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

DANIEL BJÖRKEGREN*

BROWN UNIVERSITY

This paper develops a method to estimate and simulate the adoption of a network good. I estimate demand for mobile phones as a function of individuals' social networks, coverage, and prices, using transaction data from nearly the entire network of Rwandan mobile phone subscribers at the time, over 4.5 years. I estimate the utility of adopting a phone based on its eventual usage: subscribers pay on the margin, so calls reveal the value of communicating with each contact. I use this structural model to simulate the effects of two policies. A requirement to serve rural areas lowered operator profits but increased net social welfare. Developing countries heavily tax mobile phones, but standard metrics that neglect network effects grossly understate the true welfare cost in a growing network, which is up to 3.14 times the revenue raised. Shifting from handset to usage taxes would have increased the surplus of poorer users by at least 38%.

JEL Classification Codes: O33, L96, O180, L51

Keywords: network goods, infrastructure, information technology

---

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

Revision May 22, 2017.

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

1. Introduction

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

Firms and governments use many different policies to guide the provision and adoption of network goods. While theoretical work provides intuition about network effects, there is little empirical work to guide policy choices.[1] Empirical work has been limited for three reasons. It is costly to measure an entire network using traditional data sources. It is also difficult to identify network effects: one individual may adopt after a contact adopts because the contact provides network benefits, or because connected individuals share similar traits or are exposed to similar environments. And even if these two issues are overcome, it is difficult to evaluate policies, which can cause effects to ripple through the entire network. As a result, there remain open questions about how to design policies that better capture the spillover benefits associated with network effects, as well as policies that overcome suboptimal provision arising from high concentrations in industries providing network goods.

This paper overcomes these limitations by combining a new empirical approach with rich data from nearly an entire country's remote communication system. I use 5.3 billion transaction records from Rwanda's dominant mobile phone operator, which

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

held over 88% of the market, during a period of dramatic expansion. I estimate a structural model of demand for mobile phones, and then demonstrate how this model can be used to simulate the effects of policies.

My empirical approach has three parts:

First, acknowledging that the utility of owning a mobile phone is derived from its usage, I model the utility of using a phone. I infer the value of each connection from subsequent interaction across that connection. This approach bypasses most of the simultaneity issues that result from inferring value from correlations in adoption. In the Rwandan system I study, 99% of accounts are prepaid: the person placing a call pays for it on the margin, by the second.[2] A subscriber must value a connection at least as much as the cost of calls placed across it.[3] Further, because the firm changed calling prices and increased the quality of service, I can identify the underlying demand curve for communication across each link.

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

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

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

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

affect the prices of handsets or calls, coverage, or adoption directly, in ways that can vary across time, or target specific regions or nodes in the network.

The estimation portion of my approach has parallels with Ryan and Tucker (2012). Ryan and Tucker (2012) infer the value of a videoconferencing system from the number of calls placed across it, in a setting with no cost of adoption or of placing calls, and a static firm policy. In contrast, in my first step I use changes in firm policy to identify the value of each link, accounting for the potential endogeneity of these changes as the network expands. Because I model the full structure of the network, my results factor in the position of each node in the network: an isolated node will tend to have less spillover effects than a central node, independent of the intensity of its links. Finally, while Ryan and Tucker (2012) cannot feasibly compute counterfactual equilibria in even a small network, my approach can tractably compute equilibria in large networks, and can thus be used to evaluate general policies. I turn my method to policy questions facing developing countries.

The spread of mobile phones across the developing world has been dramatic: between 2000 and 2011, the number of mobile phone subscriptions in developing economies increased from 250 million to 4.5 billion (ITU, 2011). Improvements in communication, through mobile phones as well as associated services such as mobile money and mobile internet, have the potential to knit even remote villages into the global economy. But these technologies are easily taxed and thus represent a public finance opportunity: the mobile industry contributed an average of 7% of government revenues in sub-Saharan Africa as early as 2007 (GSMA, 2008). Developing countries thus face a tension between generating revenue and extending service, particularly to rural and low income areas ('a paramount concern' in the words of former World Bank ICT Director Mohsen Khalil). Governments typically manage this tension with a set of telecom-specific taxes, and regulations and programs that encourage access to the rural poor. However, there is little evidence to guide the design of these policies, and standard approaches that do not account for network effects can give misleading estimates.

I use my approach to evaluate two policies.

I analyze the welfare implications of providing coverage to rural areas. A social planner would expand coverage until the point where building any marginal set of towers would not improve welfare. Firms may stop building before reaching this point: in a competitive market, some of the benefits of expanding coverage will spill over into competitors' networks.[4] And regardless of market structure, firms are unlikely to internalize all of the value generated for consumers: price discrimination is limited practically, and often also by regulation. Depending on the shape of private and social benefits from expansion, it may be optimal for a government to require the provision of coverage to areas that are unprofitable to serve. I find that in Rwanda, a government coverage obligation led to the building of roughly 11% of the rural towers active by 2009, which were unprofitable for the firm but welfare improving for the country.

I also use my model to evaluate the potential of telecom taxation to generate government revenue. I find that the baseline tax regime, which taxed both handsets and usage, had a substantial welfare cost that would be underestimated if network effects were ignored. For example, measures based on micro elasticities commonly used in the industry (e.g., GSMA, 2008) would suggest that taxing handsets imposes a welfare cost below \$1.22 per dollar of government revenue; but I find that including network effects the true cost is above \$2.93, and as high as \$5.21 in the early stages of the network. Further, baseline taxes heavily burden poorer users: the lowest half of users receive only 2% of consumer surplus, but account for 19-20% of government revenue. In 2010, the Rwandan government shifted taxes from handsets to usage. I find that shifting these taxes earlier would have increased the consumer surplus of the lowest half of users by at least 38% without substantially lowering the surplus of the upper half of users.

This paper connects with several literatures:

This paper studies classic network goods, whose value depends directly on the network of other users. Several empirical studies have measured the existence or extent of network effects in different environments (for example, Brynjolfsson and

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

Kemerer (1996); Goolsbee and Klenow (2002); Ackerberg and Gowrisankaran (2006); Tucker (2008)). Conceptually related are goods with indirect network effects, such as platforms and video formats (Ohashi 2003; Gowrisankaran et al. 2010; relatedly, Lee 2013): popular platforms tend to be better served by sellers, so adopters benefit indirectly from additional users.

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

The next section describes the expansion of mobile phone networks worldwide and in Rwanda. Section 3 describes the data. Section 4 presents stylized facts about mobile phone usage in Rwanda. Section 5 introduces a model of phone adoption and usage, Section 6 describes how it is estimated, and Section 7 describes how it is used to simulate counterfactual policies. Section 8 analyzes operator incentives to provide service in rural areas, and Section 9 analyzes telecom taxation. 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. Handsets were initially expensive, but became accessible to poor consumers as component costs decreased. Operators adapted to this broader base of potential subscribers by expanding coverage beyond urban centers and reducing usage prices.

Rwanda between 2005-2009 is an attractive setting to study the spread of mobile phones in developing countries. Because the Rwandan regulator restricted entry, the market during this period was extremely concentrated: the mobile operator whose data I use held above 88% of the market, and its records reveal nearly the entirety of the country's remote communication. There are few alternatives for remote communication: the fixed line network is small (with penetration below 0.4%), and mail service is insignificant.[5] The data on which this project is based is long enough to capture both adoption and use decisions for a substantial fraction of the population, as well as substantial variation in prices and coverage.

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) (all figures reported in real 2005 USD unless noted). 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 Rwandan Genocide, the new government spurred the development of a mobile phone network. An exclusive license was given to a multinational operator, which started operations in the capital, Kigali, in 1998. Service quickly spread from Kigali to other urban centers. Two changes influenced further rollout.

As global handset prices declined, it became cheaper to adopt a phone: 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 $13.

The regulator also induced a change in market structure. In 2003, the government announced it would provide a license to a second operator, which entered the market

[5]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).

Table 1. Household Characteristics (Nationally Representative)

	All Households		Households with Mobile Phones	
	2005	2010	2005	2010
Consumption per capita (real)	\$243.21	\$264.56	\$849.67	\$394.71
Monthly spending on airtime	-	\$2.43	-	\$5.28
Rural	0.85	0.86	0.23	0.75
Has electricity	0.05	0.10	0.62	0.22
Owns fixed line phone	0.008	0.003	0.14	0.007
Owns mobile phone	0.05	0.40	1.00	1.00
Owns radio	0.46	0.63	0.93	0.84
Owns television	0.02	0.05	0.41	0.12
Proportion of households	1.00	1.00	0.05	0.40

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

in 2005. This second operator was unsuccessful: it reached a maximum of 20% market share for a brief period after the end of my data, and in 2011 its license was revoked for failure to meet obligations. In combination with providing a second license, the government attached minimum coverage obligations to the first operator's license.[6]

From 2005-2009, the dominant operator reduced the real price of phone calls by 76% and nearly quadrupled the number of cell towers, increasing coverage from 60% to 95% of the country's land area. These changes 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 characteristics of the Rwandan population and these changing

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

demographics of phone owners and Figure 1 shows the changes in prices, coverage, and network adoption.

3. Data

This project uses several data sources:[7]

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 during this period. This data includes nearly every call, SMS, and top up made over 4.5 years by the operator's mobile phone subscribers, numbering approximately 400,000 in January 2005 and growing to 1.5 million in May 2009. For each transaction, the data reports: anonymous identifiers for sender and receiver, corresponding to the phone number and handset, time stamps, the location of the cell towers used, and call duration.[8] I aggregate durations to the monthly level.

Coverage: A rollout plan, $\mathbf{z} = \{(t_z, x_z, y_z)\}_z$, is defined by tower build dates and geographical coordinates. I consider the baseline rollout, $\mathbf{z}_0$, as well as counterfactual rollouts. I create coverage maps by computing the areas within line of sight of the towers operational in each month, a method suggested by the operator's network engineer. Elevation maps are derived from satellite imagery recorded by NASA (Jarvis et al., 2008; Farr et al., 2007).

Instrument for Coverage: I also create an instrument for coverage based on the interaction of geographical features with the electric grid: the incidental coverage that would arise had the operator built towers along the entire grid.

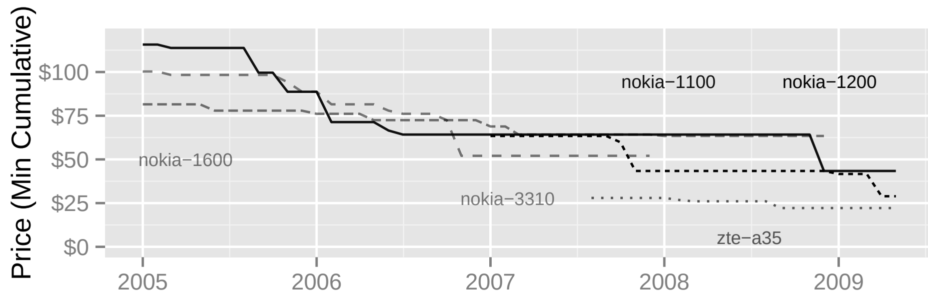
Individual locations and coverage: Because handsets are mobile, an individual may make calls from several locations, such as a village and the capital. I infer the geographical coordinates of each subscriber i's set of most used locations, $\{(x_{il}, y_{il})\}_l$, based on their eventual calls, using an algorithm analogous to triangulation (a version of Isaacman et al. (2011) that I have modified to improve performance in rural

[7]More details about measurement are described in the Supplemental Appendix.

