correlated with the unobserved benefit of adoption. In the model in the paper, for simplicity I do not consider differences in handset models, but for the majority of subscribers (960,854 out of 1,503,369) I know both the model of handset used and the price series specific to that model. As shown in Appendix E.2, although price differences between models could be quite large, their functionality was quite similar: most differences were in unobservable quality. To compare handset quality, I measure the price of each subscriber's chosen handset model as of the same date, January 2009. As shown in the last two columns, there is a correlation between coverage and this measure of handset quality (0.12 in 2005 and 0.10 in 2009): individuals who have higher coverage also have higher quality handsets. The correlation between this measure and the instruments is smaller: it is quite small for the coverage instruments (-0.03 for slope and 0.02 for incidental coverage); it is larger for the fraction of contacts subsidized (-0.07): individuals who have many contacts receiving subsidized handsets tend to have slightly lower quality handsets.

TABLE B.1. Correlations with Excluded Mechanisms by Instrument

Correlation	Instruments: Slope	Instruments: Incidental Coverage	Instruments: Fraction contacts subsidized	Comparison Test: Coverage January 2005	Comparison Test: Coverage January 2009
Number of contacts (Degree)	-0.04	-0.02	-0.01	0.03	0.11
Clustering coefficient	0.04	0.01	0.02	-0.11	-0.15
Price of handset model purchased, as of January 2009	-0.03	0.02	-0.07	0.12	0.10
N*	1,503,369	280,533	452,211	1,503,369	1,503,369
Sample	All	Primary location ≥ 5 km from electric grid	Subscribing after January 2008	All	All

All correlations have a p-value of 0.00. *: I can match the specific handset model a node is affiliated with to a price for 960,854 nodes. Correlations with handset price are computed on this subset.

APPENDIX C. ESTIMATING MISSING CELL TOWER LOCATIONS

I have location data for most high-volume cell towers; however, for some towers the location information is missing. Here I describe a new procedure to estimate the locations of missing towers based on call handoffs with known towers.

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

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

More formally, the procedure has two steps. The first step predicts the coordinates of an unknown tower $x_u = (x_u^{\text{long}}, x_u^{\text{lat}})$ by computing a weighted average of the coordinates of the set of known towers K, with weights w_{ku} of the relationship between k and u:

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

I use a simple metric based on the number of handoffs between towers k and u[86]:

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

However, while these predicted tower locations can lie anywhere, towers are generally built on ridges, to provide better coverage. The second step of the procedure adjusts the predicted location to lie on the nearest ridgeline. Since Rwanda is very hilly, this adjustment tends to be quite small (the mean adjustment is 0.59km, and the maximum is 2.15km); however, the adjustment can have a significant effect on predicted coverage since coverage is sensitive to the elevation of the tower.

To gauge the precision of tower estimates, I estimate the locations of known towers using leave-one-out sampling. Results are summarized by tower density in Figure C.1. The method performs best in areas of high tower density. The cluster of points to the right represents towers located in the capital.

The mean error in predicted location is 10.3km, with a standard deviation of 15.8km. The cluster of towers with high densities represents towers in the capital city. The method lacks a force that pushes predicted locations away from other towers; it might be possible to address this with a correction factor.

This problem is a special case of a more general inference problem on a network. We have a set of nodes $K \cup U$. There is a property of interest x_i, which is known for nodes $k \in K$ but not for $u \in U$. Nodes are connected by edges which are weighted by some distance metric, w_{ij}. I predict $\hat{x}_u$ by exploiting homophily on the network: nodes that are close as represented by w_{ij} are also close in x_i.

Appendix D. Inferring Subscriber Locations

The call data reports the location of the cell tower used at the start and end of each call. From the sequence of cell towers used, it is possible to infer an individual's location.

Before describing the algorithm used, it is relevant to discuss how a transaction is routed to a tower over a GSM network. At any point during a transaction, a mobile phone handset sends packets of information to one cellular tower, using electromagnetic waves. This tower

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

Figure C.1. Prediction Error

routes these packets to the rest of the network using either fiber optic cables or a different electromagnetic frequency; the packet is sent to a tower near the receiver and ultimately delivered to the receiver's handset.

Handsets tend to transmit information to the closest unobstructed tower, so that the tower used represents the closest approximation to the individual's location at that point in time. Calls can bounce between towers due to call traffic, variation in the weather, if a tower is down, or if the handset is in motion. The maximum technical range of a GSM tower is 35 km, but in areas of higher tower density the range is reduced to lower interference.

There is a literature on inferring a subscriber's location based on usage traces (González et al., 2008; Isaacman et al., 2010, 2011; Blumenstock et al., 2011). Most of this literature is based in the developed country context, and much is from urban settings. This paper's setting has several unique features relative to this body of work:

- Most work on location inference analyses mature networks using short panels with a fixed set of towers, but in this setting the tower network was rapidly expanding. In 2005, rural coverage was sparse; by 2009 essentially the entire country was covered. Since the measure of location is based on towers used, measures of location at early points in the data tend to be more biased towards the locations of towers, which tend to be in urban areas. By the end of the data period there is enough density in rural areas to identify rural users' locations.
- Usage is sparse. Rich location data is available for the small group of users who make many transactions; however, many users make very few transactions.
- Behavioral trends. It is common for users to make transactions immediately following a top up. Since users are observed only when transactions are made, this would tend to bias

location measures towards locations where they top up (market centers or urban areas). This would result in a valid measure of location where the user would like to make calls from conditional on the locations of agents, which is the location measure I am interested in for coverage. It is not clear how accurate it would be as a measure of a user's location of residence.

These make it clear that it is important to differentiate between location of residence (for matching to household surveys) and locations where phone service is desired (for estimation and simulation).

I implement a modified version of the 'important places' algorithm as detailed by Isaacman et al. (2011), which for each user identifies one or more important places where they spend time. The algorithm appears to work quite well: the paper reports validation results from the United States showing that the identified places were within 3 miles of reported places for 88% of a small validation sample of users, with a median error of 0.9 miles.

I have modified the algorithm to improve its performance in rural areas.

To find the important places for individual i, the algorithm proceeds as follows:

(1) The towers that i has ever used, X_i, are sorted by the number of days i used that tower, n_{ix}

(2) The most used tower forms the start of a new cluster, located at that tower's location.

(3) If the next most used tower falls within a distance threshold of the cluster, it is added to that cluster, and the cluster's location moves to its new centroid (weighted by the days each tower is used). If the tower does not fall within the threshold, it forms a new cluster. The original paper uses a fixed threshold of 1 mile, with which they obtain good results in an urban setting. To allow for good performance in urban and rural areas (high and low tower densities), I compute an adaptive threshold specific to each tower related to the density of towers nearby. In considering the distance from tower x to a cluster, I use a threshold equal to the distance from x to the 9th most distant tower as of May 2009. This adaptive threshold allows the algorithm to smoothly incorporate a large radius of spatial information in rural areas and a narrow radius in urban areas.

(4) The previous step is repeated for each tower: if the nearest cluster is within this tower's threshold, the tower is assigned to that cluster and that cluster's centroid is updated; if the nearest cluster is further away, the tower is assigned to a new cluster.

(5) After all towers have been placed in clusters, each cluster is ranked by the combined days that the individual made calls from that cluster (counting each day only once if transactions were made on multiple towers within that same cluster).

This algorithm has advantages for this setting: it uses the full panel of data, which improves precision when transactions are sparse, and works well with an expanding network: estimates simply become more precise as tower density increases. One weakness is that it does not account for migration: if an individual moved to a far away place, it would be counted as a new cluster; but if an individual moved to a location close by, it would result in one cluster accounting for both locations.

The determination of clusters could be disturbed if there were measurement error in the tower locations (e.g., a tower placed with error may end up bridging two clusters that should be separate). For this reason, in determining a subscriber's location I ignore the use of towers for which I have only a predicted location.

Appendix E. Data Appendix

E.1. **Handset models.** The call data reports a type allocation code (TAC) for each handset that was used, which can be mapped to the make and model. The GSM Association has an official registry matching type allocation codes to handset models, but access is restricted. I instead use an independent registry (Mulliner, 2013) to match handsets in the data to their model names. Using this database I am able to match all but 44 359 of the 1 377 836 total handsets used in the call data.

I match model names to characteristics and price series obtained from several sources, detailed in the following two sections.

E.2. **Handset characteristics.** Handset characteristics are gathered from two independent web sites, phonearena.com and gsmarena.com. For each handset I am able to obtain characteristics that are important for this setting, including battery life, the presence of an FM radio, flashlight, display quality (number of colors), and camera (number of megapixels).

E.3. **Handset prices.** In order to estimate the handset purchase equation, I compile a dataset of each handset model and how its price has changed over time.

Handsets can be purchased through the operator directly or through third parties. Over the period covered in the phone data, 1 333 477 handsets were activated, and 134 834 were listed in operator sales records, suggesting the operator sold approximately 10% of handsets.

I assemble one price series for each handset, making the assumption that at any point in time the price of a particular handset is uniform across the country. This is not wholly unreasonable, as the handsets all are imported through a small number of distributors.

The project uses three sources of historical handset prices: retail prices posted on the operator's website from 2004-2012, the operator's internal sales database from 2005-2012, and historical records from an independent shop in Kigali covering sales between 2005-2009.

Some of the operator's prices bounce around over time. Since I am interested in the nationwide price available to handset purchasers and not the specific prices offered by the operator, I assume that price decreases are not reversed: I take the cumulative minimum price as a measure of the price level of each handset. This would be invalid if observed price decreases represented temporary sales, or if there were factors that drove prices up (such as any spillovers from the post-election violence in Kenya of 2007-2008).

E.3.1. Integrating price series. I first assemble a panel of prices by handset model. If the operator's price is available for a given model-month, I use that. I fill in any missing data points with prices from the independent shop. This sometimes leaves observation gaps within a given model.

