Social Networks and Internal Migration*

Joshua Blumenstock†
University of California, Berkeley

Xu Tan‡
University of Washington

September 20, 2017

PRELIMINARY - DO NOT CITE OR CIRCULATE WITHOUT PERMISSION

Abstract

How does the structure of an individual's social network affect his or her decision to migrate? Economic theory suggests two prominent mechanisms — as conduits of information about jobs, and as a safety net of social support — that have historically been difficult to differentiate. We bring a rich new dataset to bear on this question, which allows us to adjudicate between these two mechanisms and add considerable nuance to the discussion. Using the universe of mobile phone records of an entire country over a period of four years, we first characterize the migration decisions of millions of individuals with extremely granular quantitative detail. We then use the data to reconstruct the complete social network of each person in the months before and after migration, and show how migration decisions relate to the size and structure of the migrant's social network. We use these stylized results to develop and estimate a structural model of network utility, and find that the average migrant benefits more from networks that provide social support than networks that efficiently transmit information. Finally, we show that this average effect masks considerable heterogeneity in how different types of migrants derive value from their social networks.

*We thank Lori Beaman, Matthew Jackson, David Miller, Matthew Olckers, Mark Rosenzweig, Ott Toomet, Yves Zenou, as well as seminar participants at U.C. Berkeley, the NSF Conference on Network Science in Economics, the Economic Demography Workshop, and the Barcelona GSE Summer Forum for helpful feedback. We are grateful for financial support from the UW Royalty Research Fund Grant 65-7397. All errors are our own.

†University of California, Berkeley, jblumenstock@berkeley.edu

‡University of Washington, tanxu@uw.edu

1 Introduction

Migrants play a central role in bringing an economy towards a more efficient use of its resources. In many contexts, however, a range of market failures limit the extent to which people can capitalize on opportunities for arbitrage through migration. Recent literature documents, for instance, cases where information does not reach the migrant (Jensen, 2012), households lack insurance against the risk of migration (Bryan, Chowdhury and Mobarak, 2014), and source communities discourage exit (Beegle, De Weerdt and Dercon, 2010, Munshi and Rosenzweig, 2016). The resulting underinvestment in migration leads to the misallocation of capital, and can have severe consequences for the overall economy.[1]

The decision to migrate depends on the extent to which the migrant is connected to communities at home and in the destination. Much of the existing literature has focused on how strong ties to the destination community can facilitate migration by providing access to information about jobs (Borjas, 1992, Topa, 2001, Munshi, 2003) and material support for reecent arrivals (Munshi, 2014). The role of the home network is more ambiguous. On the one hand, robust risk sharing networks can partially insure against the risk of temporary migration (Morten, 2015), making it easier for people to leave. On the other hand, strong source networks can also discourage permanent migration if migrant households are subsequently excluded from risk sharing networks (Munshi and Rosenzweig, 2016).

While there is thus general consensus that social networks play an important role in migration decisions, the exact nature of this role is unclear. This ambiguity stems, at least in part, from a lack of reliable data on both migration and the structure of social networks. Migration is difficult to measure, particularly in developing countries where short-term migration is common and reliable household survey data is limited (Deshingkar, Grimm and Migration, 2005, McKenzie and Sasin, 2007, Carletto, de Brauw and Banerjee, 2012, Lucas, 2015). Social network structure is even harder to observe.[2] Recent empirical work

[1] Bryan, Chowdhury and Mobarak (2014), for instance, link underinvestment in migration to seasonal famine.

[2] For instance, the rich survey data used in Banerjee et al. (2013, 2014) cost $250,000 to collect - cite ARD

on networks relies on survey modules that ask respondents to list their social connections, but this approach is necessarily limited in scope and scale. Thus, much of the literature on networks and migration relies on indirect information on social networks, such as the (plausible) assumption that individuals from the same hometown, or with similar observable characteristics, are more likely to be connected that two dissimilar individuals.[3]

We leverage a novel source of data to provide detailed insight into the role of social networks in the decision to migrate. Using several years of data capturing the entire universe of mobile phone activity in Rwanda, we track the internal migration decisions of roughly one million unique individuals, as inferred from the locations of the cellular towers they use to make and receive phone calls. We link these migration decisions to the structure of each migrant's social network, as inferred from the set of people with whom he or she interacts over the phone network. Merging the geospatial and network data, we observe the migrant's connections to his home community, his connections to all possible destination communities, as well as the complete higher-order structure of the network (i.e., the connections of the migrant's connections).