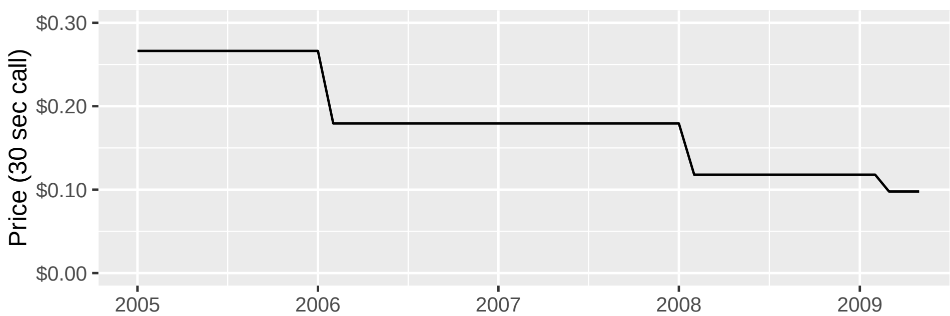
[8]Some months of data are missing; from the call records: May 2005, February 2009, and part of March 2009. The locations of 12% of tower identifiers are missing from this data; I infer their location based on call handoffs with known towers using a procedure I have developed (Bjorkegren, 2014).

Figure 1. Variation in Data

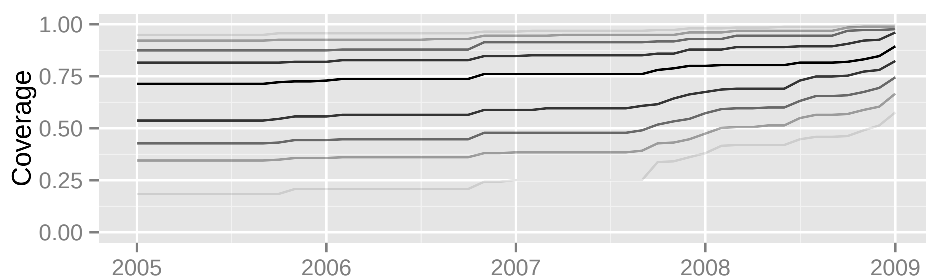
Handset Prices (top 5 retail models, nominal)



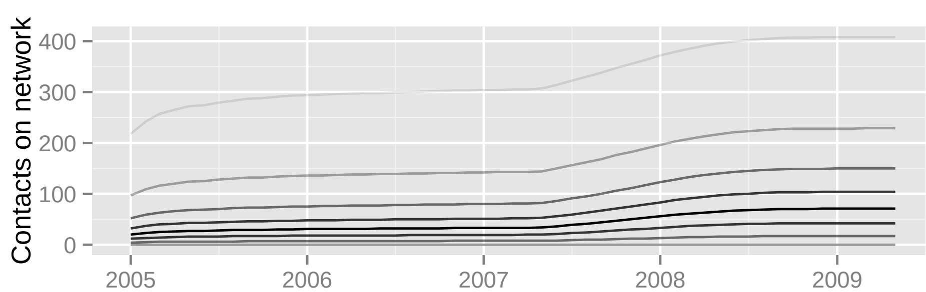
Calling Price (30 second call at peak rate, nominal)



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. Nominal prices shown in USD.

areas). Because individuals may use the phone in the area surrounding each location, I compute a smoothed coverage map, where $\phi_t(x, y; \mathbf{z})$ represents an average of the coverage available near (x, y) under rollout plan $\mathbf{z}$, weighted by a two-dimensional Gaussian kernel with radius 2.25km. I compute an individual specific coverage sequence by taking the average of the coverage at each individual's important locations weighted by the days spent at each location, d_{il}: $\phi_{it}(\mathbf{z}) = \frac{\sum_l \phi_t(x_{il}, y_{il}; \mathbf{z}) \cdot d_{il}}{\sum_l d_{il}}$.

Handset prices: I create a monthly handset price index $p_t^{handset}$ based on 160 popular models in Rwanda, adjusting for quality and weighting each model by the quantity activated on the network.

Household surveys: I report background statistics from nationally representative household surveys, including DHS and government surveys (EICV) from 2005 and 2010, and a technology usage survey (Stork and Stork, 2008).

4. Patterns of mobile phone use

The primary unit of observation is an account, which corresponds to a phone number. I assume that each account is associated with a unitary entity such as an individual, firm, or household. This assumption would be violated if the composition of people using a handset changed over time.[9] For ease of exposition 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 for remote communication, which I refer to as the communication graph. I model the utility of communicating with a fixed potential set of contacts, which may represent family, friends, or business contacts.

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

[9] There were an average of 1.03 accounts per phone owner (Gillwald and Stork, 2008), and little reason to change accounts (there was one majority operator, and unused prepaid accounts are reactivated when more credit is added). If the composition of people using a particular phone changes in response to adoption (say, if a couple initially shares a phone but later obtains separate phones and splits its communication), then the communication graph I estimate will be similar to a weighted average of the underlying networks. See Supplemental Appendix S1 for more details.

reveal how much j would be willing to pay. To capture this asymmetry, I take the communication graph to be a fixed, directed network, and present results under different assumptions of the value of incoming calls. I assume the utility obtained from a contact is independent of the state of other contacts on the network.

Adoption. In Rwanda as in many developing countries it would be difficult to enforce service contracts at scale, so nearly all accounts are prepaid. Joining the network entails opening an account, which was easy and cheap (about $1), and investing in a handset, which was expensive (offered at retail price). Most handsets were mainstream, imported models, with little differentiation in terms of features. Handsets could be purchased from the operator or independent sellers; local price trends are consistent with global trends. I treat the handset market as perfectly competitive and handset prices as exogenous. Individuals plan ahead when considering adoption: when asked in 2007, 89% of individuals without phones planned to purchase a phone in the future (Stork and Stork, 2008). I model adoption as an optimal stopping problem, where individuals decide how long to delay adoption.

Additional Simplifications. I make a number of simplifications for tractability and due to data limitations. I focus on voice calls between the primary operator's subscribers, and do not explicitly model utility from SMS, missed calls, international calls, and calls from payphones.[10] I ignore the second operator, which had less than 12% of the market during this period, and the fixed network, which had roughly 1% of the market. Any value these omissions provide to the primary operator's subscribers will be captured in a residual when I estimate the adoption decision. Mobile money was not available during this period, and mobile internet availability was negligible.

[10]From the data it is not possible to match the sender and receiver of a given SMS. Though important in other contexts, in Rwanda text messaging or SMS was high priced ($0.10 per message) and represented less than 13% of revenue and 16% of transactions. Only calls that are answered incur a charge; so subscribers may communicate simple information with missed calls (Donner, 2007). But 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).

5. Model

This section introduces a model of handset adoption and usage. Since the utility of owning a phone is derived from making calls, I begin in reverse, with usage.

Let $\bar{G}$ be the communication graph of Rwanda (a directed social network), with N nodes representing all individuals in the country. A directed link $ij \in \bar{G}$ indicates that i would have a potential desire to call j via phone. I assume links are fixed.

Let $S_t \subseteq N$ be the set of individuals with phones in month t. Since I observe only individuals who adopt phones by the end of my data T, I observe the set $S_T \subseteq N$. There may be other individuals who would have adopted if conditions were more favorable, so I will tend to underestimate the effects of counterfactuals that speed up adoption (discussed further in the simulation section). Since I only observe a link if a call was placed, I observe the subgraph, $G^T \subseteq \bar{G}$, where $ij \in G^T$ if i has called j by period T.[11] As shorthand, define $G_i = \left\{ j \,\middle|\, ij \in G^T \right\}$ as i's set of contacts.

Calling Decision. In each period t that he has a phone, individual i can call any contact j that currently subscribes, $j \in G_i \cap S_t$. i draws a communication shock $\epsilon_{ijt} \overset{iid}{\sim} F_{ij}$ representing a desire to call j; this desire might be high after an event or while coordinating a meeting, or low if there is little information to share (its link-specific distribution will be specified later). Given the shock, i chooses a total duration $d_{ijt} \geq 0$ for that month, earning utility:

$$u_{ijt} = \max_{d \geq 0} \left[\frac{1}{\beta_{cost}} v_{ij}(d, \epsilon_{ijt}) - c_{ijt} d \right]$$

where $v(d, \epsilon)$ represents the benefit of making calls of total duration of d, β_{cost} converts between utils and money, and c_{ijt} represents the per-second cost.

I model the benefit of making calls as:

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

where the first term represents a linear benefit; $\gamma > 1$ controls how quickly marginal returns decline, and α controls the intercept of marginal utility, and thus the fraction of months for which no call is placed. (This functional form was chosen to satisfy 8 reasonable properties for the utility from telephone calls; see Appendix A.)

[11] This will miss any links between subscribers where there is a latent desire to communicate but no call has been placed by T ($G^T \subseteq \bar{G}^T$). See Supplemental Appendix S16.

The marginal cost is:

$$c_{ijt} = p_t + \beta_{coverage}\phi_{it}\phi_{jt}$$

where p_t is the per-second calling price (including any tax), and $\beta_{coverage}\phi_{it}\phi_{jt}$ represents a hassle cost when the caller or receiver have imperfect coverage. An individual's coverage $\phi_{it} \in [0, 1]$ is derived from the fraction of the area surrounding his most used locations receiving cellular coverage in month t.

The benefit of an additional second of duration across a link is decreasing, so i will call j until the marginal benefit equals the marginal cost, at duration:

$$d(\epsilon, p_t, \phi_{it}, \phi_{jt}) = [\epsilon\,(1 - \beta_{cost}\,(p_t + \beta_{coverage}\phi_{it}\phi_{jt})) - \alpha]^{\frac{1}{\gamma-1}} \tag{2}$$

which increases with the desire to communicate (ϵ) and decreases with cost. If the desire to communicate is not strong enough, i does not call: $d_{ijt} = 0$ when $\epsilon_{ijt} \leq \underline{\epsilon}_{ijt} := \frac{1+\alpha}{1-\beta_{cost}\left(p_t+\beta_{coverage}\phi_{it}\phi_{jt}\right)}$.