Once a model has been introduced, I assume its price declines predictably. I fill interior gaps (after the first observed price) with the price from the nearest observed month. I fill trailing gaps (after the last month I observe a price) with predicted declines from a regression describing how within-model prices decline over time. Specifically, I estimate the relation $p_t^h = \alpha_h + \beta t + \epsilon_{ht}$, with a fixed effect for each model and a general time trend. I find an average price decline of \$0.54 per month.

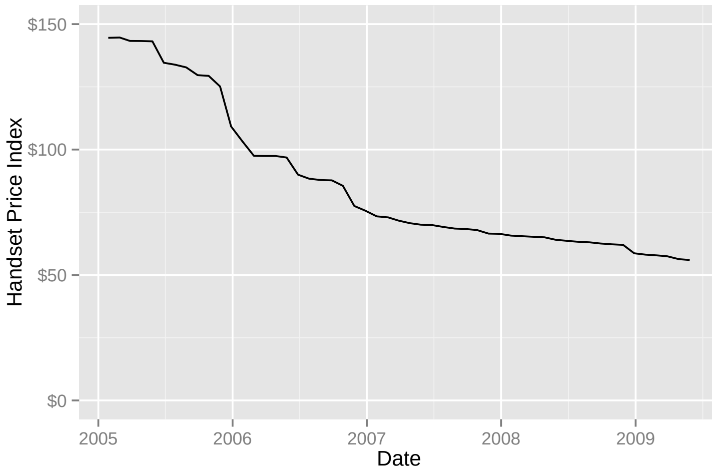
Many handsets were introduced during this time period. For these models I construct alternatives for the periods before that model was introduced. Although the most popular handset models are similar in observed characteristics, there is substantial variation in contemporaneous prices, suggesting differences in unobserved quality. The model could be enriched to allow for handset model choice, which would allow subscribers to choose a different handset in a prior period. For simplicity I omit handset model choice. Instead, when the selected handset was not available in prior periods, I use the price from alternative handsets of similar quality. I fill leading gaps (before that handset model was available for sale) by using the price series of the adjacent model of higher quality. I select the handset h' that was next most costly to h in the first time t when both were available (and thus presumably of higher quality). I fill in missing prices in this manner starting backwards from the end of the data towards the beginning.

E.3.2. Assembling a price index. The prior step generates a price series p_t^h for 160 handset models, where missing values before the introduction are imputed using the prices of the handset of nearest higher quality. From these series I assemble a single price index, weighting by the total quantity of each handset model activated in the data, Q_h:

$$p_t^{handset} = \frac{\sum_h Q_h \cdot p_t^h}{\sum_h Q_h}$$

Figure E.1 shows the resulting price index.

Figure E.1. Handset Price Index



ModelName	Date.First	Count.Activated	Radio	Flashlight	Battery_Standby_h	Camera_mp	Screen.Colors	Price.Highest	Price.Lowest
nokia-1110i	2006-05-10	128247	0	0	380	0.0	2	47.10	47.10
nokia-1200	2007-08-27	126597	0	1	390	0.0	2	63.41	28.92
nokia-1100	2004-12-31	117737	0	1	400	0.0	2	115.70	43.39
motorola-c113	2005-01-01	72290	0	0	450	0.0	2	28.00	28.00
zte-a35	2007-07-13	71302	0	0	150	0.0	2	27.96	22.18
nokia-1600	2005-09-05	54256	0	0	450	0.0	65536	81.52	63.41
nokia-3310	2004-12-31	51783	0	0		0.0	2	100.27	52.06
nokia-2310	2005-11-09	43893	1	0	400	0.0	65536	135.87	108.70
nokia-1110	2005-09-28	30341	0	0	380	0.0	2	53.99	30.85
nokia-2300	2004-12-31	29356	1	0	400	0.0	2	144.93	72.46
nokia-1650	2007-12-30	24210	1	1	420	0.0	65536	99.64	65.22
nokia-1112	2006-06-21	18182	0	0	380	0.0	2	48.59	48.59
motorola-c115	2004-12-31	15109	0	0	250	0.0	2	77.13	41.65
nokia-2100	2004-12-31	14473	0	0	150	0.0	2	192.83	86.77
nokia-6070	2006-07-21	13566	1	0	300	0.1	65536	163.04	115.70
nokia-2600	2005-01-01	13204	0	0	250	0.0	4096	131.12	72.46
siemens-a36	2004-12-31	11057	0	0		0.0	2	86.96	63.41
motorola-c200	2004-12-31	9422	0	0	192	0.0	2	110.51	38.57
nokia-5110	2004-12-31	8711	0	0		0.0	2	298.91	262.68
nokia-2626	2007-03-11	8293	1	0	300	0.0	65536	86.96	63.41

TABLE E.1. Top 20 Handset Models

Appendix F. Simplifications

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 0.53[87], 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.

Handset sharing. Given the high cost of handsets, sharing is common. 55% of phone owners report they allow others to use their handset regularly. There are two types of sharing: handset sharing and account sharing:

An individual may open an account but use it with others' handsets, by inserting their SIM card. This allows them their own phone number and balance, but it is difficult to receive calls. This practice is rare: 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.

It is more common that a person borrows another's handset and account.[88] These borrowing patterns cannot be observed in the data, and are difficult to ask about in surveys, so I omit the possibility of account borrowing. This is less of an omission than it may seem: borrowing is a hassle: it is difficult to receive calls, and making a call incurs a charge on the borrowed account. That subscribers are willing to spend significant sums of money to buy a personal handset suggests that borrowing is a poor substitute. Thus, although borrowing is common, the actual volume of calls due to borrowing is likely to be low. In the presence of borrowing, my model would allocate the surplus from borrowed calls accrues to the handset owner. The main impact of borrowing is on the outside option: I denote the utility of not having a handset as zero, but when phones can be borrowed, calls with high enough value can still be placed. This would have two effects. First, access to a borrowed handset is likely to be erratic, which would add additional fluctuations to the utility of the outside option. If the intervals chosen for the revealed preference estimation are long enough, they smooth

[87]There 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.
[88]This pattern would include the use of payphones that run on the mobile network.

out these idiosyncratic fluctuations. Second, waiting to purchase a handset would be more attractive; so simulated adoption dates would be biased forwards in time.

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.

Appendix G. Functional Form of Calling Utility

The form of calling utility matters for two reasons: it 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. The function should satisfy the following properties derived from theory and intuition:

(1) Price enters linearly, so that the duration choice is separable across contacts.
(2) *No utility from no call:* zero duration yields zero utility: $v(0, \epsilon) = 0$
(3) *Diminishing returns to duration:* $v(d, \epsilon)$ is concave in d
(4) *If a call is placed, it provides some utility:* 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.
(5) *Even if calls were free, you wouldn't talk forever:* there is bounded demand under zero price: $\frac{\partial v}{\partial d}(\bar{d}, \epsilon) = 0$ for some $\bar{d}$.
(6) *If the price is high enough, you wouldn't want to talk:* given a particular ϵ, there is a cost $\bar{c}$ for which d^* is zero:

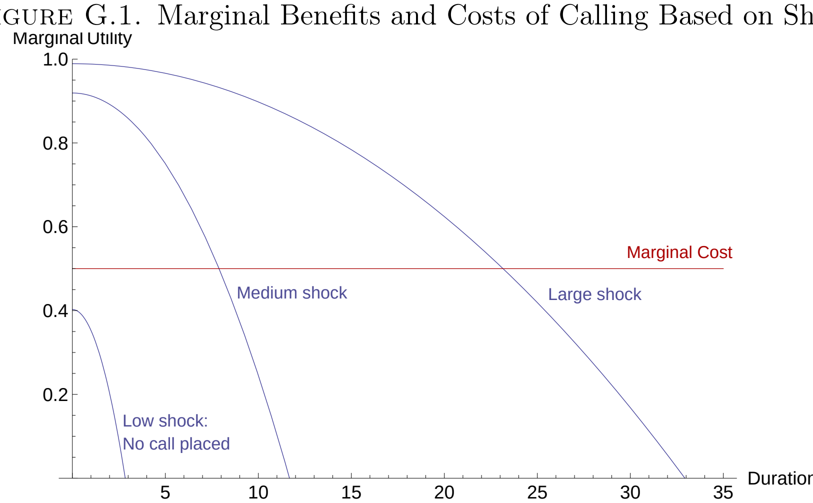
$$\left(\frac{\partial v}{\partial d}\right)^{-1} (\bar{c}, \epsilon) = 0$$

(7) *Changing the price of a call affects the extensive decision to call:* this requires that marginal utility be finite at zero:

$$\frac{\partial v}{\partial d}(0, \epsilon) < \infty$$

(8) *Changing the marginal price of a call affects longer calls more than short calls*: $\frac{\partial^2 d^*}{\partial c \partial \epsilon} < 0$
(9) *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.

FIGURE G.1. Marginal Benefits and Costs of Calling Based on Shock



(10) *Relationships with higher information flows provide more utility:* the optimized utility is increasing in the optimal duration:

$$\frac{\partial}{\partial d} v(\bar{d}, \epsilon(\bar{d})) > 0$$

In this paper I use the following specification:

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

which satisfies the above properties.[89] This functional form results in a marginal benefit of calling as depicted in Figure G.1.