We first use these data to validate a common hypothesis in prior theoretical and empirical work: that individuals are more likely to migrate to destinations to which they have stronger social ties. We measure both the extensive margin of number of unique contacts as well as the intensive margin of the frequency of communication with those contacts, and find that the probability of migration is increasing in both. The relationship is monotic with constant elasticity, such that the probability of migration roughly doubles as the number of contacts in the destination doubles.[4] Similarly, we observe that stronger networks in the home community will make a migrant less likely to leave, which is consistent with a story where individuals fear being ostracized from inter-family insurance networks (Munshi

[3]Identification strategies are varied: for example, Beaman (2012) uses data on resettled political refugees; Munshi (2003) use rainfall shocks at origin; ...;

[4]Superficially, this result diverges from a series of studies that predict a eventual negative externalities from network size, as when members compete for information and opportunities (Calv-Armengol, 2004, Calv-Armengol and Jackson, 2004, Beaman, 2012). Note that our data only permit a short-run analysis, so we cannot test, for instance, the heterogeneity by vintage of network members as predicted in these models.

and Rosenzweig, 2016). Our data indicate a decreasing and monotonic relationship between migration rates and the extent of the home network.

To provide structure to these and subsequent results, we develop a model that relates internal migration decisions to social network structure. The model characterizes the migration decision as, ceteris paribus, a tradeoff between the utility an agent receives from the home network and the utility received from a potential destination network, net an idiosyncratic cost of migrating. How agents derive utility from their social networks is not known a priori, and we show that certain simple models of network utility – such as a model of information diffusion as in Banerjee et al. (2013) – cannot explain important characteristics of our data. For instance, we document – to our knowledge for the first time – the role that more distant network connections play in migration. Namely, we find that an individual is more likely to migrate to a destination where her friends have more friends. However, contrary to what would be expected in most models of information diffusion, this effect does not persist after controlling for the number of friends in the destination. For example, if both Joe and Jane have the same number of contacts living in a destination community, but Joe's contacts have more contacts in the destination than Jane's contacts do, we find that Joe is no more likely to migrate to the destination than Jane.

These observations lead us to develop a structural model of the utility migrants obtain from their social networks. This model, while highly stylized, allows us to differentiate between two of the primary mechanisms articulated in prior work on networks and migration: the potential for the network to transmit information to the migrant (for instance about jobs and opportunities), and the potential for the network to support dynamic cooperation (as with risk sharing and favor exchange). In this model, information is treated as a diffusion process with possible loss of information; cooperation as a strategic interaction where agents randomly meet their connected neighbors over time, and when two agents meet, they each contribute effort to a joint project. Effort is determined endogenously by the network structure, so the model allows us to describe in equilibrium how network structure affects the social value that agents get from the network, which in turn affects the decision to migrate.[5]

Together, our model and data make it possible to generate and test several hypotheses about the role of social networks in migration that the existing literature has been unable to explore. Notably, we show that social connections generate positive externalities. That is, if two agents form a link or increase their interaction over the existing link, their common neighbors (in addition to themselves) receive strictly higher utilities from the network. This generates the testable predictions that an agent should be more likely to migrate if her connections in the destination form more links among themselves, and if the frequency of interaction between common neighbors increases. As we show, each of these predictions is supported by the data, although the shape of the migration response function is not always linear or monotonic.

[Structural estimation results here.]

Our final set of results explores heterogeneity in the migration response to social network structure. We separately study the role of the network in migration between and across rural and urban areas, in short- and long-distance moves, and in temporary vs. permanent migrations. While the main effects described above are generally consistent in each of these sub-populations, the shape and magnitude of the migration response differs significantly by migration type.