Then, calls from i to j in period t have expected duration:

$$E_\epsilon d_{ij}(p_t, \boldsymbol{\phi}_t) = \int_{\underline{\epsilon}_{ijt}}^{\infty} d(\epsilon, p_t, \boldsymbol{\phi}_t) \cdot dF_{ij}(\epsilon) \tag{3}$$

and provide expected utility:

$$E_\epsilon u_{ij}(p_t, \boldsymbol{\phi}_t) = \int_{\underline{\epsilon}_{ijt}}^{\infty} \left[d(\epsilon, p_t, \boldsymbol{\phi}_t) \cdot \left(\frac{1}{\beta_{cost}} \left(1 - \frac{\alpha}{\epsilon}\right) - p_t - \beta_{coverage}\phi_{it}\phi_{jt} \right) - \frac{1}{\beta_{cost}\epsilon} \frac{d(\epsilon, p_t, \boldsymbol{\phi}_t)^\gamma}{\gamma} \right] dF_{ij}(\epsilon) \tag{4}$$

where $\boldsymbol{\phi}_t$ is the vector of coverage for all individuals. If they are valued, incoming calls provide utility:

$$E_\epsilon \tilde{u}_{ji}(p_t, \boldsymbol{\phi}_t) = \int_{\underline{\epsilon}_{jit}}^{\infty} \left[d(\epsilon, p_t, \boldsymbol{\phi}_t) \cdot \left(\frac{1}{\beta_{cost}} \left(1 - \frac{\alpha}{\epsilon}\right) - \beta_{coverage}\phi_{it}\phi_{jt} \right) - \frac{1}{\beta_{cost}\epsilon} \frac{d(\epsilon, p_t, \boldsymbol{\phi}_t)^\gamma}{\gamma} \right] dF_{ji}(\epsilon)$$

Altogether, each month i is on the network, he receives expected utility from each contact who is also on the network:

$$E_\epsilon u_{it}(p_t, \boldsymbol{\phi}_t, \boldsymbol{x}_{G_i}) = \sum_{j \in G_i \text{ and } x_j \leq t} E_\epsilon u_{ij}(p_t, \boldsymbol{\phi}_t) + w \cdot E_\epsilon \tilde{u}_{ji}(p_t, \boldsymbol{\phi}_t)$$

where x_j represents j's adoption time and $w \in [0, 1]$ specifies whether recipients value incoming calls. Each month that i is not on the network he receives utility zero.

Adoption Decision. Conditional on the adoption decisions of others, an individual's adoption decision is an optimal stopping problem. At period t, i knows the current price of a handset, $p_t^{handset}$ (including any tax). He believes that in period $x > t$, the handset price will be exactly $E_t p_x^{handset}$, call price $E_t p_x$, and coverage $E_t \boldsymbol{\phi}_x$. (Beliefs are assumed to be point masses completely described by their expectation.) He believes that his contacts will adopt at times $E_t \boldsymbol{x}_{G_i}$, and expects the utility of adopting at time x to be:

$$E_t U_i^x(E_t \boldsymbol{x}_{G_i}) = \delta^x \left[\sum_{s \geq x}^{\infty} \delta^{s-x} E_\epsilon u_{is}(E_t p_s, E_t \boldsymbol{\phi}_s, E_t \boldsymbol{x}_{G_i}) - E_t p_x^{handset} + \eta_i \right] \tag{5}$$

i adopts at the first month x_i where he expects adopting immediately to be more attractive than waiting, given his belief about when his contacts will adopt:

$$\min_{x_i} s.t. \left[U_i^{x_i}(E_{x_i} \boldsymbol{x}_{G_i}) \geq \max_{s > x_i} E_{x_i} U_i^s(E_{x_i} \boldsymbol{x}_{G_i}) \right] \tag{6}$$

An individual's type η_i captures any residual heterogeneity explaining why i adopted at x_i. It may include optimization error, or actual utility or costs. I do not restrict the distribution of η_i (specifically, it need not be mean zero, and can be arbitrarily correlated across network neighbors), but to make simulation tractable I do require that each individual's type is constant over time and across counterfactuals.[12]

Network Adoption Equilibrium. Initial adopters (S_0) are held fixed.[13] Each other individual i decides on an adoption time $x_i \in [1, ..., \bar{T}]$, for some $\bar{T} \geq T$. The number of potential states of the network is large ($2^{|S \setminus S_0|} > 2^{1{,}000{,}000}$); I maintain tractability with the following definition:

An **equilibrium** Γ is defined by adoption dates $\boldsymbol{x} = [x_i]_{i \in S}$ such that each individual $i \in S \setminus S_0$ adopts optimally according to Equation 6, with beliefs consistent with when his contacts adopt ($E_{x_i} \boldsymbol{x}_{G_i} = \boldsymbol{x}_{G_i}$).

This definition corresponds with a Nash equilibrium of the game where each individual simultaneously announces their adoption date x_i at the beginning of time

[12]The assumption that type is constant would be violated if, for example, a handset provided status value that changed over time, or a person preferred to purchase a handset in a certain month because he was flush with cash or had more calling needs. I expect any such changes to be dwarfed by the large changes in fundamentals over this period. In Monte Carlo simulations in the Supplemental Appendix, I find that the model performs well even in the presence of small idiosyncratic shocks, and tends to attenuate results for large shocks.

[13]Since I infer adoption from transactions, I assume subscribers with transactions between January and March of 2005 are initial adopters. I also hold fixed the adoption of individuals whose η_i I cannot back out (see Estimation).

(a complete information static game). This implies that individuals do not anticipate how later adopters will respond to their actions, because later adopters may not condition their strategy on actions in prior periods.[14] Despite its simplicity, this definition allows for rich behavior: a perturbation of utility that causes one individual to change their adoption date can shift the equilibrium, inducing ripple effects through potentially the entire network.

I also formalize expectations about the future. In addition to correctly forecasting the dates their contacts adopt, individuals also correctly forecast call prices ($E_t p_x = p_x$), and coverage ($E_t \boldsymbol{\phi}_x = \boldsymbol{\phi}_x$).[15] However, because a handset becomes sunk at the time of purchase, forecasts of future prices can sway the adoption decision. I assume that at each period t, individuals learn the current handset price and expect the price in future periods to decline at an exponential rate consistent with the overall decline over this period:

$$E_t p_x^{handset} = \omega^{x-t} p_t^{handset}$$

for $\omega = \left(\frac{p_T^{handset}}{p_0^{handset}}\right)^{\frac{1}{T}}$.[16]

If at the point of adoption, an individual forecasts utility differently than specified here, the error will be captured in his type η_i, as long as the error is fixed across time and counterfactuals. To assess the importance of forward looking behavior, in the Supplemental Appendix I also estimate and simulate results under a myopic model where individuals do not consider the future in their adoption decision.

---

[14]This results in an open loop equilibrium; see for example Fudenberg and Levine (1988).

[15]Coverage in principle could be forecasted: the operator was forced to provide a 5-year coverage rollout plan to the regulator, and was fined for a deviation.

[16]Note that this definition introduces a slight inconsistency: when i decides whether to adopt in period t, he will not know future handset prices, but will know the adoption dates of his future contacts, whose decisions will have incorporated future handset prices. I tolerate this inconsistency for the sake of tractability.

6. Estimation

Individuals choose when to adopt a mobile phone and, if they adopt, how to use the phone. The usage decision reveals the value of each connection. The adoption decision reveals any residual factors affecting adoption (individual types η_i).

Identification. Traditional approaches towards network goods estimate the value of each connection indirectly, based on correlations in adoption. For example, consider individual i who has one link, does not consider the future ($\delta = 0$), and is deciding whether to adopt, $a_i \in \{0, 1\}$. i will adopt if the value exceeds the cost:

$$a_i = I\left(\theta_{ij} a_j + \eta_i \geq p_t^{handset}\right)$$

where θ_{ij} is the value of the link if j also adopts. It is difficult to estimate θ_{ij} from correlations in adoption: each individual's adoption depends on the other, as well as any unobserved shocks, which are likely to be correlated (the reflection problem, Manski 1993). Approaches that instrument for adoption tend to rely on very particular variation, and yield crude measures of value.[17]

Rather than inferring θ_{ij} from correlated adoption, I measure it directly. A link provides value because it enables calls:

$$\theta_{ij} = E_\epsilon u_{ij}(p_t, \boldsymbol{\phi}_t)$$

I identify a link's value by how its usage changes in response to changes in the cost of communicating. The value of all links together represent the value of the network.[18]

My approach requires that the latent desire to communicate (ϵ_{ijt}) is uncorrelated with costs (p_t and $\phi_{it}\phi_{jt}$, which both improve over time). As the network grows, the composition of subscribers changes, and the operator may adjust prices and coverage in response. I absorb compositional changes by using only within-link variation to estimate the response of usage to costs. My identification assumption implies that the value of communicating with a particular contact does not otherwise trend over time, or depend on who else has adopted. I test this assumption by analyzing

[17]For example, Tucker (2008) identifies the value of a videoconferencing system using variation in television watching partly driven by the World Cup. Instrumental variable approaches do not capture rich heterogeneity, or account for how the cost of using a link affects its value.

[18]This approach has parallels with Ryan and Tucker (2012), who analyze adoption of a videoconferencing system from its use. But in that context, individuals face no cost of use or adoption.

changes in calling patterns across links; results are consistent with these factors being negligible.[19] Apart from these restrictions, communication shocks can be arbitrarily correlated between any links in the network.

After the network portion of utility (θ_{ij}) is estimated, it is straightforward to back out any residual heterogeneity affecting adoption, η_i. These types may be arbitrarily correlated between nodes, but are fixed over time and across counterfactuals.

Calling Decision. I use data on phone calls to estimate the country's latent communication graph (the call shock distributions F_{ij}), the shape of the utility function (γ and α), and how usage responds to cost (β_{cost} and $\beta_{coverage}$). I use maximum likelihood.

Estimation Procedure. First, I specify the distribution for call shocks $\epsilon_{ijt} \overset{iid}{\sim} F_{ij}$. 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 $1 - q_i$:

$$F_{ij}\left[\epsilon\right] = (1 - q_i) + q_i \Phi\left(\frac{\ln\left(\epsilon\right) - \mu_{ij}}{\sigma_i}\right)$$

where $\Phi(\cdot)$ represents the standard normal CDF.[20] In each period t, for each link between subscribers, I observe a duration $d_{ijt} \geq 0$. Equation 2 recovers the underlying call shock ϵ:

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

given coverage under the baseline rollout plan $\boldsymbol{\phi}_t(\mathbf{z}_0)$. If the call shock was not high enough to place a call, the month will have no call ($d_{ijt} = 0$), with likelihood $F_{ij}\left[\epsilon\left(1 \text{ second}\left|\cdot\right.\right)\right]$. The likelihood of calls of total duration d_{ijt} is $F_{ij}\left[\epsilon\left(d_{ijt} + 1\left|\cdot\right.\right)\right] - F_{ij}\left[\epsilon\left(d_{ijt}\left|\cdot\right.\right)\right]$.