[89]Two additional desirable property would be the following:
1. Higher information flows provide diminishing marginal utility: this would imply that the optimized utility is concave in the optimal duration: $\frac{\partial^2}{\partial d^2} v(\bar{d}, \epsilon(\bar{d})) < 0$. However, it is surprisingly difficult to find a functional form that satisfies this property in addition to the others. Since $\epsilon(d)$ is increasing in d, the concentrated function $v(d, \epsilon(d))$ is more convex in d than v; in order for it to still be concave, the marginal utilities corresponding to different ϵ's either need to cross for some value of d or bend away from the origin at low ϵ's and towards the origin at high ϵ's. The former implies that for a given level of prices, there is a maximum duration regardless of the shock ϵ (which is problematic). One candidate that does the latter is a power-error function $v(d, \epsilon) = -(d + e)^{\epsilon} + \alpha d + e^{\epsilon}$; however the resulting mapping $\epsilon(d)$ uses Lambert functions (the solution to the equation $x \ln x = y$), which do not have an analytic solution, which causes issues in estimation.
2. Some information is important enough you will share it regardless of the cost: this would imply for any cost c there is a ϵ for which $d^* > 0$. When costs are linear, this property requires ϵ to shift the intercept of marginal utility upwards without bound. However, with the current specification, this increases the convexity of the concentrated v function, which is undesirable. In this specification there exists a cost $\bar{c}$ such that $d^* = 0$ for all ϵ.

Appendix H. Estimation Notes

I split the estimation procedure into two steps, first estimating all parameters jointly for a subset of nodes and then each individual's parameters conditional on the estimated common parameters.

The first step itself is broken down into two substeps. The inner loop, Step B_i, chooses link-specific parameters Θ_i to maximize the likelihood of the calls from i, conditional on unified parameters Γ: $\hat{\Theta}_i(\Gamma) = \arg\max_{\Theta_i} \ln L_i(\Gamma, \Theta_i)$. The outer loop, Step A, chooses unified parameters to maximize the entire likelihood, concentrating out the optimal link-specific parameters: $\hat{\Gamma} = \arg\max_{\Gamma} \sum_{i \in S} \ln L(\Gamma, \hat{\Theta}_i(\Gamma))$. The B_i steps can be computed in parallel on small portions of the likelihood, which greatly improves computational performance. Due to some initial problems fitting extremely long calls, I estimated this dropping the 1% of links with calls longer than 30 minutes.

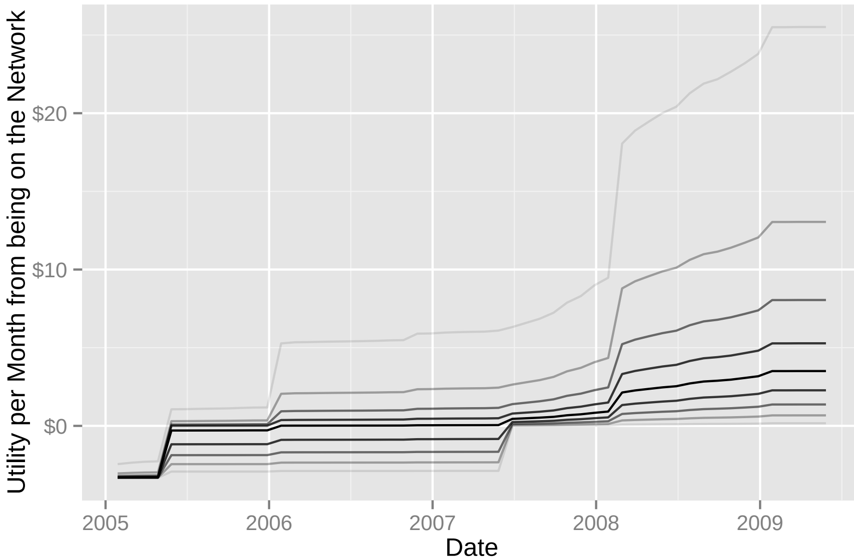
There is an issue with this approach relating to edge cases of coverage. The form of utility function implies that there is a cutoff level of cost (hassle cost of coverage and prices) above which no calls will be placed, regardless of the shock (I discuss this implication further in Appendix G). A random subset is unlikely to include the envelope of observations representing the highest cost instances under potential coefficient estimates. If common parameters are estimated off of such a subset, when applied to a set of links with a higher cost instance they could imply that an observed duration has zero likelihood (since the cost lies above the estimated cutoff). This would not be a problem if the full estimation could be done jointly. To correct for this, I constrain the estimation of the subproblem so that it is consistent with a nonzero probability of calling at the envelope of high cost instances.

Counterfactual utility. In estimating adoption, I compute the utility that an individual would have received in a given period had they owned a phone, holding fixed others adoption decisions. In Figure H.1 I graph the quantiles of this quantity, u_{it}, over time given the observed adoption sequence.

I make several observations:

In the first 4 months (January - April 2005), the operator charged a monthly fee of \$2.26 which was higher than the utility most eventual subscribers would have gotten from the network at that point, so most quantiles of utility are negative. After April 2005, the operator switched to a minimum top up requirement, requiring that each subscriber had to top up their balance with a minimum of \$4.53 every 30 days to keep their account active. Most eventual subscribers use their phones less than this minimum, and this policy appears to have been a substantial barrier to adoption. I assume this requirement incurred a utility cost of half of the extra amount that a consumer would have had to top up. This minimum requirement was active until June 2007. The graph shows that during this period, some eventual subscribers used the network enough to make their utility of being on the network positive, while others would still obtain negative utility from keeping a phone active.

Figure H.1. Quantiles of Monthly Utility u_{it} for Eventual Subscribers



In February 2006, effective calling prices were reduced through the introduction of per second billing, leading to a jump in utility especially for those who would have made more calls. In June 2007, the minimum top up requirement was lifted; after this point essentially all charges were on the margin, so that all utilities are weakly positive. In February 2008, the operator reduced calling prices. Other, continuous changes in utility are driven by changes in coverage and by contacts joining the network.

In selecting an adoption date, consumers weigh the future stream of utility from joining the network against the current price of a handset. Figure E.1 shows the handset price index.

Appendix I. Simulation Notes

Extrapolation after End of Data. The calling data ends at $\bar{T}^{data} = 53$ but I set $\bar{T} = 89$, three years beyond, corresponding to the last month I have handset price data. For $t \leq \bar{T}^{data}$, I use the formulation of expected utility as described by the model. For $t > \bar{T}^{data}$, there is not enough data to completely populate this model, so I use aggregate data on the expansion of the network. I assume the utility is a multiple of the utility from the last period, where the factor γ_t is derived from the increase in adoption from regulator statistics.

$$u_{it} = \gamma_t \cdot u_{i\bar{T}^{data}} \quad \text{for all } t > \bar{T}^{data}$$

This assumption is needed because the full network of calls after $\bar{T}^{data}$ is not observed; it implies that the increase in benefits accrues in the same proportion to each node in the network. I focus on how benefits improve due to increased adoption. I fit a sigmoid function to the total number of mobile subscribers as measured by the regulator, assuming a saturation point of 70%, resulting in predicted subscribers at month t of $\left|\hat{S}_t\right| = \frac{0.7 \cdot 11{,}000{,}000}{1+e^{-0.052(t-72.8)}}$. I then compute $\gamma_t = \frac{\left|\hat{S}_t\right|}{\left|\hat{S}_{\bar{T}^{data}}\right|}$, the proportional increase in subscribers over the number of subscribers in the last period with full data. This factor overstates the increase in benefits in two ways: at high levels of penetration the marginal benefit of an additional subscriber is likely declining, and the additional subscriptions measured by the regulator double count individuals who hold accounts with multiple operators, which becomes an issue after the third operator joins and the market becomes more competitive. It understates the increase in benefits in that it does not account for price declines associated with increasing competition (coverage remained relatively stable after $\bar{T}^{data}$). Overall, I believe the understatement and overstatement roughly net out. I assume that γ_t becomes stable in year 2025.

Handling subsidized nodes. Because the government adoption subsidy was time limited, recipients faced a large discontinuity in the cost of joining the network, requiring special care.

I back out recipients' idiosyncratic benefits η_i using two assumptions:

Recipients did not delay adoption in order to receive a subsidy. This would hold if either the subsidy was unanticipated, or if recipients anticipated the subsidy but expected it to be less generous than it was, which is reasonable given that recipients made fewer payments than originally intended. I back out the upper bound of the idiosyncratic benefit $\bar{\eta}_i$ assuming that two months prior to the subsidy, the recipient would still have waited to adopt even if the subsidy were not available. That is, I compute the upper bound for subsidy recipients using the standard price series not including the subsidy. Note that this assumes that the subsidy was not substantial enough to greatly shift expectations about future adoption.

Recipients preferred taking the subsidy at the point of adoption to purchasing any time in the following 4 years. Because the subsidy was time limited and very attractive, it causes a large discontinuity in the utility of adopting for recipients. For nonrecipients I compare the chosen date against local deviations of two months to back out bounds on the idiosyncratic benefit η_i. For nonrecipients these local deviations are near optimal, but for recipients local deviations may be far from the optimal: the next most optimal adoption date may be far later. Thus, for subsidized nodes I use global comparisons after adoption to back out η_i. Because the optimality of the chosen date should hold across the entire estimated set of baseline equilibria, this requires an extra step. To compute the lower bound of idiosyncratic benefits for subsidy recipients, I first simulate the baseline equilibrium holding recipients' adoption dates fixed. I then back out the benefits that would be consistent with adopting during the period of the subsidy over adopting in any following period, in all baseline equilibria:

$$\underline{\eta}_i = \max_{\Gamma} \max_{1 \le K \le \bar{T}-\tau_i} \left\{ -\frac{1-\delta}{1-\delta^K} \left[\sum_{k=0}^{K-1} \delta^k E \left(\sum_{j \in G_i \cap S_{\tau_i+k}} u_{ij\tau_i+k} + w \cdot E u_{ji\tau_i+k} \right) - \beta^{handset} (p^h_{\tau_i} - \frac{18.94}{552} - \delta^K p^h_{\tau_i+K}) \right] \right\}$$

I then use these lower bounds for the final simulations, in which subsidized nodes reoptimize their adoption dates.