Since our approach to studying migration with mobile phone data is new, we perform a large number of specification tests to calibrate for likely sources of measurement error and to test the robustness of our results. In particular, one limitation of our approach is that we lack exogenous variation in the structure of an individual's network, so that network structure may be endogenous to decisions regarding migration. We address this concern in two principal ways. First, we derive structural properties of the migrant's social network in the period prior to migration. Our results change little even when we reconstruct each

[5] See Ali and Miller (2016) for a related approach, which builds on past observations that social sanctions can improve commitment in risk sharing (Chandrasekhar, Kinnan and Larreguy, 2014, Karlan et al., 2009) and may strengthen job referral networks (Heath, 2016).

migrant's social network using communications data from several months prior to the date of migration. Second, we leverage the vast quantity of data at our disposal to control for a robust set of network characteristics and better isolate the structural parameter of interest. For instance, we condition on the number of common neighbors when analyzing the effect of the frequency of communication between common neighbors. Thus, while having a large number of contacts in a destination may be endogenous to migration, and likely migrants may even select contacts who are connected to each other, we assume migrants will be less able to control the extent to which those contacts communicate amongst themselves.

This paper makes two primary contributions. First, we contribute to a growing literature on the economic value of social networks (cf. Jackson, Rodriguez-Barraquer and Tan, 2012, Banerjee et al., 2013, **?**). Our model connects this literature to research on migration in developing countries, and indicates that a model combining information and cooperation is more consistent with the data than a model of information diffusion. Second, we contribute to empirical research on the determinants of internal migration in developing countries (cf. Bryan, Chowdhury and Mobarak, 2014, Morten, 2015, Lucas, 2015). In this literature, it has historically been difficult to empirically characterize the relationship between social networks and migration; our data make it possible to directly test several conjectures in the prior literature, and to develop new insight into the relationship between social network structure and the decision to migrate.

2 A strategic model of migration

People can derive utility from their social networks in myriad ways (Jackson, 2010). We focus our model on two stylized features of social networks that the literature has consistently shown to play an important role in the decision to migrate. The first is the potential for the social network to provide the migrant with access to information about jobs, new opportunities, and the like (Topa, 2001, Calv-Armengol and Jackson, 2004, Banerjee et al., 2013), which we denote as u^I. The second is the utility agents derive from interactions involving repeated cooperation, such as risk sharing and social insurance (Munshi and Rosenzweig, 2016, Jackson, Rodriguez-Barraquer and Tan, 2012), which we denote by u^C. We describe these in turn below, and then develop a strategic model of migration that allows for both factors to influence the migration decision.

2.1 Utility from information

A robust theoretical and empirical literature studies processes of information diffusion on networks.[6] We build on recent efforts by Banerjee et al. (2013) to model the value of information as a diffusion process with possible loss of information.

In this model, agents meet with their neighbors repeatedly for T periods. When they meet, they share information with each other with probability q. Thus the amount of information agent i gets by the end of period T is the ith entry of the vector

$$DC(G; q, T) \equiv \sum_{t=1}^{T} (qG)^t \cdot \mathbf{1} \tag{1}$$

in which G is the adjacency matrix of the network.

Several intuitive predictions can be derived from such a model: an agent receives additional utility from each friend in the network, additional (q-discounted) utility from each friend of those friends, q^2-discounted utility from the friends of the friends' friends, and so forth. As we later show, not all of these predictions are supported by our data, and other strong correlates of migration are not easily reconciled with this model.

2.2 Utility from repeated cooperation

Consider a population of N players, $N = \{1, \ldots, n\}$, who are connected in an undirected network G, with $ij \in G$ if agent i and j are connected (we abuse the notation of G slightly).

[6]See Jackson and Yariv (2010) for a summary of both mechanical and strategic models of communication and diffusion.

Denote agent i's neighbors as $N_i = \{j : ij \in G\}$, and her degree as $d_i = |N_i|$.