[19] I evaluate whether the duration of calls across a link changes with the time since an individual adopted, or as more of the sender's and receiver's contacts join the network, after controlling for cost. For the median subscriber, the change in duration associated with either the change in time and contacts on the network is less than 5% of the change associated with the changes in prices and coverage over this time period. See Supplemental Appendix S5.

[20] Two parameters control the likelihood of a month without a call: α controls the part that depends on the cost of the call, and q_i controls the part that does not.

These are combined into the log-likelihood function:

$$\ln L(\Theta) = \sum_i \sum_t \sum_{j \in S_t \cap G_i} \quad 1_{\{\text{call placed}_{ijt}\}} \ln\left(F_{ij}\left[\epsilon\left(d_{ijt}+1\,|p_t,\phi_{it},\phi_{jt}\right)\right] - F_{ij}\left[\epsilon\left(d_{ijt}\,|p_t,\phi_{it},\phi_{jt}\right)\right]\right) + \left[1 - 1_{\{\text{call placed}_{ijt}\}}\right] \ln F_{ij}\left[\epsilon\left(1 \text{ second}\,|p_t,\phi_{it},\phi_{jt}\right)\right] \tag{7}$$

The full sample has 1,525,061 nodes, 414.5 million links, and a total of 15 billion link-month duration observations. The calling decision has 7 types of parameters. I assume that the shape and sensitivity parameters are common to all links (γ, α, β_{cost}, $\beta_{coverage}$). I allow the parameter scaling the shock distribution (σ_i), and the probability of no call at any price ($1 - q_i$) to vary at the individual level. I allow the shock distribution to be shifted at the link level, with structure:

$$\mu_{ij} = \mu_i + \mu_{\max(x_i,x_j),\overline{\phi_{it}\phi_{jt}}}$$

which includes an individual mean term μ_i, and a cost fixed effect for each combination of link adoption date ($\max\{x_i, x_j\}$) and average coverage ($\overline{\phi_{it}\phi_{jt}}$), discretized to 519 combinations. This term ensures that price and coverage sensitivity are identified off of within-link changes in calling.[21]

I estimate these parameters in two steps. First, I jointly estimate the common parameters, cost fixed effects, and distribution parameters for a random 0.5% subset of nodes and their full set of links.[22] Then, I impose the common parameters and cost fixed effects to estimate the remaining distribution parameters (q_i, σ_i, μ_i; 4.6 million altogether). Conditional on the common parameters and cost fixed effects, an individual's distribution parameters affect only his own likelihood, so this last step is computationally much less demanding than performing a full joint estimation.

I use the estimated model to compute the expected duration and utility along each link using Monte Carlo integration of Equations 3 and 4.

[21]To see how this formulation addresses the selection issue, note that the numerator inside the standard normal CDF in $F_{ij}\left[\epsilon\left(d\right)\right]$ can be written: $\ln\left(d^{\gamma-1}+\alpha\right) - \ln\left(1 - \beta_{cost}\left(p_t + \beta_{coverage}\phi_{it}\phi_{jt}\right)\right) - \mu_{ij}$. Consider the decomposition $\mu_{ij} = \mu_i + e_{ij}$ where μ_{ij} is the true link mean and e_{ij} is an error. If the estimated specification only includes μ_i, and if e_{ij} is correlated with p_t or $\phi_{it}\phi_{jt}$, then estimates of β_{cost} and $\beta_{coverage}$ may be biased. While each link faces the same price series over time, links that join the network later face lower prices on average: the price path faced by a link is summarized by the link's adoption date. Each link faces a different coverage path; I approximate these with the average joint coverage using 10 bins.

[22]The approximately 0.5% sample is 8,000 nodes with 1,317,539 links and 39 million link-month observations (2.6 million with calls).

Adoption Decision. Consider the utility i would have received had he adopted a different month. At time x_i, i bought a handset rather than waiting K months. Holding fixed the actions of others, Equation 6 implies $U_i^{x_i}(\boldsymbol{x}_{G_i}) \geq E_{x_i} U_i^{x_i+K}(\boldsymbol{x}_{G_i})$. This implies that the expected utility of being on the network during the following K months must have exceeded the expected drop in handset prices:[23]

$$\sum_{k=0}^{K-1} \delta^k E_\epsilon u_{ix_i+k}(p_{x_i+k}, \boldsymbol{\phi}_{x_i+k}, \boldsymbol{x}_{G_i}) + (1-\delta^K)\eta_i \geq p_{x_i}^{handset} - \delta^K E_{x_i} p_{x_i+K}^{handset} \tag{8}$$

Similarly, i could have purchased a handset earlier. At time $x_i - K$, i chose to wait K months, implying $U_i^{x_i-K}(\boldsymbol{x}_{G_i}) \leq E_{x_i-K} U_i^{x_i}(\boldsymbol{x}_{G_i})$. Thus the expected utility from the K months prior to purchase must have been less than the expected drop in handset prices:

$$\sum_{k=1}^{K} \delta^{K-k} E_\epsilon u_{ix_i-k}(p_{x_i-k}, \boldsymbol{\phi}_{x_i-k}, \boldsymbol{x}_{G_i}) + (1-\delta^K)\eta_i \leq p_{x_i-K}^{handset} - \delta^K E_{x_i-K} p_{x_i}^{handset} \tag{9}$$

These inequalities imply bounds for each individual's type: $\underline{\eta}_i \leq \eta_i \leq \bar{\eta}_i$, where:

$$\begin{aligned}
\underline{\eta}_i &= \frac{1}{1-\delta^K}\left[p_{x_i}^{handset} - \delta^K E_{x_i} p_{x_i+K}^{handset} - \sum_{k=0}^{K-1} \delta^k E_\epsilon u_{i,x_i+k}(p_{x_i+k}, \boldsymbol{\phi}_{x_i+k}, \boldsymbol{x}_{Gi}) \right] \\
\bar{\eta}_i &= \frac{1}{1-\delta^K}\left[p_{x_i-K}^{handset} - \delta^K E_{x_i-K} p_{x_i}^{handset} - \sum_{k=1}^{K} \delta^{K-k} E_\epsilon u_{i,x_i-k}(p_{x_i-k}, \boldsymbol{\phi}_{x_i-k}, \boldsymbol{x}_{Gi}) \right]
\end{aligned} \tag{10}$$

I set $K = 2$ months.[24] Note that the future after $x_i + K$ cancels out of these expressions, so that results are not sensitive to the evolution of the network beyond that point. These conditions are necessary for equilibrium and are valid in the presence of multiple equilibria. During months that extra fees were charged, I incorporate the fee schedule.[25] I set the discount factor to the inverse of the average real interest rate

[23]The model implies that i correctly forecasts the first K months of utility and his expectation of the continuation flow does not change between x_i and $x_i + K$. Both options provide the same continuation flow of utility after $x_i + K$, so they differ only in the utility provided in the first K months. Available top up amounts also changed during this period, which I do not model.

[24]I select K to balance two forces: lower values produce tighter bounds; higher ones smooth any shocks around their adoption date that are unaccounted for.

[25]Before June 2007, subscribers needed to add roughly \$4.53 in credit per month to keep their account open; I factor this in as a hassle cost. Actually opening an account entails purchasing a SIM card, which cost roughly \$1 itself plus the cost of an initial top up. See Supplemental Appendix S11 for more information.

in Rwanda over this period: $\delta = (\frac{1}{1.07})^{1/12} \sim 0.9945$ (source: World Bank). I am able to recover η_i's for 0.8m nodes adopting between $x_i \in [K, T-K]$ [26]

Results. Parameter estimates are reported in Table 2. The calling decision implies $\beta_{cost} = 0.200$, which corresponds with an average price elasticity of -0.67; and $\beta_{coverage} = -3.85$, which corresponds with an average elasticity of 0.47 for either sender or receiver's coverage. Predicted durations are highly correlated with observed durations (by month: 0.95, by node: 0.90). The second panel shows the distribution of the parameters of the communication graph. The third panel shows the bounds on individual types $[\underline{\eta}_i, \bar{\eta}_i]$ implied by the adoption model, assuming individuals do not value incoming calls ($w = 0$). The median individual's adoption is consistent with expecting to receive \$0.78-1.02 of additional value from a handset each month, beyond that represented by the call model (shown in Table 2). However, types are heterogeneous: the 25th percentile expects to receive \$0.12-0.63 less value, and the 75th percentile \$1.42-1.55 more.

Validation of Utility. As a robustness check, I test whether the value implied by calls corresponds with the value implied by adoption. I form Equations 8 and 9 into moment inequalities, using instrumental variables that shift adoption based on the cost of providing coverage to different areas due to Rwanda's hilly geography, in a manner similar to Yanagizawa-Drott (2014), as well as in the fraction of contacts who join in response to a government adoption subsidy program. Results suggest these valuations of calls and adoption are consistent under $w = 0$, but under $w = 1$, the utility of calls is double counted. (See Appendix B.) I proceed with $w = 0$ as a base specification and evaluate the case where $w = 1$ for robustness in the Supplemental Appendix.

[26] I do not back out bounds for those receiving a rural handset subsidy in 2008 (for whom it is difficult to value the purchase), and whose activation does not coincide with the adoption of a new handset (altogether these account for 5% of last period durations). In simulations, I compute changes in the call model for all nodes, and hold the adoption of these fixed; doing so will tend to attenuate the results of policy counterfactuals.

TABLE 2. Parameter Estimates

		Parameter		**Estimate**		**Standard Error**			
Call Model	Common Parameters	γ		2.289		0.002			
		α		97.897		0.471			
		β_{cost}		0.200		0.0004			
		$\beta_{coverage}$		-3.845		0.037			
		Quantile:		**5%**	**25%**	**50%**	**75%**	**95%**	**Number**
	Fixed Effect Parameters	$\mu_{\max(x_i,x_j),\overline{\phi_{it}\phi_{jt}}}$		-1.71	-1.45	-1.12	-0.65	0.00	519
		SE		0.31	0.31	0.31	0.32	0.32	
	Node Parameters	μ_i		1.95	3.01	3.66	4.95	6.83	1.5m
		SE		0.32	0.32	0.32	0.33	0.33	
		σ_i		0.80	1.30	1.62	1.98	2.58	1.5m
		SE		0.04	0.02	0.02	0.03	0.07	
		q_i		0.06	0.27	1.00	1.00	1.00	1.5m
		SE		0.00	0.01	0.02	0.00	0.00	
									$N_{link-months} = 15$ billion
Adoption Model		$\bar{\eta}_i \cdot (1-\delta)$	\$	-4.57	-0.12	1.02	1.55	4.19	0.8m
		SE		0.63	0.16	0.08	0.00	0.04	
		$\underline{\eta}_i \cdot (1-\delta)$	\$	-5.89	-0.63	0.78	1.42	3.83	0.8m
		SE		0.69	0.19	0.08	0.01	0.03	
									$N_{adoptions} = 0.8$ million

Call model parameters are estimated in a two step maximum likelihood procedure, based on Equation 7. For noncommon parameters I report the quantiles of estimates and the standard errors of these quantiles for a joint estimation subsample, based on bootstrap draws using the empirical covariance matrix. See Supplemental Appendix for comparative statics that interpret these estimates. Individual types are backed out from Equation 10, assuming individuals only receive value from outgoing calls ($w = 0$). I report types multiplied by $1 - \delta$ to obtain monthly flows.