Note that the resulting $\underline{\eta}_i$ for subsidized nodes represents a knife edge case. When computing the lower bound equilibrium, a small perturbation that causes one recipient to choose an alternative adoption date over the subsidized date would trickle through the network and cause many of the other recipients to give up the subsidy. In theory, the lower bound $\underline{\eta}_i$ could be instead backed out based on the assumption that recipients would prefer taking the subsidy even if no other recipients did, but that would require solving a computationally intensive fixed point (since obtaining the equilibrium when the subsidy is not taken up requires an estimate for $\underline{\eta}_i$). Instead, I address this with two restrictions:

- I hold the adoption dates of 30 subsidy recipients for which the bounds cross ($\underline{\eta}_i > \bar{\eta}_i$) fixed, since resolving these bounds would tip the equilibrium.
- When computing counterfactuals that do not relate to the subsidy, I hold fixed the adoption dates of the 41,225 subsidy recipients. If I instead allowed them to adjust their adoption, in the lower bound equilibrium many would choose not to take the subsidy, and I would conflate the effect of the subsidy and the alternate counterfactual.

Both of these restrict adjustment and thus will attenuate the estimated effect of a policy change.

Appendix J. Computation

The raw transaction records on which this project is based represent approximately 2 terabytes of data when compressed, which is too large to read into memory. To process them, I read in transactions line by line using scripts written in Python. Each script aggregates portions of the data into a data structure in memory for a specific question, and then writes that data structure to disk for further analysis. Most of the analysis uses these intermediate data structures generated from the raw data.

Most of the computational steps are fairly straightforward and can be run in parallel, with two exceptions. Some care is needed to parallelize the estimation of the unified parameters in the calling decision; this is detailed in Section 6 and Appendix H. The simulation method also requires special care due to the fact that decisions are interlinked and the objects that must be stored in memory are substantial (altogether, the entire process requires about 115 GB of memory).[90] While in principle the simulation method is parallelizable, obtaining a substantial boost in performance over a serial algorithm would require more access to a cluster than typical scheduling constraints allow.

Computation is performed on two systems: a dedicated server supported by the Center for Complex Network Research at Northeastern University, and the Odyssey2 cluster supported by the FAS Science Division Research Computing Group at Harvard University. Both systems have enough RAM to keep relevant data structures in memory without the need to

[90]Among other objects, it requires storing the links of the graph as well as the time series of utility for each link. The latter is required because I model utility as arising from links with different latent intensities (so one row is needed for each link) and whose calling depends on the changing coverage of both sender and receiver (so each link's utility has a different time path).

swap to disk (512 GB, and 256 GB per node). Most analysis is written in Python. I use the numpy and scipy packages for scientific computation, and the multiprocessing package to parallelize computation. Reduced form results for regressions with over 1 million observations are computed with incremental least squares using the pyrerp package written by Nathaniel Smith. I use a variety of optimizers: NLOPT and KNITRO, as well as optimizers built in to scipy. Where it significantly improves speed, innermost loops are written in C++, using the Boost library for scientific calculations; this code is inlined using the instant package. Some final analysis is coded in R, and graphs are produced using the ggplot2 package. ArcMap is used for GIS analysis; some further processing was done using Python and the Geospatial Data Abstraction Library.

Appendix K. Robustness

Below I present simulation results under the assumption that incoming calls are valued the same as outgoing calls ($w = 1$). Based on the comparison with the adoption decision, this assumption appears to roughly double count the surplus utility from calls; however, this double counting is accounted for in the estimation of the adoption decision. In the adoption decision I estimate the ratio of estimated usage utility to handset price at the point of adoption; the coefficient $\beta^{handset}$ scales up to absorb the overstated call utility: I find $\beta^{handset} = 0.3640$, 2.6 times the estimate when $w = 0$. Since I use this estimate to convert between calling utils and money, it roughly undoes the double counting. (If alternately we were concerned that $\beta^{handset}$ was affected by expectations or liquidity constraints, and that the price sensitivity derived from calling β^{call} represented a better estimate of dollar value of a util of communication, we could multiply the utility estimates by the ratio $\frac{\beta^{handset}}{\beta^{call}}$, which is 1.80 when $w = 1$ or 0.68 when $w = 0$. Doing so would assign zero value to nonvoice communication.) As a result, the effect of including utility from incoming calls is roughly to reallocate the surplus from calls evenly between sender and receiver. The choice of w primarily affects links with asymmetric communication.

In the following tables I present results when incoming calls are valued the same as outgoing calls. Table K.1 shows the results of rural adoption subsidy when $w = 1$ (the analogue of Table 9). Results are qualitatively similar but impacts are smaller: the program cost \$569,741, increased bounds on consumer surplus by \$909,519 (lower bound) and \$542,585 (upper bound), and revenue by \$795,872 (lower) and \$36,755 (upper). On net, it increased bounds on welfare by \$1,135,650 (lower) and \$9,599 (upper). Table K.2 shows the results of rural expansion when $w = 1$ (the analogue of Table 10). The cost of building and operating the towers was \$467,186. It increased bounds on consumer surplus by \$282,335 (lower) and \$235,312 (upper), and revenue by \$401,769 (lower) and \$326,238 (upper). On net, it increased bounds on welfare by \$216,918 (lower) and \$94,364 (upper).

Table K.1. Robustness: Impact of Adoption Subsidy Program: Call Surplus Allocated Equally between Sender and Receiver

		All nodes	Nodes by network distance to subsidized node 0	1	≥2
Number		1,503,369	41,225	728,347	733,797
Adoption Time (mean)					
Simulation					
... with subsidy	month	[26.96, 22.12]	[37.38, 37.38]	[26.17, 21.73]	[27.17, 21.64]
... no subsidy, only proximal effect of removal	month	[27.29, 22.17]	[49.06, 39.29]	[26.17, 21.73]	[27.17, 21.64]
... no subsidy, proximal and ripple effects	month	[27.35, 22.17]	[49.65, 39.30]	[26.26, 21.73]	[27.18, 21.64]
Total Impact of Subsidy	month	-0.39, -0.05	-12.27, -1.93	-0.10, -0.00	-0.01, 0.00
... proximal effect of removal	month	-0.32, -0.05	-11.68, -1.91	0.00, 0.00	0.00, 0.00
... additional ripple effect	month	-0.07, -0.00	-0.59, -0.01	-0.10, -0.00	-0.01, 0.00
Revenue (total)					
Simulation					
... with subsidy	million $	[215.50, 236.16]	[1.09, 1.19]	[132.15, 145.86]	[82.27, 89.11]
... no subsidy, only proximal effect of removal	million $	[214.99, 236.13]	[0.78, 1.16]	[131.95, 145.85]	[82.26, 89.11]
... no subsidy, proximal and ripple effects	million $	[214.71, 236.12]	[0.74, 1.16]	[131.73, 145.85]	[82.23, 89.11]
Total Impact of Subsidy	million $	0.80, 0.04	0.34, 0.03	0.42, 0.01	0.03, 0.00
... proximal effect of removal	million $	0.51, 0.04	0.31, 0.03	0.20, 0.01	0.00, 0.00
... additional ripple effect	million $	0.29, 0.01	0.03, 0.00	0.22, 0.01	0.03, 0.00
Consumer Surplus (total)					
Simulation					
... with subsidy	million $	[40.74, 48.04]	[1.48, 1.60]	[28.81, 33.91]	[10.44, 12.53]
... no subsidy, only proximal effect of removal	million $	[40.03, 47.50]	[0.91, 1.07]	[28.67, 33.90]	[10.44, 12.53]
... no subsidy, proximal and ripple effects	million $	[39.83, 47.50]	[0.89, 1.07]	[28.52, 33.90]	[10.42, 12.53]
Total Impact of Subsidy	million $	0.91, 0.54	0.59, 0.54	0.30, 0.01	0.02, 0.00
... proximal effect of removal	million $	0.72, 0.54	0.57, 0.54	0.15, 0.01	0.00, 0.00
... additional ripple effect	million $	0.19, 0.00	0.02, 0.00	0.15, 0.00	0.02, 0.00

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 30 subsidized nodes that have crossed bounds for η_i (for details see Appendix I). 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 K.2. Robustness: Impact of Rural Service Expansion: Call Surplus Allocated Equally between Sender and Receiver

		All nodes	Nodes by change in coverage	
			> 0.5%pt coverage change	≤ 0.5%pt coverage change
Number		1,503,369	160,154	1,343,215
Adoption Time (mean)				
Simulation				
... with expansion	month	[26.96, 22.12]	[32.03, 26.17]	[26.36, 21.63]
... no expansion, only immediate effect on calls	month	[26.96, 22.12]	[32.03, 26.17]	[26.36, 21.63]
... no expansion, full impact including adoption	month	[27.00, 22.14]	[32.20, 26.24]	[26.38, 21.65]
Total Impact of Expansion	month	-0.04, -0.02	-0.17, -0.08	-0.02, -0.01
Revenue (total)				
Simulation				
... with expansion	million $	[215.50, 236.16]	[9.64, 11.31]	[205.86, 224.85]
... no expansion, only immediate effect on calls	million $	[215.33, 236.02]	[9.60, 11.27]	[205.73, 224.74]
... no expansion, full impact including adoption	million $	[215.10, 235.84]	[9.55, 11.24]	[205.55, 224.60]
Total Impact of Expansion	million $	0.40, 0.33	0.09, 0.08	0.31, 0.25
... immediate effect on calls	million $	0.17, 0.15	0.04, 0.04	0.13, 0.10
... added effect through adoption	million $	0.23, 0.18	0.05, 0.03	0.18, 0.14
Consumer Surplus (total)				
Simulation				
... with expansion	million $	[40.74, 48.04]	[2.31, 2.94]	[38.43, 45.10]
... no expansion, only immediate effect on calls	million $	[40.60, 47.89]	[2.25, 2.88]	[38.34, 45.01]
... no expansion, full impact including adoption	million $	[40.46, 47.81]	[2.22, 2.85]	[38.24, 44.95]
Total Impact of Expansion	million $	0.28, 0.24	0.10, 0.09	0.19, 0.15
... immediate effect on calls	million $	0.15, 0.16	0.06, 0.07	0.09, 0.09
... added effect through adoption	million $	0.14, 0.08	0.04, 0.02	0.10, 0.06

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 30 subsidized nodes that have crossed bounds for η_i (for details see Appendix I). 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.