Each pair of connected agents, $ij \in G$, is engaged in a partnership ij that meets at random times generated by a Poisson process of rate $\lambda_{ij} > 0$. When they meet, agent i and j choose their effort levels a_{ij}, a_{ji} in $[0, \infty)$ as their contributions to a joint project.[7] Player i's stage game payoff function when partnership ij meets is $b(a_{ji}) - c(a_{ij})$, where $b(a_{ji})$ is the benefit from her partner j's effort and $c(a_{ij})$ is the cost she incurs from her own effort. The benefit function b and the cost function c are smooth functions satisfying $b(0) = c(0) = 0$. All players share a common discount rate $r > 0$, and the game proceeds over continuous time $t \in [0, \infty)$.

We write the net value of effort a as $v(a) \equiv b(a) - c(a)$, and we assume that it grows in the following manner.

Assumption **1.** *The net value of effort $v(a)$ is strictly increasing and weakly concave, with $v(0) = 0$. Moreover, $v'(a)$ is uniformly bounded away from zero.*

Assumption 1 implies that higher effort is always socially beneficial; concavity means it is better for partners to exert similar effort, holding their average effort constant. The following assumption articulates that higher effort levels increase the temptation to shirk.

Assumption **2.** *The cost of effort c is strictly increasing and strictly convex, with $c(0) = c'(0) = 0$ and $\lim_{a\to\infty} c'(a) = \infty$. The "relative cost" $c(a)/v(a)$ is strictly increasing.*

Strict convexity with the limit condition guarantees that in equilibrium effort is bounded (as long as continuation payoffs are bounded, which we assume below). Increasing relative cost means a player requires proportionally stronger incentives to exert higher effort.

As has been documented in several different real-world contexts, we assume agents have only local knowledge of the network.[8] Each agent only observes her local network, including her neighbors and the links among them (in additional to her own links). To be precise,

[7]The variable-stakes formulation is adopted from Ghosh and Ray (1996) and Ali and Miller (2016).

[8]Examples in the literature include Krackhardt (1990), Casciaro (1998) and Chandrasekhar, Breza and Tahbaz-Salehi (2016).

agent i observes her neighbors in N_i and all links in $G_i = \{jk : j, k \in \{i\} \cup N_i\}$. Moreover, we consider locally public monitoring, such that each agent learns about her neighbors' deviation, and we assume this information travels instantly.

Homogenous meeting frequency

As a benchmark, we start with the case that λ is the same across agents. Following the definition from Jackson, Rodriguez-Barraquer and Tan (2012), a link ij is *supported* if they have at least one common neighbor $k \in N_i \cap N_j$, and ij is *m-supported* if they have m common neighbors. There are critical effort levels, for supported and unsupported links.

Unsupported cooperation. Consider a strategy profile in which each of i and j exerts effort level a_0, if each has done so in the past; otherwise, each exerts zero effort.

$$b(a_0) \leq v(a_0) + \int_0^\infty e^{-rt} \lambda v(a_0) dt. \tag{2}$$

The incentive constraint is binding at effort level a_0^*.

Supported cooperation. Consider a triangle i, j, k and a strategy profile in which each of them exerts effort level a_1, if each has done so in the past; otherwise, each exerts zero effort.

$$b(a_1) \leq v(a_1) + 2 \int_0^\infty e^{-rt} \lambda v(a_1) dt. \tag{3}$$

The incentive constraint is binding at effort level a_1^*. Notice that the future value of cooperation is higher in a triangle, $2 \int_0^\infty e^{-rt} \lambda v(a_1) dt$, so it can sustain higher level of efforts $a_1^* > a_0^*$ and everyone gets a strictly higher utility.

Higher-order supported cooperation. Let $m(ij) = |N_i \cap N_j|$ be the number of agents who are common neighbors of i and j. Consider i and j sharing m common neighbors and a strategy profile in which each of them exerts effort level a_m if each has done so in the past;

otherwise, each exerts zero effort.

$$b(a_m) \leq v(a_m) + \int_0^\infty e^{-rt} \lambda \left[v(a_m) + mv(a_1^*) \right] dt, \tag{4}$$

in which i and j assume other pairs cooperate on at least a_1^* when they are supported. The incentive constraint is binding at effort level a_m^*. Following the same argument, more common neighbors can sustain a higher level of cooperation between i and j, that is a_m^* strictly increases in m.