7. Simulation of Network Good Adoption

To identify equilibria, I use an iterated best response algorithm. I start with a candidate adoption path $\boldsymbol{x}^0$ and sequentially allow each individual to optimize their adoption date, conditional on the adoption dates of others, until the path converges.[27]

An equilibrium $\Gamma(\boldsymbol{\eta})$ is a function of the vector of individual types, $\boldsymbol{\eta} = [\eta_i]_i$. There tend to be multiple equilibria for two reasons: each individual's type is backed out as a set, and individuals may coordinate on being optimistic or pessimistic about others' adoption.

I derive bounds for the entire set of equilibria by exploiting its lattice structure: i's optimal adoption date x_i is weakly monotonic in both η_i and x_j.[28] The equilibrium with lowest possible adoption, $\underline{\Gamma}$, can be identified by setting each individual's type η_i to its lower bound, and starting with a pessimistic candidate adoption path: $\boldsymbol{x}^0 = \bar{T}$ (initially individuals expect everyone else to completely delay adoption). The highest possible equilibrium, $\bar{\Gamma}$, can be identified by setting each individual's type to its upper bound and starting with an optimistic candidate adoption path: $\boldsymbol{x}^0 = 0$ (initially individuals expect everyone else to adopt immediately).[29]

My data ends at May 2009 (T); although I report outcomes only up until T, the end date affects the analysis in two ways. First, counterfactuals that make adoption less attractive may induce individuals in my data ($i \in S^T$) to delay adoption beyond T. To account for this, I allow individuals to adopt up until $\bar{T} = T + 36$, three years beyond the end of my data, extrapolating utility for the additional months. (This extrapolation serves only to determine whether an individual adopts before or after T.)[30] Second, counterfactuals that make adoption more attractive could induce individuals not in my data ($i \in N \backslash S^T$) to adopt before T. This could potentially happen in the second counterfactual on taxes; in the Supplemental Appendix I do a

[27] See Appendix C.

[28] A higher type η_i weakly decreases i's optimal adoption date, and a decrease in i's adoption date x_i weakly decreases j's optimal adoption date. This follows from the lattice structure of $\boldsymbol{x}$ and because $U^{x_i}(\eta_i, \boldsymbol{x}_{-i})$ has increasing differences in x_i and x_j, or is supermodular in $\boldsymbol{x}$; see Topkis (1978) and Milgrom and Shannon (1994).

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

[30] I assume that after T, the utility provided to each node grows at the same rate as the total number of subscribers (as reported by the regulator). See Supplemental Appendix S11 for more details.

subsample analysis which suggests that my results are likely to understate the true effects.[31]

Outcomes. For each equilibrium Γ, I compute the expected net present value of consumer surplus through T, as of January 2005:

$$U_{net}^{\Gamma} = \sum_{i \in S \text{ and } x_i \leq T} \left[\sum_{t \geq x_i}^{T} \delta^t E_\epsilon u_{it}(p_t, \boldsymbol{\phi}_t, \boldsymbol{x}_{G_i}) - \delta^{x_i} p_{x_i}^{handset} - \delta^T p_T^{handset} \right]$$

which is net of calling, handset, and hassle costs. I compute handset costs assuming each handset is provided by a competitive market at marginal cost, and is sold back at the end of the data at the prevailing price. I omit the type η_i that enters the individual's adoption decision, because this term may pick up errors that do not represent the utility individuals receive.

The government earns expected revenue from taxes on handsets ($\tau_{it}^{handset}$) and usage (τ_{it}^{usage}):

$$R_G^{\Gamma} = \sum_{i \in S \text{ and } x_i \leq T} \left[\delta^{x_i} \tau_{ix_i}^{handset} p_{x_i}^{handset} + \sum_{t \geq x_i}^{T} \delta^t \tau_{it}^{usage} p_t \sum_{j \in G_i \cap S_t} E_\epsilon d_{ij}(p_t, \boldsymbol{\phi}_t) \right]$$

And the firm earns expected revenue from calls:

$$R_F^{\Gamma} = \sum_{i \in S \text{ and } x_i \leq T} \sum_{t \geq x_i}^{T} \delta^t (1 - \tau_{it}^{usage}) p_t \sum_{j \in G_i \cap S_t} E_\epsilon d_{ij}(p_t, \boldsymbol{\phi}_t)$$

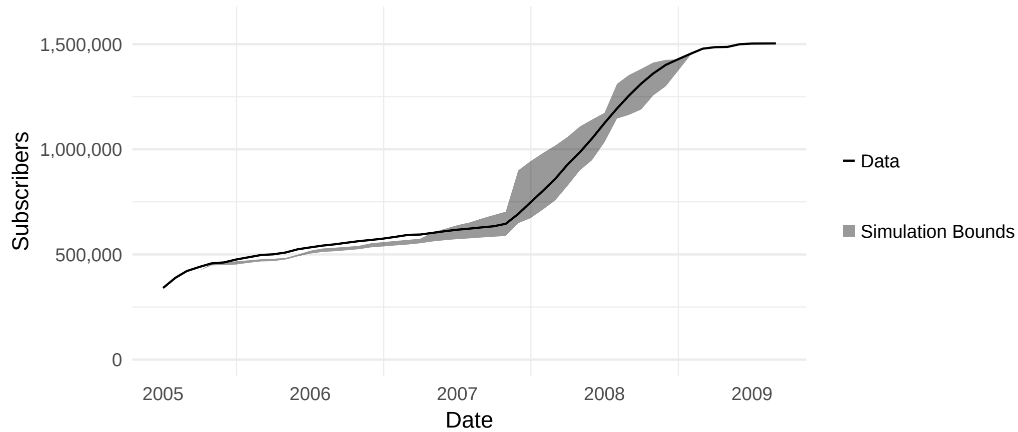
In simulations, I report each outcome Y at the high equilibrium ($Y^{\bar{\Gamma}}$) and low equilibrium ($Y^{\underline{\Gamma}}$).[32] I measure policy impacts by reporting the change in the lowest and highest equilibria ($Y^{\bar{\Gamma}} - Y_{baseline}^{\bar{\Gamma}}$ and $Y^{\underline{\Gamma}} - Y_{baseline}^{\underline{\Gamma}}$).[33]

[31] Additionally, reducing the cost of communication (e.g., by lowering usage taxes) could induce calls between individuals who hadn't called each other within my data (making additional links active beyond those I observe).

[32] Because there is a monotonic relationship between adoption date and utility, these outcomes bound all possible equilibrium outcomes for call utility. They also bound all possible equilibrium outcomes for firm revenue R_F when usage tax is constant over time. The high and low equilibrium outcomes of U_{net} and R_G may not bound all possible equilibrium outcomes, though I expect deviations to be minor.

[33] A more natural measure of policy impact would be bounds on the changes in revenue and utility across the range of equilibria; however, this measure is computationally prohibitive because adoption decisions are interlinked. Monte Carlo simulations in the Supplemental Appendix provide a test of this approach.

Figure 2. Simulation Fit



Baseline Simulation Results. The simulation matches the general trend of the data, shown in Figure 2. In the simulated equilibrium, the mobile phone system provides a total benefit between $474-516m over the 4.5 years I observe, discounted to 2005 dollars. These benefits are split among the operator (33%), the government (13%), and consumers (54%, net of handset, calling, and hassle costs).[34] While the welfare gains of mobile phones has been documented in specific sectors (Jensen, 2007; Aker, 2010; Jack and Suri, 2014), to my knowledge these are the first micro-identified estimates of the total welfare generated by a developing country mobile phone network.

8. Application: The Provision of Service to Rural Areas

For communication services, a key question for regulators is whether—and if so, how—to ensure service to poor and remote communities. Governments use a wide variety of policies, including tax-and-transfer schemes, service obligations, and universal service funds (GSMA, 2013).

A social planner would expand coverage until the point where building any marginal set of towers would not improve welfare. Firms may stop building before reaching this point: they are likely to internalize only a fraction of the benefits of expanding

[34]The government also collects a general 30% tax on firm revenues after deductions; since I do not observe deductions I do not compute this explicitly; some additional firm revenues should be transferred to the government.

the network. Whether it is optimal for a government to regulate service to remote areas depends crucially on both the shape of social and private benefits. Both are difficult to measure due to spillovers induced by geographical interconnectedness and network effects. I use my method to measure both objects for an expansion in rural service in Rwanda induced by a coverage requirement.

Impact of Rural Expansion in Rwanda. In Rwanda, the regulator required a rollout plan culminating in near-complete coverage. I do not attempt to isolate the impact of specific obligations, as they are likely to have been anticipated by the operator and formed in ongoing discussions. I would ideally compare the revenue and consumer surplus generated under the actual rollout to that generated by the rollout that maximizes profits in absence of regulation. Determining this profit maximizing rollout is computationally infeasible. Instead, I simulate a suggestive counterfactual, where the operator does not build marginal, unprofitable towers.

I compute this counterfactual by first ranking each tower z by a proxy of how desirable it was for the firm to build, and then finding a desirability cutoff below which tower construction was unprofitable. I rank each constructed tower z by the empirical revenue of the transactions that were transmitted through it ($\bar{R}_z$: 'average direct baseline revenue'). The distribution of monthly revenue by tower is shown in Figure 3a. This provides a rough gauge of the desirability of a tower, but does not capture the causal impact of building a tower on revenue: it omits substitution between towers and the effect of coverage on adoption. I determine the causal impact using my simulation method.

I propose a cutoff, starting with the $n = 10$ lowest revenue rural towers, and compute the progression of coverage omitting this set of rural towers ($\mathbf{z}_{(n)} = \{z \in \mathbf{z}_0 | \bar{R}_z > \bar{R}_{(n)}$ or z is urban$\}$, where $\bar{R}_{(n)}$ is the nth order statistic of $\bar{R}$ for rural towers). Given each counterfactual rollout plan, I compute each individual's time series of coverage $\phi_{it}(\mathbf{z}_{(n)})$, the resulting link utilities and durations, and then simulate a new 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. To evaluate whether a deviation is profitable, I compute the annualized cost of building and operating each tower building plan $C_F(\mathbf{z})$ using estimates provided by the regulator.[35]

Figure 3 shows results from simulations with $n \in \{0, 10, 20\}$. Direct baseline revenue of these towers are highlighted in Figure 3a. Full simulation results are presented in 3b. The 10 rural towers with lowest direct baseline revenue constructed between 2005 and January 2009 were unprofitable (11% of all rural towers). In contrast, the 11-20th lowest were profitable.[36] Thus, in absence of the coverage obligation I would expect the operator to not build these unprofitable rural towers, which had an annualized net present cost over this time period of \$271,356 in 2005. Figure 3c shows these towers; several cover border crossing points, for which there was an explicit coverage requirement.