Appendix L. Predicting the Benefits of Mobile Internet Adoption

I model the adoption of mobile internet analogously to that of adopting a mobile phone for voice services. In this context, most potential adopters of mobile internet will be current mobile phone subscribers. Let $\tilde{S}_t$ represent the set of subscribers to internet service at time t.

For an existing mobile phone subscriber, adopting mobile internet entails upgrading to a smartphone with internet capability. Similar to voice use, there is no fee or contract associated with opening a data plan; data usage is charged on the margin, per kilobyte or fraction thereof.[91] I model this decision in analogue to the initial adoption decision: the utility to i of upgrading at time τ is given by:

$$\tilde{U}_i^{\tau} = \sum_{t=\tau}^{\infty} \delta^{t-\tau} \tilde{u}_{it} - \beta_i^{handset} (p_{i\tau}^{\tilde{h}} - p_{i\tau}^{h_i})$$

where $\tilde{u}$ represents the additional benefit of owning a smartphone over a feature phone, $p_{i\tau}^{\tilde{h}}$ represents the price of the smartphone, and $p_{i\tau}^{h_i}$ represents the trade-in value of the feature phone that i currently owns.

While the utility derived from being on the voice network is derived almost entirely from the utility of communicating with contacts, I assume that the utility from internet access has two components: a fixed benefit $\tilde{w}_i$ of being able to access monolithic web pages (such as YouTube, newspapers, etc.), and a network benefit derived from enriched communication with one's social contacts (representing e-mail, discussion forums, and social networking):

$$\tilde{u}_{it} = \tilde{w}_{it} + \sum_{j \in G_i \cap \tilde{S}_t} \tilde{u}_{ijt}$$

For current internet subscribers it would be possible to estimate the internet usage utilities $\tilde{w}_i$ and $\tilde{u}_{ijt}$ analogously to how call utilities were estimated in this paper. However, our interest here is to predict how nonsubscribers would use the internet.

In order to proceed, note two things. First, future mobile internet users are likely to be current mobile phone users, so their call data is available. Second, call data represents a rich projection of human behavior on a network; indicators derived from this data are likely to be predictive of mobile internet use. The total call volume between i and j itself is likely predictive of internet interaction between i and j, but there are a host of finer grained indicators on the edge ij that can more finely capture the form of relationship: the physical distance between the contacts, the contacts they have in common, the balance of who pays for calls, and whether the direction of calling is mediated by missed calls. Phone data can

[91]There are discounts for purchasing bundles of megabytes or days of unlimited use.

populate a set of individual characteristics X_i and link characteristics X_{ij}. I then seek to predict utility derived from mobile internet usage as a function of these characteristics:

$$\tilde{u}_{ijt} = f\left(X_i, X_j, X_{ij}, \tilde{p}_t, \tilde{\phi}_{it}\right) \quad \text{and} \quad \tilde{w}_{it} = g\left(X_i, \tilde{p}_t, \tilde{\phi}_{it}\right)$$

where $\tilde{p}_t$ is the price of data usage, and $\tilde{\phi}_{it}$ is the data coverage available to i.

The relations f and g can be estimated in one of two ways. For a technology that has been partially adopted, it can be estimated using actual usage data for the subset of nodes that currently subscribe, and predicted for nonsubscribers. This assumes that f and g are the same between these two groups. It is also possible to collect survey evidence that illuminates these quantities, for example, asking respondents to describe how valuable it would be to communicate with different contacts using the internet, as well as how they currently communicate with those contacts. This approach can be used on nonsubscribers directly and thus does not require the assumption that the equations are stable; however, it may be difficult for nonsubscribers to predict how they would use services that are unfamiliar.

These estimated relations make it possible to predict the latent utility derived from mobile internet usage for each individual $\tilde{w}_{it}$ and link $\tilde{u}_{ijt}$. The resulting predicted benefit network can be itself analyzed to explore the shape of the network and identify clusters of nodes to subsidize. A fuller analysis can be performed by using the simulation method outlined in this paper. Doing so requires an estimate of consumer expectations of future adoption. Expectations can be gathered for a sample from a survey instrument (asking respondents to predict when, if ever, each of their contacts would adopt mobile internet) and then predicted for the entire population. Populating the model with this information then makes it possible to run full counterfactual targeting simulations. Any key nodes that are identified can be targeted by an operator for discounts or marketing messages.

Table 1. Household Characteristics (Nationally Representative)

	All Households		Households with Mobile Phones	
	2005	2010	2005	2010
Fraction of households	1.00	1.00	0.05	0.40
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
Has piped water	0.02	0.05	0.38	0.11
Owns mobile phone	0.05	0.40	1.00	1.00
Owns fixed line phone	0.008	0.003	0.14	0.007
Owns radio	0.46	0.63	0.93	0.84
Owns television	0.02	0.05	0.41	0.12
Owns computer	-	0.02	-	0.04
Number of mobile phones	-	0.68	-	1.52
Household members	5.0	4.50	6.13	4.91

Sources: Consumption and last two rows: 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. A dash indicates that that question was not asked in that survey round.

Table 2. Monthly Usage (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 recharge	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.

TABLE 3. Determinants of Calling

	Duration - seconds per month, outgoing			
Contacts.Live.From		0.07026		0.05822
		(0.001383)		(0.00166)
Contacts.Live.To			0.05568	0.02247
			(0.001427)	(0.001713)
Price USD/minute	-142.4	-138.0	-138.9	-137.4
(event study)	(2.611)	(2.612)	(2.612)	(2.613)
Coverage.From	28.38	20.08	27.49	21.14
	(1.113)	(1.125)	(1.113)	(1.128)
Coverage.To	47.68	45.95	41.08	43.58
	(1.125)	(1.126)	(1.138)	(1.14)
Dummy: price regime 2	0.1827	-0.6579	-0.473	-0.7784
	(0.1155)	(0.1167)	(0.1168)	(0.1171)
Dummy: price regime 3	1.724	-1.895	-1.051	-2.395
	(0.1616)	(0.1766)	(0.1765)	(0.1806)
Link fixed effects	×	×	×	×
Month fixed effects	×	×	×	×
$N_{observations}$	15,616,407	15,616,407	15,616,407	15,616,407
N_{links}	564,672	564,672	564,672	564,672
R^2	0.0019	0.0021	0.0020	0.0021

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 one month window before and after. Dummies are included for the other months within each price regime, and month fixed effects are included to control for base monthly differences. 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.

Table 4. Parameter Estimates

Calling Decision

Unified Parameters	Standard Error
γ	0.0006
α	0.3292
β_{call}	0.0001
$\beta_{coverage.from}$	0.0051
$\beta_{coverage.to}$	0.0053
$\beta_{coverage.interaction}$	0.0079

Communication Graph

	Quantile:	0.01	0.25	0.50	0.75	0.99
Links (124.6m)	μ_{ij}	1.60	3.52	4.40	5.14	7.32
	SE(μ_{ij})	0.12	0.30	0.39	0.51	1.64
	N per link	6	19	45	52	53

	Quantile:	0.01	0.25	0.50	0.75	0.99
Nodes (1.5m)	σ_i	0.13	0.49	0.67	0.95	2.01
	SE(σ_i)	0.01	0.02	0.04	0.06	0.28
	q_i	0.06	0.21	0.44	0.82	1.00
	SE(q_i)	0.00	0.01	0.02	0.04	0.39
	N per node	13	227	637	2,464	27,725
Overall	N per parameter	6	21	41	46	51
	$N_{observations}$	4 billion				

Adoption Decision

	Parameter	Estimate
Adoptions (1m)	$\beta^{handset}$	0.1379

Usage decision parameters are estimated in a two step maximum likelihood procedure. In the first step, shape parameters are estimated jointly with node and link parameters for a random subset of 1,500 nodes and their 92,386 associated links, representing a total of 2,467,574 link-month observations. In the second step, shape parameters are held fixed while node and link parameters are estimated for the full sample. The second table reports the quantiles of estimates, quantiles of standard errors, and quantiles of observations per node and link. Since each node has two parameters plus one parameter per link, the number of observations per parameter will be lower. Standard errors reported in this table assume that there is no covariance between unified parameters and communication graph parameters, for computational reasons. The number of adoptions is lower than the number of nodes because some initial adopters joined the network before the start of the data.

Table 5. Call Model Comparative Statics

Price	Coverage			0.01	0.25	0.50	0.75	0.99
		Shock variance σ_i				Median: 0.67		
		Cost-independent censoring q_i				0.44		
		Quantile		0.01	0.25	0.50	0.75	0.99
		Shock mean μ_{ij}		1.60	3.52	4.40	5.14	7.32
Price	**Coverage**	**Outcome (Expectation)**						
		Duration conditional on call		8.1 sec	24.5 sec	43.2 sec	90.3 sec	11.1 min
		Probability of call		0.00..	0.04	0.22	0.37	0.43
Lowest	Full	Call cost	$	0.00...	0.004	0.04	0.13	1.14
		Hassle cost	$	0	0	0	0	0
		Net utility	$	0.00...	0.002	0.03	0.15	1.71
		Duration conditional on call		6.5 sec	18.7 sec	29.5 sec	58.6 sec	7.2 min
		Probability of call		0.00..	0.01	0.11	0.28	0.43
Lowest	Half	Call cost	$	0.00...	0.0007	0.01	0.06	0.74
		Hassle cost	$	0.00...	0.0004	0.01	0.04	0.40
		Net utility	$	0.00...	0.0002	0.01	0.04	0.68
		Duration conditional on call		6.3 sec	17.8 sec	27.2 sec	53.4 sec	6.5 min
		Probability of call		0.00..	0.01	0.08	0.25	0.43
Highest	Full	Call cost	$	0.00...	0.0009	0.01	0.06	0.74
		Hassle cost	$	0	0	0	0	0
		Net utility	$	0.00...	0.0001	0.01	0.04	0.68

Statistics computed on on 1% random subsample of nodes. The shock variance σ_i and cost-independent censoring parameter q_i are set to their medians, and outcomes from the model are shown for the range of shock means μ_{ij}.

Table 6. Allocation of Subsidized Handsets by District

Household properties		District Mean: Participating	District Mean: Nonparticipating	Difference p-value
Rural		0.94	0.73	0.04
Consumption per capita		$204	$334	0.03
Handsets allocated	Total	3556.8	0	0.00
	Per Household	0.05	0.00	0.00
Own mobile phone	2005	0.04	0.12	0.07
	2010	0.40	0.47	0.17
	Difference	0.36	0.36	0.76
N		15	15	

Sources: Handset allocations: Banque Rwandaise de Développement; other columns: EICV 2 and 3 surveys, National Institute of Statistics, 2005-2006, 2010-2011.

Table 7. Subsidy Allocation and Change in Phone Ownership

	Increase in Fraction of Households Owning Phones, 2005-2010			
Fraction allocated handsets in 2008	2.00	2.17	3.39	3.39
	(1.08)	(1.12)	(1.53)	(0.898)
Fraction owning phones in 2005		0.0645		0.0368
		(0.115)		(0.884)
Intercept	0.35	0.343	0.316	0.315
	(0.015)	(0.0197)	(0.0265)	(0.0451)
Districts included:	All	All	Participating	Participating
R^2	0.10	0.11	0.30	0.30
N	30	30	15	15

Estimates computed using ordinary least squares; robust standard errors in parentheses.

Table 8. Usage by Subsidy Recipients

		Accounts	**Accounts**	**Subsidy Recipients**
		All	Adopting 1-5.2008	Adopting 1-5.2008
Number		1,503,369	309,379	41,225
Rural	Mean	0.44	0.54	0.76
	SD	0.50	0.50	0.43
Calls	Mean	40.0	37.5	37.7
per month	Median	24.1	26.1	28.7
	SD	59.0	48.9	34.0
Duration	Mean	27.6	18.1	16.4
minutes per month	Fraction to accounts subscribing after 1.2008	24%	33%	35%
	SD	92.2	47.1	23.0
Number of Contacts	Mean	105.8	57.5	62.2
(Degree)	SD	159.9	73.4	42.8
Clustering Coefficient	Mean	0.068	0.081	0.082
	SD	0.066	0.070	0.057

Rural is defined as an account's mode tower being located in a rural area

Table 9. Impact of Adoption Subsidy Program

		All nodes	Nodes by network distance to subsidized node		
			0	1	≥2
Number		1,503,369	41,225	728,347	733,797
Adoption Time (mean)					
Simulation					
... with subsidy	month	[27.08, 22.58]	[37.38, 37.38]	[26.24, 22.05]	[27.34, 22.28]
... no subsidy, only proximal effect of removal	month	[27.69, 22.63]	[59.46, 39.39]	[26.24, 22.05]	[27.34, 22.28]
... no subsidy, proximal and ripple effects	month	[27.96, 22.64]	[62.77, 39.43]	[26.58, 22.06]	[27.38, 22.28]
Total Impact of Subsidy	month	-0.89, -0.06	-25.39, -2.05	-0.34, -0.01	-0.05, -0.00
... proximal effect of removal	month	-0.61, -0.06	-22.08, -2.01	0.00, 0.00	0.00, 0.00
... additional ripple effect	month	-0.28, -0.01	-3.31, -0.04	-0.34, -0.01	-0.05, -0.00
Revenue (total)					
Simulation					
... with subsidy	million $	[214.82, 235.27]	[1.09, 1.19]	[131.67, 145.16]	[82.06, 88.92]
... no subsidy, only proximal effect of removal	million $	[213.02, 235.18]	[0.29, 1.15]	[130.69, 145.11]	[82.04, 88.92]
... no subsidy, proximal and ripple effects	million $	[212.13, 235.15]	[0.23, 1.15]	[129.97, 145.08]	[81.94, 88.92]
Total Impact of Subsidy	million $	2.68, 0.13	0.86, 0.04	1.70, 0.08	0.13, 0.01
... proximal effect of removal	million $	1.80, 0.09	0.80, 0.04	0.98, 0.05	0.02, 0.00
... additional ripple effect	million $	0.88, 0.03	0.06, 0.00	0.72, 0.03	0.10, 0.00
Consumer Surplus (total)					
Simulation					
... with subsidy	million $	[53.79, 64.79]	[1.60, 1.74]	[37.38, 44.52]	[14.81, 18.53]
... no subsidy, only proximal effect of removal	million $	[51.74, 64.19]	[0.42, 1.19]	[36.52, 44.47]	[14.79, 18.53]
... no subsidy, proximal and ripple effects	million $	[50.84, 64.16]	[0.34, 1.18]	[35.82, 44.45]	[14.68, 18.53]
Total Impact of Subsidy	million $	2.95, 0.63	1.25, 0.56	1.57, 0.07	0.13, 0.01
... proximal effect of removal	million $	2.05, 0.60	1.17, 0.56	0.86, 0.05	0.02, 0.00
... additional ripple effect	million $	0.89, 0.03	0.08, 0.00	0.71, 0.02	0.11, 0.00

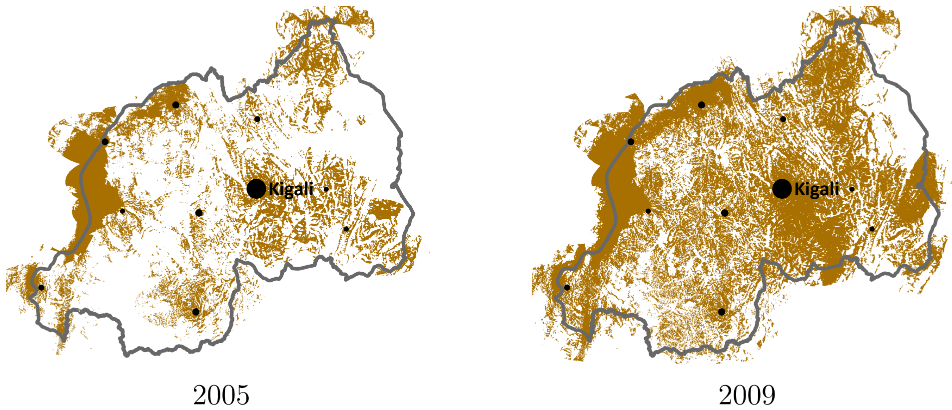
Results in each cell reported for the lower bound and upper bound estimate of the equilibrium. Impacts represent the difference in these bounds. I hold fixed the adoption decision of 30 subsidized nodes that have crossed bounds for η_i (for details see Appendix I). Utility and revenue reported in 2005 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 10. Impact of Rural Service Expansion

		All nodes	Nodes by change in coverage	
			> 0.5%pt coverage change	≤ 0.5%pt coverage change
Number		1,503,369	160,154	1,343,215
Adoption Time (mean)				
Simulation				
... with expansion	month	[27.08, 22.58]	[32.18, 26.80]	[26.47, 22.08]
... no expansion, only immediate effect on calls	month	[27.08, 22.58]	[32.18, 26.80]	[26.47, 22.08]
... no expansion, full impact including adoption	month	[27.11, 22.60]	[32.33, 26.87]	[26.49, 22.09]
Total Impact of Expansion	month	-0.03, -0.02	-0.16, -0.07	-0.02, -0.02
Revenue (total)				
Simulation				
... with expansion	million $	[214.82, 235.27]	[9.62, 11.25]	[205.20, 224.02]
... no expansion, only immediate effect on calls	million $	[214.65, 235.12]	[9.58, 11.20]	[205.07, 223.92]
... no expansion, full impact including adoption	million $	[214.44, 234.94]	[9.53, 11.17]	[204.91, 223.77]
Total Impact of Expansion	million $	0.37, 0.33	0.09, 0.08	0.28, 0.25
... immediate effect on calls	million $	0.17, 0.15	0.04, 0.04	0.13, 0.11
... added effect through adoption	million $	0.20, 0.18	0.05, 0.03	0.16, 0.14
Consumer Surplus (total)				
Simulation				
... with expansion	million $	[53.79, 64.79]	[3.27, 4.15]	[50.52, 60.64]
... no expansion, only immediate effect on calls	million $	[53.60, 64.58]	[3.20, 4.06]	[50.40, 60.52]
... no expansion, full impact including adoption	million $	[53.45, 64.47]	[3.16, 4.04]	[50.30, 60.43]
Total Impact of Expansion	million $	0.34, 0.32	0.12, 0.11	0.22, 0.21
... immediate effect on calls	million $	0.19, 0.21	0.07, 0.08	0.12, 0.13
... added effect through adoption	million $	0.15, 0.11	0.04, 0.03	0.10, 0.08

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 Appendix I). 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. Coverage change is calculated based on coverage in January 2009.