I estimate the effect of the regulation by taking the difference between the baseline ($n = 0$) and the counterfactual where these unprofitable towers are not built ($n = 10$), during the sample period.[37] Table 3 presents the results for adoption months, revenue, and consumer surplus. In the rows of Table 3, 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 (allowing changes to adoption, including network effects).

The first column of Table 3 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. The expansion moves the

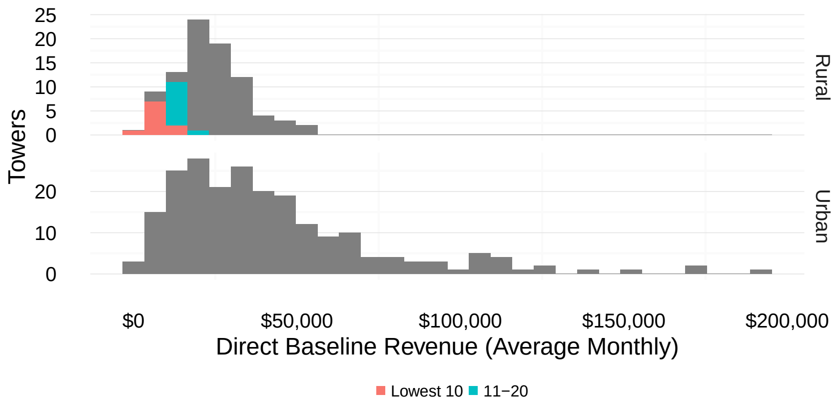
[35]Costs are given by $C_F(\mathbf{z}) = \sum_{\mathbf{z}} \sum_{s \geq t_z} \delta^s \left[K_{base} + 1_{\{z \text{ is off grid}\}} K_{off-grid}\right]$. The total annualized cost of building and operating a tower in Rwanda is \$31,165 per year, plus \$18,078 for towers that are far from the electric grid that must be powered by generators, based on financial data provided by operators to the regulator (RURA, 2011). The total cost of ownership includes operating expenses, annualized depreciation, and a 15% cost of capital. Depreciation assumes lifespans of 15 years for towers, 8 years for electric grid access, and 4 years for generators.

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

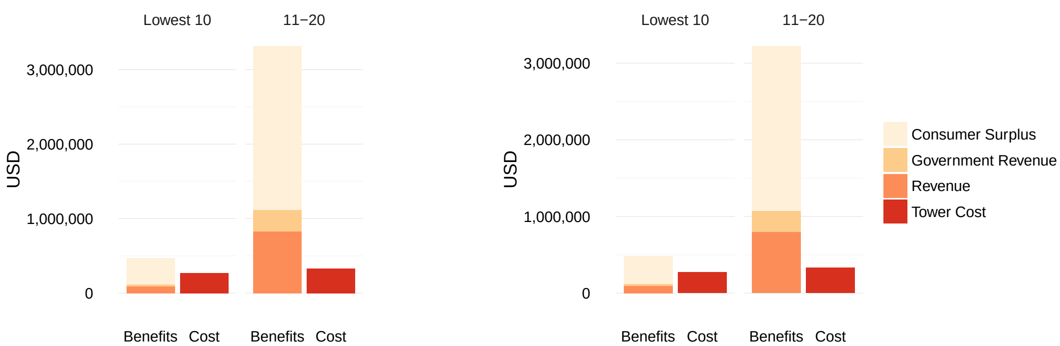
[37]Although the operator would not know the revenue generated by each tower ex-ante, they would likely have precise estimates. Note that based on the data I have I cannot compute revenues and consumer surplus beyond May 2009; the full impact could be more positive (e.g., if demand is dynamic or there is a first mover advantage) or more negative (if the unprofitable towers continue to lose money in the future).

FIGURE 3. Rural Expansion

(a) Baseline Tower Revenue

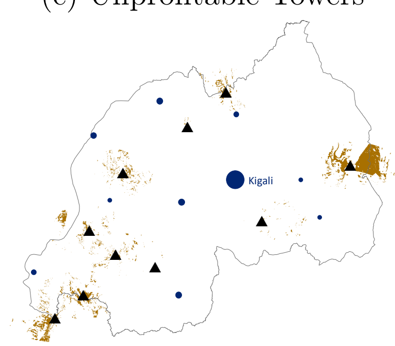


(b) Simulated Costs and Benefits of Building Low Revenue Towers



Low Equilibrium High Equilibrium

(c) Unprofitable Towers



Revenue includes domestic voice calls originating at that tower, billed by the average basket of prepaid rates, averaged over all months the tower was operational. For the counterfactual, I drop the 10 lowest revenue rural towers built during the data. In the map, cities are denoted by dots, dropped towers are denoted by triangles, and lost coverage is shaded.

Table 3. Impact of Rural Service Expansion

		All nodes	Nodes by change in coverage $> 1\%$pt	$\leq 1\%$pt
Number		1,503,675	82,523	1,421,152
Adoption Time (mean)				
Baseline with expansion	month	[24.104, 22.498]	[30.021, 28.200]	[23.760, 22.166]
Total Impact of Expansion	month	-0.005, -0.005	-0.026, -0.028	-0.004, -0.003
Revenue (total)				
Baseline with expansion	million $	[156.32, 172.51]	[2.75, 3.04]	[153.57, 169.48]
Total Impact of Expansion	million $	0.09, 0.09	0.02, 0.02	0.07, 0.07
... immediate effect on calls	million $	0.07, 0.07	0.02, 0.02	0.05, 0.05
... added effect through adoption	million $	0.02, 0.02	0.00, 0.00	0.02, 0.02
Consumer Surplus (total)				
Baseline with expansion	million $	[255.10, 275.07]	[3.19, 3.40]	[251.91, 271.67]
Total Impact of Expansion	million $	0.35, 0.36	0.08, 0.08	0.28, 0.28
... immediate effect on calls	million $	0.32, 0.33	0.07, 0.07	0.25, 0.26
... added effect through adoption	million $	0.03, 0.03	0.00, 0.00	0.03, 0.03
Government Revenue (total)				
Baseline with expansion	million $	[62.51, 68.23]	[1.84, 1.98]	[60.67, 66.25]
Total Impact of Expansion	million $	0.03, 0.03	0.01, 0.01	0.02, 0.02
... immediate effect on calls	million $	0.02, 0.02	0.01, 0.01	0.01, 0.02
... added effect through adoption	million $	0.01, 0.01	0.00, 0.00	0.01, 0.01

Baseline reported for the lower bound and upper bound estimate of the equilibrium. Impacts represent the difference in these bounds between the baseline and that with the expansion removed. Utility and revenue reported in 2005 U.S. Dollars, discounted at a rate of δ. 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. I present results split by change in coverage as of January 2009.

former's adoption forward by an average of 0.03 months, and the latter's adoption forward by 0.004 months in the lower equilibrium and 0.003 in the upper. I find:

The expansion was unprofitable for the operator (by construction). Building the towers shifted bounds on the operator's profits downward by $181,243 (lower equilibrium) and $177,677 (upper). It would have reduced profits even if the government rebated the associated tax revenue back to the firm.

Rural expansion improved welfare. Building the lowest revenue towers shifted bounds on welfare upward by $200,150 (lower equilibrium) and $213,639 (upper).

The benefits generated by the expansion was dispersed. Approximately 77% of revenue comes from, and 78% of benefits accrue to, individuals whose personal coverage was not substantially affected.

Due to this dispersion, it would have been difficult for communities to finance tower construction themselves. A common question is whether citizens would be willing to raise local taxes to finance local infrastructure improvements. But in this case most of the consumer surplus accrues to individuals in other areas. If the most affected citizens banded together to raise money for the towers, they would incur a utility loss: bounds on their consumer surplus would have declined by roughly $195,088, despite generating benefits for consumers in other locations, the operator, and the government.

In the Supplemental Appendix I perform additional checks. I evaluate two alternate models, one where recipients also receive surplus from incoming calls ($w = 1$), and a myopic model where individuals do not consider the future in their adoption decision. These yield similar results. Removal of the coverage obligation could cause the firm to reoptimize other margins, but I find limited incentives for a supply side response in calling price in a test based on firm optimization.

9. Application: Telecom Taxation in Developing Countries

Generating public revenue is a perennial challenge for developing countries, which tend to be confined to a small set of feasible instruments (Gordon and Li, 2009). However, even in countries with very little other capacity to collect revenue, the telecom sector is large, formal, and easy to tax.

Developing country governments recognize this convenient source of revenue: the mobile industry contributed an average of 7% of government revenue in sub-Saharan Africa as early as 2007 (GSMA, 2008).[38] In addition to standard taxes, governments charge spectrum license fees and specific taxes on telecom equipment, mobile handsets, and airtime. While it is clear that this emerging sector provides a public

[38] For a sample of 19 countries from which data is available.

finance opportunity for poor countries, it is unclear how to best exploit it. There is a widespread concern that countries may continue to tax telecom heavily in the short term at the expense of long term growth. The former Director of ICT at the World Bank, Mohsen Khalil, cautions: "the indirect benefits to the economy of having affordable access to telecommunications services far outweigh any short-term benefit to the budget."

Two key tax policies affecting adoption are the handset tax and tax on usage. Together with import duties on telecom equipment, these represented 66% of tax revenue from telecom in sub-Saharan African countries in 2006 (GSMA, 2008)[39]. Sub-Saharan African consumers faced an average handset tax of 31% and usage tax of 20% (in Rwanda, 48% and 23%, respectively).[40]

Many telecoms argue that high handset taxes slowed the adoption of feature phones, and are currently inhibiting the adoption of smartphones and thus the internet. Some argue that lowering handset taxes would boost adoption and lead to higher long term tax revenue, and some countries have tried eliminating handset taxes, including Kenya (eliminated 2009, reinstated 2013), Senegal (eliminated 2009), and Rwanda (eliminated 2010). But there is little evidence guiding how, and how much, to tax telecom.

This application simulates adoption under a variety of tax policies. My estimates are likely to be conservative because simulations that make adoption more attractive may induce individuals outside my data to adopt sooner (I gauge this by showing results for shorter periods in the Supplemental Appendix). I assume complete passthrough of handset taxes, which would result if handsets were offered by a competitive market with free entry. For airtime taxes, I evaluate results under the two extremes of complete and no passthrough.

Simulation results are presented in Table 4. The first row presents the baseline tax regime (48% handset tax and 23% airtime tax).[41] The following rows present the effects of altering the handset tax, and then altering the usage tax. I present the

[39]For the 15 countries from which data is available.

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

[41]I report tax rates as a proportion of the pretax amount; the model uses the equivalent posttax amount for τ.

immediate effect of the change in tax policy on individuals, without allowing changes to ripple through the network ('proximal effect') as well as the full effect after all nodes have reoptimized ('ripple effects'). The columns present the revenue accruing to the telecom and government, and the net surplus accruing to consumers in the upper and lower equilibria. I find:

Taxing a growing network imposed a substantial welfare cost. Handset taxes raised \$12m at the cost of \$22m in consumer surplus and \$14m (lower equilibrium) or \$17m (upper) in telecom revenue. This corresponds with an average welfare cost of \$2.93 (\$3.14) for each dollar of government revenue raised. This is a higher cost than estimates of marginal cost of public funds from the literature, of 1.21 for sub-Saharan Africa and 1.37 for Rwanda (Auriol and Warlters, 2012), suggesting it would have been preferable to use alternative instruments to raise these revenues. Since in this model telecoms earn no revenue from handset sales, the entire effect on telecom revenue is driven indirectly, by reduced usage.

Micro elasticities grossly understate the welfare costs of handset taxes in a growing network. Network ripple effects account for up to 63% of the effect of handset taxation on telecom revenues (\$7.59m or \$10.59m). Additionally, ripple effects generate additional government revenue and consumer surplus that would be neglected by a model that only considered individual responses (\$2.3m or \$3.2m in government revenue and \$11m or \$12m in consumer surplus). Naïve estimates based on micro elasticities that treat phones like standard goods would suggest the average cost of raising a dollar of government revenue from handset taxes would be much lower–\$1.22 in the lower equilibrium and \$1.04 in the upper. Under these estimates handset taxes would have looked more attractive than other tax instruments, as reported by Auriol and Warlters (2012).

The welfare cost of taxation may be heterogeneous over time as the network expands. I also present simulations that remove taxes for different time periods. There are two differences across time periods—the pace of potential adoption differs, and I observe the dynamic effects of early taxes over a longer time horizon. These effects combine to suggest that welfare costs during the early period of adoption were very high. The welfare cost per dollar raised for taxing handsets from 2007-2009 was \$2.43 per dollar raised (\$1.85). If taxes during that period were lifted, the welfare cost for taxing handsets from 2005-2007 would be \$5.21 (\$13.27). Similarly, the welfare cost for taxing usage, assuming complete passthrough, was \$2.56 (\$2.35) from 2007-2009, and if the tax during that period were lifted, \$4.25 (\$5.16) from 2005-2007.

Usage taxes may have imposed a similar welfare cost. The telecom may choose whether to pass changes in usage taxes through to consumers. If there were no passthrough, a change in usage tax would have directly revenue from the telecom without distortions. If there were complete passthrough, usage taxes would have caused distortions similar in magnitude to handset taxes. In that case, usage taxes would have raised more revenue, \$46m (lower equilibrium) or \$51m (upper), but at the cost of \$72m (lower) or \$75m (upper) in consumer surplus, and \$55m (lower) or \$61m (upper) in telecom revenue. This is also a substantial welfare cost: \$2.77 (or \$2.67) for each dollar of government revenue raised. The naïve estimate of the welfare cost of a usage tax is underestimated but less so than for a handset tax (\$2.12 or \$2.09).

Handset taxes impose a high cost on poorer users. I explore the distributional implications of taxes in Table 5. Although I do not observe household income or consumption, representative survey data suggest that consumers with lower airtime usage have lower consumption per capita.[42] I show revenues and consumer surplus for the entire sample, and then for the top half of users (above median average daily duration at baseline) and bottom half (below). Under the baseline tax regime, the bottom 50% of users account for only 7% of firm revenue and receive only 2% of consumer surplus, but account for 19-20% of government revenue. Since all adopters pay the fixed cost of a handset regardless of usage, poor adopters end up paying a substantial portion of the tax burden. Eliminating handset taxes would raise the surplus obtained by these consumers by 56%.

Shifting taxes from handsets to usage would have improved welfare of poorer users by at least 38%. While it may have improved welfare to eliminate these taxes and recoup government revenue with alternate instruments, I also consider

[42]Among households with mobile phones in the 2010 Rwandan EICV, those in the top half of airtime expenditure have 3.7 times more overall consumption than those in the bottom half.

Table 4. Telecom Taxation

Simulation		Revenue ($m) Telecom	Revenue ($m) Government	Consumer Surplus ($m)	Avg. Welfare Cost per Dollar of Public Funds ($)
Baseline (23% Usage Tax, 48% Handset Tax)		[156.32, 172.51]	[62.51, 68.23]	[255.10, 275.07]	
Impact of Policies					
Handset Tax: Removal	**Passthrough**				
Total Effect	Complete	14.21, 16.75	-12.21, -12.33	21.52, 21.94	2.93, 3.14
... only proximal effect	Complete	6.62, 6.16	-14.47, -15.49	11.03, 9.98	1.22, 1.04
... additional ripple effects	Complete	7.59, 10.59	2.27, 3.16	10.49, 11.95	-
By time period:					
... from 2007-2009	Complete	8.73, 7.63	-10.01, -10.94	15.56, 12.65	2.43, 1.85
... also 2005-2007	Complete	5.48, 9.12	-2.20, -1.39	5.96, 9.29	5.21, 13.27
Usage Tax: Removal	**Passthrough**				
Total Effect	Complete	55.14, 60.63	-45.82, -50.72	71.66, 74.75	2.77, 2.67
	None	46.69, 51.53	-46.69, -51.53	0.00, 0.00	1.00, 1.00
... only proximal effect	Complete	42.76, 47.13	-46.11, -50.96	55.16, 59.36	2.12, 2.09
	None	46.69, 51.53	-46.69, -51.53	0.00, 0.00	1.00, 1.00
... additional ripple effects	Complete	12.38, 13.50	0.29, 0.23	16.50, 15.39	-
	None	0.00, 0.00	0.00, 0.00	0.00, 0.00	-
By time period:					
... from 2007-2009	Complete	44.39, 47.27	-40.13, -44.98	58.25, 58.45	2.56, 2.35
... also 2005-2007	Complete	10.75, 13.35	-5.69, -5.74	13.41, 16.30	4.25, 5.16

Results in each cell reported for the lower bound and upper bound estimate of the equilibrium. Utility and revenue reported in 2005 U.S. Dollars, discounted at a rate of δ. 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.

the case where the government must earn a fixed amount of revenue from telecom. In 2010, the Rwandan government lowered handset taxes and raised usage taxes. I consider the effect of an earlier shift: eliminating handset taxes and raising usage taxes to 30% starting in 2005. If the increase in usage taxes were passed through, this would raise more government revenue ($1.79m or $3.61m) and increase the consumer surplus accruing to the bottom half of users by at least 38%, without substantial harm to the top half of users (-$0.55m in the low equilibrium or +$0.12m in the high), though

Table 5. Shifting Telecom Taxes - Distribution

Tax Regime			Revenue ($m)		Consumer
Handset	Usage	Sample Split	Telecom	Government	Surplus ($m)
Baseline					
48%	23%	All	[156.32, 172.51]	[62.51, 68.23]	[255.10, 275.07]
		Above Q50 usage	[145.45, 160.76]	[50.25, 55.31]	[249.59, 269.71]
		Below Q50 usage	[10.87, 11.75]	[12.26, 12.92]	[5.51, 5.36]
Impact of changing taxation					
0%	23%	All	14.21, 16.76	-12.20, -12.33	21.52, 21.93
		Above Q50 usage	12.91, 15.83	-3.26, -2.88	18.45, 18.93
		Below Q50 usage	1.30, 0.93	-8.94, -9.45	3.06, 3.00
0%	30%, increase passed through	All	-4.81, -3.41	1.79, 3.61	1.54, 2.15
		Above Q50 usage	-4.99, -3.19	9.63, 11.90	-0.55, 0.12
		Below Q50 usage	0.18, -0.22	-7.84, -8.29	2.09, 2.04
0%	30%, increase not passed through	All	-1.29, -0.45	3.30, 4.88	21.52, 21.93
		Above Q50 usage	-1.49, -0.22	11.13, 13.17	18.45, 18.93
		Below Q50 usage	0.19, -0.22	-7.83, -8.29	3.06, 3.00

Results in each cell reported for the lower bound and upper bound estimate of the equilibrium. Sample split by quantile of average daily usage. Utility and revenue reported in 2005 U.S. Dollars, discounted at a rate of δ. 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.

it would reduce operator revenue (by $4.81m or $3.41m). (However, the operator could actually earn more revenue by opting not to pass the tax increase on, in which case the shift in taxes would raise $3.30m ($4.88m) in government revenue, increase the consumer surplus of the bottom half of users by 56% and the top half by 7%, at the cost of a small reduction in operator revenues (-$1.29m or -$0.45m), which could be reimbursed to the firm out of the increased government revenue.) Since potential adopters who are not in my data are likely to be light users, these results likely understate the full distributional impact of a change in policy.

In the Supplemental Appendix I perform additional checks. To test the impact of observing only subscribers adopting by T, I present results for subperiods of the data, which suggest that my results may understate the welfare costs of taxation. I also evaluate two alternate models, one where recipients also receive the surplus from incoming calls ($w = 1$), and a myopic model where individuals do not consider the future in their adoption decision. In both, the welfare cost of usage taxes is similar but that of handset taxes is lower. Removing taxes could also cause a supply side response, but tests using firm optimality conditions suggest responses would not have major effects.

Mobile phones have been successful among the poor in part because usage charges are primarily marginal; these simulations suggest that governments can encourage adoption among the poor by taxing on the margin of usage rather than adoption.

However, the impact of taxation is likely to depend on the technology and the maturity of the network. These results all describe an early, growing network with relatively undifferentiated handsets sold by an independent market. Optimal policy for smartphone adoption in mature networks is likely to differ. However, across settings, more nuanced policies can collect revenue with less hinderance to adoption. An investment in a handset only generates network effects if it enables a new network service. In the time period I study, most handsets are purchased by consumers who obtain voice service for the first time. Likewise, when a voice subscriber upgrades to a smartphone they may obtain internet service for the first time. In the presence of public finance concerns, governments could tax purchases that enable new network services at a lower rate than upgrades that provide the same network services.[43]

10. Conclusion

This paper introduces a method for estimating and simulating the adoption of network goods. I overcome measurement issues that have limited empirical work on network goods by estimating the value of the network directly, from how it is used. I use a tractable framework and rich data on the adoption and usage of nearly an entire network of mobile phone users.

I demonstrate this method with two applications. I analyze the expansion of the mobile phone network into rural areas, and find that while most of the expansion of the network appears to be driven by private incentives, an obligation to provide coverage

[43]Or tax could be charged only on the difference between the price of a given handset and the price of the lowest cost handset with the same network functionality. Either approach would require categorizing handsets by network functionalities to be prioritized (e.g., allows voice, internet, or mobile wallet services).

in rural areas led to the building of otherwise unprofitable towers that improved welfare. I also simulate the effect of different tax policies. In growing networks, handset and usage taxes can inhibit adoption and are much more distortionary than naïve estimates would suggest. My results also suggest that handset taxes place a disproportionate burden on poorer phone adopters.