Figure 1. Coverage

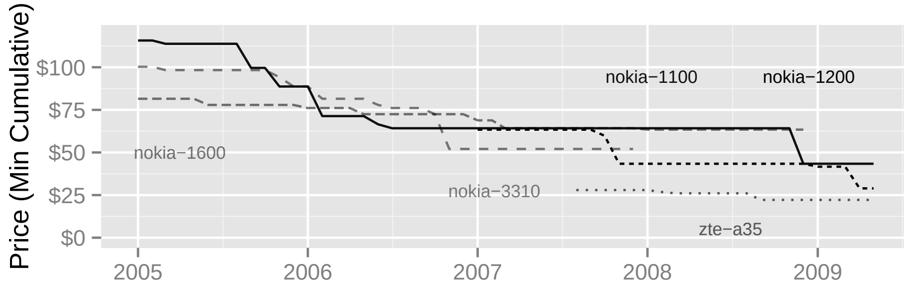


Locations with coverage are shaded. Cities are denoted by points sized by population.

Figure 2. Telephone Subscriptions in Rwanda

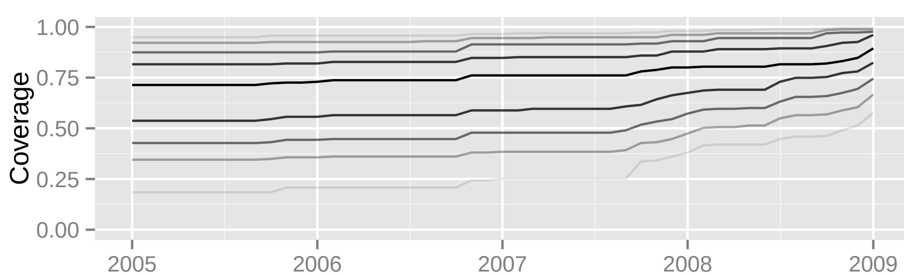
Figure 3. Variation in Data

Handset Prices (top 5 retail models)

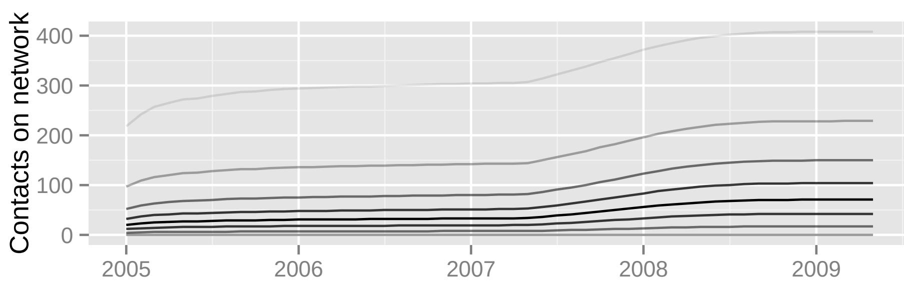


Calling Price (30 second call at peak rate)

Coverage (quantiles for eventual subscribers)

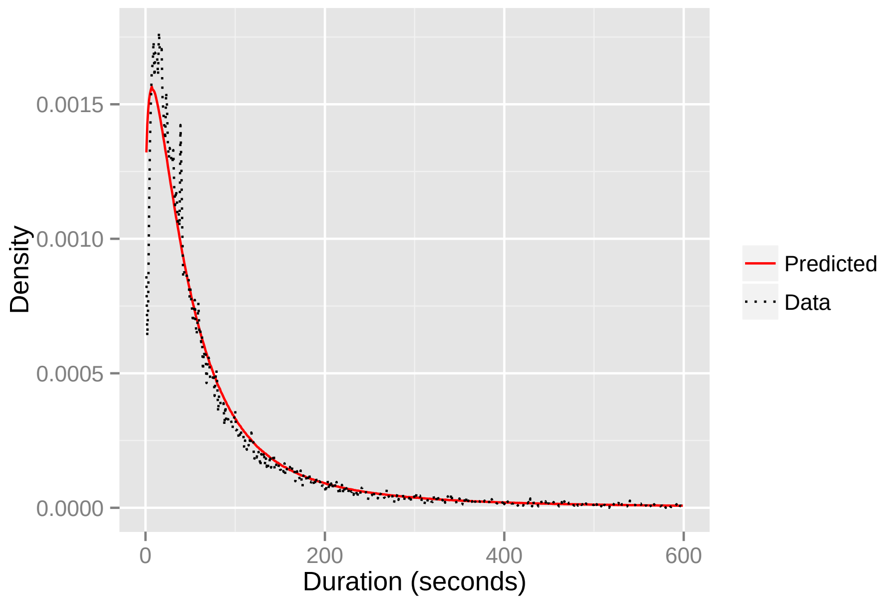


Contacts on the network (quantiles for eventual subscribers)



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.

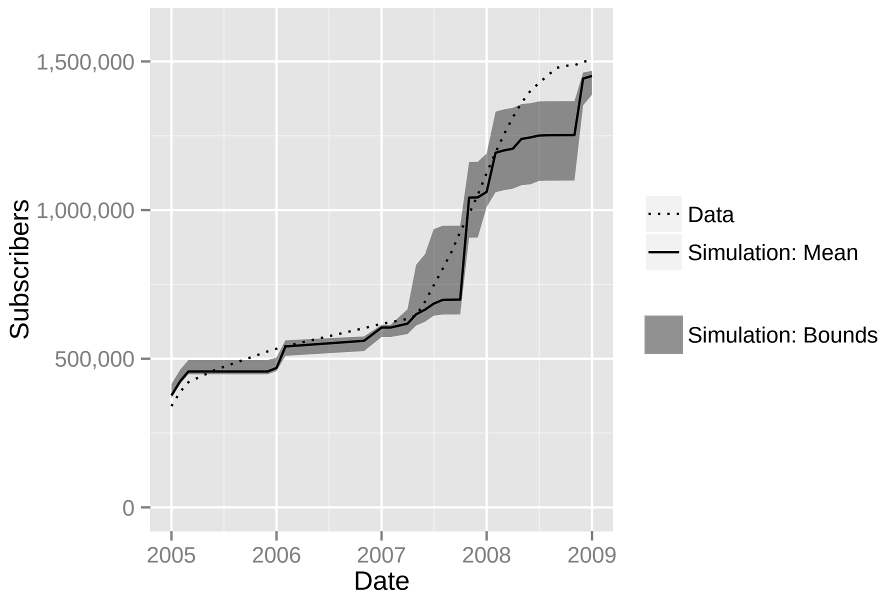
Figure 4. Call Model Fit



Months without call:	Data	88.0%
	Predicted	88.2%

Computed on random subsample of 10,000 links.

Figure 5. Simulation Fit



Fit metrics	Adoption month in data vs. adoption month under Lower equilibrium	Mean	Upper equilibrium
Correlation	0.86	0.87	0.83
Mean deviation	5.80	2.82	-0.83
Mean absolute deviation	6.63	4.56	5.08
Median deviation	5	2	-2
Median absolute deviation	5	3	4

Figure 6. Activations of Subsidized Handset Model

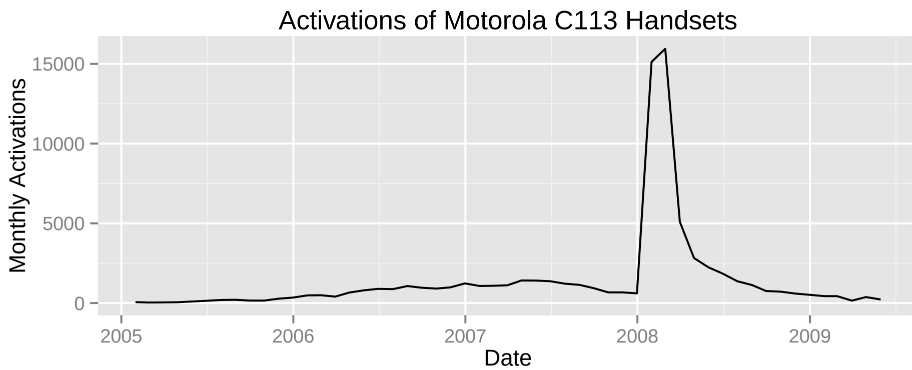
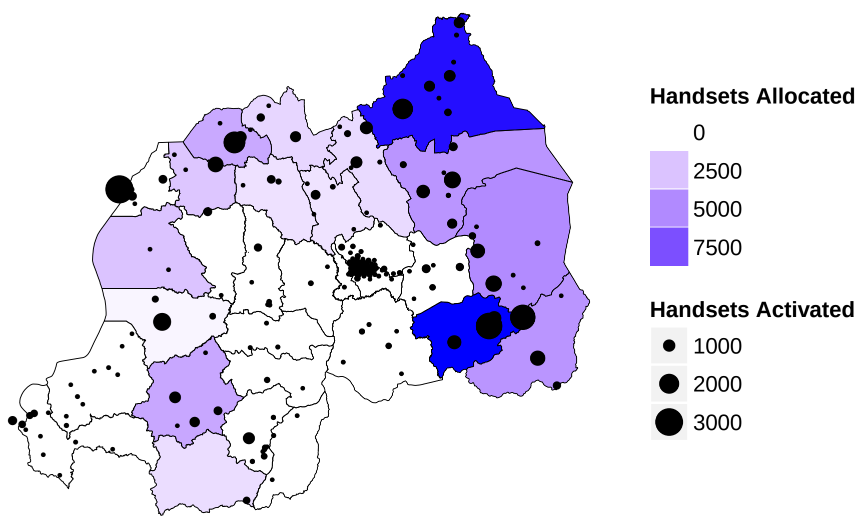
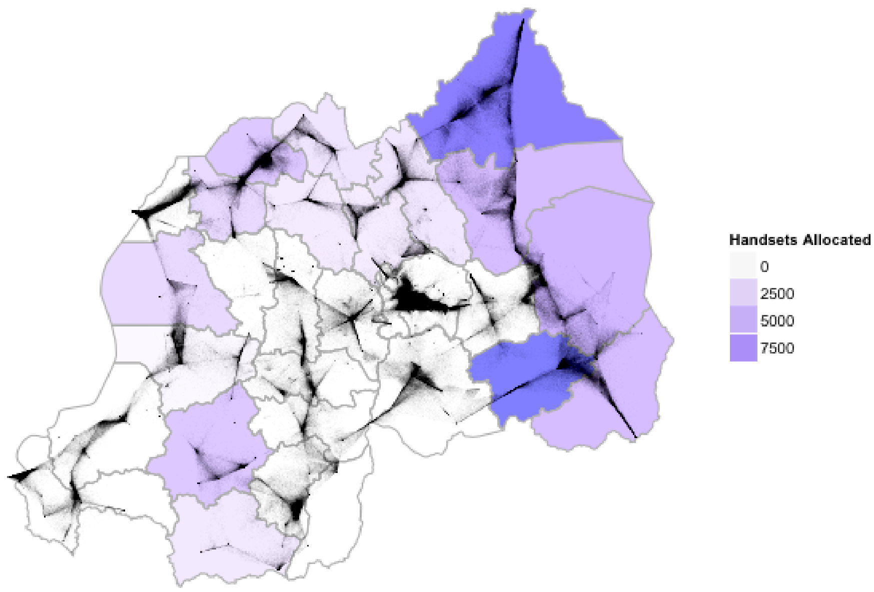