References

Ackerberg, D. A. and G. Gowrisankaran (2006): "Quantifying equilibrium network externalities in the ACH banking industry," *The Rand Journal of Economics*, 37, 738–761.

Aker, J. C. (2010): "Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger," *American Economic Journal: Applied Economics*, 2, 46–59.

Auriol, E. and M. Warlters (2012): "The marginal cost of public funds and tax reform in Africa," *Journal of Development Economics*, 97, 58–72.

Birke, D. and G. P. Swann (2010): "Network effects, network structure and consumer interaction in mobile telecommunications in Europe and Asia," *Journal of Economic Behavior & Organization*, 76, 153–167.

Bjorkegren, D. (2014): "Inferring Missing Cell Tower Locations using Call Handoffs," *White Paper*.

——— (2017): "The Joint Network Structure of Spillovers and Policy: Evidence from a Mobile Phone Subsidy in Rwanda," *Working Paper*.

Brynjolfsson, E. and C. F. Kemerer (1996): "Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market," *Management Science*, 42, 1627–1647.

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.

Fudenberg, D. and D. K. Levine (1988): "Open-loop and closed-loop equilibria in dynamic games with many players," *Journal of Economic Theory*, 44, 1–18.

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

Goolsbee, A. and P. J. Klenow (2002): "Evidence on Learning and Network Externalities in the Diffusion of Home Computers," *Journal of Law and Economics*, 45, 317–343.

Gordon, R. and W. Li (2009): "Tax structures in developing countries: Many puzzles and a possible explanation," *Journal of Public Economics*, 93, 855–866.

Gowrisankaran, G., M. Rysman, and M. Park (2010): "Measuring Network Effects in a Dynamic Environment," Tech. rep., SSRN, Rochester, NY.

GSMA (2008): "Taxation and the Growth of Mobile Services in Sub-Saharan Africa," .

——— (2013): "Universal Service Fund Study," Tech. rep., GSM Association.

Isaacman, S., R. Becker, R. Caceres, S. Kobourov, M. Martonosi, J. Rowland, and A. Varshavsky (2011): "Identifying Important Places in People's Lives from Cellular Network Data," in *Pervasive Computing*, ed. by K. Lyons, J. Hightower, and E. Huang, Springer, vol. 6696 of *Lecture Notes in Computer Science*, 133–151.

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

Jack, W. and T. Suri (2014): "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution," *American Economic Review*, 104, 183–223.

Jarvis, A., H. Reuter, A. Nelson, and E. Guevara (2008): "Hole-filled seamless SRTM data V4," *CIAT*.

Jensen, R. (2007): "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector," *Quarterly Journal of Economics*, 122, 879–924.

Jia, P. (2008): "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry," *Econometrica*, 76, 1263–1316.

Katz, M. L. and C. Shapiro (1985): "Network Externalities, Competition, and Compatibility," *The American Economic Review*, 75, 424–440.

Lee, R. S. (2013): "Vertical Integration and Exclusivity in Platform and Two-Sided Markets," *American Economic Review*, 103, 2960–3000.

Manski, C. F. (1993): "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 60, 531–542.

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

Ohashi, H. (2003): "The Role of Network Effects in the US VCR Market, 1978-1986," *Journal of Economics & Management Strategy*, 12, 447–494.

Pakes, A. (2010): "Alternative Models for Moment Inequalities," *Econometrica*, 78, 1783–1822.

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*.

Ryan, S. P. and C. Tucker (2012): "Heterogeneity and the dynamics of technology adoption," *Quantitative Marketing and Economics*, 10, 63–109.

Saloner, G. and A. Shepard (1995): "Adoption of Technologies with Network Effects: An Empirical Examination of the Adoption of Automated Teller Machines," *The RAND Journal of Economics*, 26, 479–501.

Stork, C. and M. Stork (2008): "ICT Household Survey Methodology & Fieldwork," Tech. Rep. 1, Research ICT Africa.

Topkis, D. M. (1978): "Minimizing a Submodular Function on a Lattice," *Operations Research*, 26, 305–321.

Tucker, C. (2008): "Identifying Formal and Informal Influence in Technology Adoption with Network Externalities," *Management Science*, 54, 2024–2038.

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

Yanagizawa-Drott, D. (2014): "Propaganda and Conflict: Evidence from the Rwandan Genocide," *The Quarterly Journal of Economics*, 129, 1947–1994.

Zwane, A. P., J. Zinman, E. Van Dusen, W. Pariente, C. Null, E. Miguel, M. Kremer, D. S. Karlan, R. Hornbeck, X. Gine, E. Duflo, F. Devoto,

B. Crepon, and A. Banerjee (2011): "Being surveyed can change later behavior and related parameter estimates," *Proceedings of the National Academy of Sciences.*

Appendix A. Functional Form of Calling Utility

The benefit of making calls, Equation 1, satisfies the following intuitive properties:

(1) Zero duration yields zero utility: $v(0, \epsilon) = 0$
(2) *Diminishing returns:* $v(d, \epsilon)$ is concave in d
(3) *For some values of ϵ and c, a call is placed.* The optimal duration yields nonnegative utility: $v(d^*, \epsilon) \geq 0$ where d^* solves $\frac{\partial v}{\partial d}(d^*, \epsilon) = c$ or is zero.
(4) *Even if calls were free, you wouldn't talk forever:* there is bounded demand under zero cost: $\frac{\partial v}{\partial d}(\bar{d}, \epsilon) = 0$ for some $\bar{d}$.
(5) *Changing the cost of a call affects the extensive decision to call:* this requires that marginal utility be finite at zero: $\frac{\partial v}{\partial d}(0, \epsilon) < \infty$
(6) *Changing the marginal cost of a call affects longer calls more*: $\frac{\partial^2 d^*}{\partial c \partial \epsilon} < 0$
(7) *The amount of information maps to duration:* there is a one to one mapping $\epsilon(d^*)$, which has an analytic solution that is efficient to compute.
(8) *Relationships with higher information flows provide more utility:* the optimized utility is increasing in the optimal duration: $\frac{\partial}{\partial d} v(d, \epsilon(d)) > 0$

Appendix B. Robustness: Network Value Implied by Adoption Model

As a robustness check, I test whether the value implied by usage coincides with the value implied by adoption. Taking the inequalities implied by the adoption decision (Equations 8 and 9), I pose that consumers value a dollar of call utility equally with λ dollars of handset price, and then test whether λ equals one. I form these into moment inequalities (for example, Pakes, 2010):

$$E\left[V_{mi} \left(\lambda \sum_{k=0}^{K-1} \delta^k E_\epsilon u_{i,x_i+k} - \left(p_{x_i}^{handset} - \delta^K p_{x_i+K}^{handset}\right) + (1-\delta^K)\eta_i\right)\right] \geq 0$$
$$E\left[V_{mi} \left(\lambda \sum_{k=1}^{K} \delta^{K-k} E_\epsilon u_{i,x_i-k} - \left(p_{x_i-K}^{handset} - \delta^K p_{x_i}^{handset}\right) + (1-\delta^K)\eta_i\right)\right] \leq 0$$

for a set of instruments $\{V_m\}_m$. As in the body of the paper I select $K = 2$.

Because of peer effects, straightforward estimation would face an endogeneity problem: if η_i is higher, i will adopt earlier, which will induce i's contacts to adopt earlier and raise u_{it}. Instead, I seek instruments that are correlated with the observed benefit of being on the network, u_{it}, but uncorrelated with residual shifters affecting adoption ($E\left[\eta_i|V_{mi}\right] = E\left[\eta_i\right]$). I consider two main sources of variation:

In 2008, Rwanda gave out 53,352 subsidized handsets in rural areas. This directly increased adoption among recipients, but also improved the utility of adopting for their contacts. I include the proportion of an individual's contacts who received a subsidy as an instrument. Because rural residents may find it most costly to adopt a handset, I use only variation within individuals who have a similar fraction of rural contacts. Bjorkegren (2017) finds that recipients of this program appear similar to nonrecipients, suggesting that individuals with many subsidized contacts are not likely to have differential residuals (η_i).

I also include instruments that shift the coverage an individual receives. Rwanda's topography is extremely hilly, and as a result, the cost of covering even very close villages can differ greatly. Since it is much cheaper to operate towers connected to the electric grid, a location is much more likely to receive coverage if it was visible from 35m above the existing electric grid, where a tower's transmitter would be constructed. I construct an instrument using the fictitious coverage that would have resulted had the operator built towers along the full network of power lines, using only variation for individuals at least 5 km from the grid.[44] While coverage itself is likely to be correlated with adoption residuals, this measure picks up idiosyncratic variation based on the interaction of geography with the preexisting electric grid that is unlikely to be. I instrument u_{it} with these measures for i's own location as well as those of his contacts.

The Supplemental Appendix provides more details on the construction of these instruments, tests of relevance, and suggestive tests of exclusion.

In the body of the paper I impose $\lambda = 1$ and do not restrict $E[\eta_i]$. Here, if I do not restrict $E\left[\eta_i\right]$, the implied bounds for λ include 1, but are very wide. If I impose

[44]While villages close to an electric grid will tend to have less obstructed views, they differ in other ways from those further away. However, that effect is likely to attenuate with distance more quickly than cellular coverage (cell towers have a range of up to 35 km).

$E[\eta_i] = 0$, I find that the value implied by adoption is very close to that implied by calls. If recipients do not value incoming calls ($w = 0$), the adoption decision implies that consumers value a dollar of call utility as $\lambda \in \$1.02 - 1.17$ of handset price.[45] If recipients also receive the surplus from incoming calls ($w = 1$), the model appears to double count the utility from calls: consumers value a dollar of call utility as $\lambda \in \$0.27 - 0.31$ of handset price.

Appendix C. Simulation Algorithm

(1) Propose a candidate adoption path $\boldsymbol{x}^0$
(2) Allow each individual to optimize their decision, holding fixed the adoption path of others:[46] $x_i^{k+1} = \min_t s.t. \left[U_i^t(\boldsymbol{x}_{G_i}^k) \geq \max_{s>t} E_t U_i^s(\boldsymbol{x}_{G_i}^k)\right]$
(3) Iterate until the equilibrium converges:[47] $x_i^{k+1} = x_i^k$ for all i

[45] Outlier nodes earning a very high utility would skew the estimate, so I omit the nodes earning the top 0.1% utility from the network.
[46] I hold fixed the adoption dates of initial adopters and individuals whose η_i I cannot back out (see Section 6).
[47] The algorithm sometimes reaches a cycle rather than an equilibrium. These cycles tend to involve only a handful of nodes. If the algorithm reaches a cycle, I break the cycle and note the number of nodes involved.