Figure 7. 2008 Handset Subsidy Program
Activation Locations by Handset



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

FIGURE 8. Usage Locations by Account
(a) All accounts



(b) Accounts affiliated with a subsidized handset

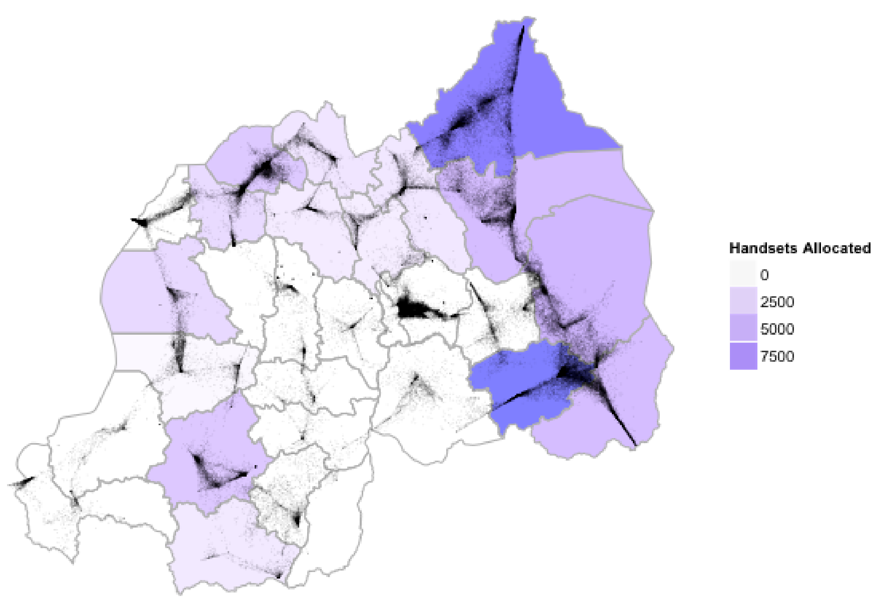
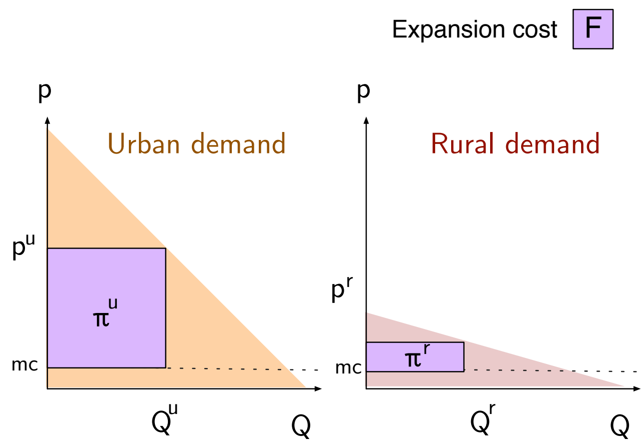
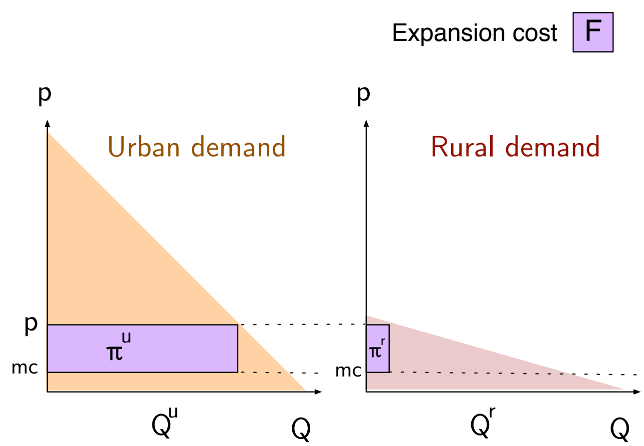


Figure 9. Incentives for Rural Expansion

(a) Operator allowed to set separate prices

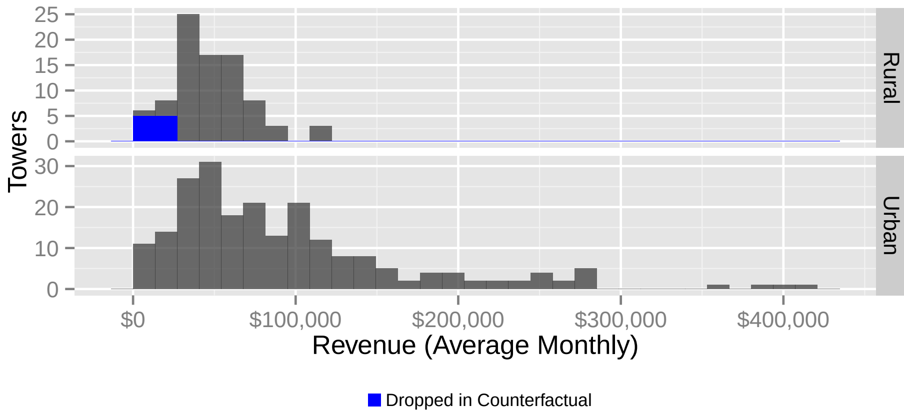


(b) Operator required to charge uniform prices



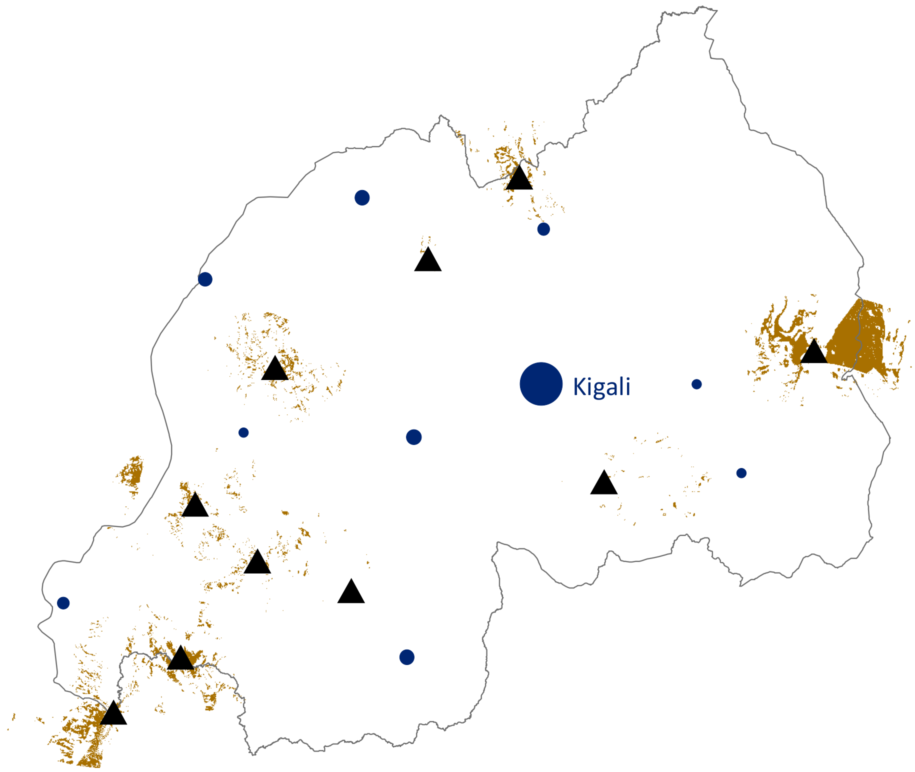
These panels show a simplified version of a monopolist operator's price setting problem, abstracting away from the specifics of pricing under capacity constraints and network effects. Assume there is a monopolist operator who currently serves an urban market, and has the option of expanding into a rural area at an expansion cost F. If the operator is allowed to set separate prices, it can treat the rural market independently, and weigh the profits resulting from an optimal price against the expansion cost, as shown in panel (a). If instead the operator is required to offer a uniform price, its decision to expand will also be affected by the urban market, as shown in panel (b).

FIGURE 10. Distribution of Tower Revenue



Includes revenue from 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 10 lowest revenue rural towers built during the data. There were other low revenue towers built before the start of the data; since the initial adopters in the data would have internalized the coverage provided by these towers in their adoption decision, I do not drop these towers.

Figure 11. Areas Losing Coverage in Counterfactual



Cities are denoted by circles, dropped towers are denoted by triangles, and locations that received coverage in the data in 2009 but not in the counterfactual are shaded.