Definition 1. *A strategy profile is measurable to local networks if the effort level between any pair of agents only depends on the local network they share.*

This makes sure that agents do not need knowledge of the outside network structure.

Definition 2. *A strategy profile is strongly robust if any pair of agents who have not deviated always cooperate at the same level, on and off the path of play.*

This property is stronger than the robustness criterion used by Jackson, Rodriguez-Barraquer and Tan (2012), which allowed for cooperation to break down among a bounded set of innocent players following a deviation by one of their neighbors.

Proposition 1. *Consider the game with homogenous meeting frequency. There exists an equilibrium, measurable to local networks and strongly robust, in which any pair of connected agents, say i and j, cooperate on $a_{m(ij)}^*$, where $m(ij) = |N_i \cap N_j|$.*

All proofs are in Appendix A1. It is easy to see that there is an even simpler equilibrium in which there are two cooperation levels, high level for supported links and low level for unsupported links.

Corollary 1. *Consider the game with homogenous meeting frequency. There exists an equilibrium, measurable to local networks and strongly robust, in which any pair of connected agents, cooperate on a_1^* if the link is supported, and on a_0^* otherwise.*

Intuitively, the more common neighbors a pair of agents have, the higher utility they can get from their cooperation. That is, a^*_m increases in the number of common neighbors m. Thus, in the equilibrium above each agent gets a strictly higher utility if she forms more links, or if her neighbors form more links among themselves.To generalize this intuition, we now show that if an agent's degree, support or clustering increases,[9] then she can get a higher utility from the network. In particular, while there are many possible equilibria, we restrict our attention to those in which each agent gets a positive expected payoff from each link.

PROPOSITION **2.** *Consider two networks, G and $G' = G \cup \{ij\}$ such that $ij \notin G$. For any equilibrium Σ_G in network G, there is an equilibrium $\Sigma_{G'}$ in network G', such that everyone gets a weakly higher utility in $\Sigma_{G'}$ and for any agent $k \in \{i, j\} \cup (N_i \cap N_j)$ in network G, k must get a strictly higher utility in $\Sigma_{G'}$.*

The proposition shows that each link not only benefits its two agents, but also exhibits positive externalities. First of all, the link ij gives agent i and j each a higher utility due to this new cooperation opportunity. As a result, they get a higher utility from cooperation and thus they face a higher punishment if they deviate. This additional punishment then sustains i and j's incentives to cooperate at a higher level with their common neighbors $k \in N_i \cap N_j$, who can observe the link. So k can get a higher utility once i and j are connected.

Heterogeneous meeting frequency

In the data, we can measure the communication frequency between any pair of agents. To examine its effect on the utility one gets from the network, we now allow heterogeneous meeting frequency, and λ_{ij} is locally observed by i, j and their common neighbors in $N_i \cap N_j$.

First, in a bilateral partnership, agents get a higher utility when they meet more often. Let $a^*_0(\lambda_{ij})$ be the unsupported effort level when only i and j are connected and they meet with the frequency λ_{ij}. The counterpart to equation (2) becomes $b(a_0) \leq$

[9] "Support" is defined as the fraction of one's links that are supported; "clustering" is defined as the fraction of pairs of one's neighbors that are connected.

$v(a_0) + \int_0^\infty e^{-rt}\lambda_{ij}v(a_0)dt$. The incentive constraint is binding at effort level $a_0^*(\lambda_{ij})$. It is easy to verify that $a_0^*(\lambda_{ij})$ increases in λ_{ij}, which implies as the meeting frequency increases, the utility agent i and j can obtain from their cooperation increases.

This is also true in an arbitrary network, such that an agent can get a higher utility if her interaction frequency with her neighbors increases and/or the interaction frequency between two of her neighbors increases.

PROPOSITION **3.** *Consider the game with heterogeneous meeting frequencies, and increase the frequency on one and only one link $\lambda'_{ij} > \lambda_{ij}$. For any equilibrium Σ_λ, there is an equilibrium $\Sigma_{\lambda'}$, such that everyone gets a weakly higher utility in $\Sigma_{\lambda'}$ and for any agent $k \in \{i,j\} \cup (N_i \cap N_j)$, k must get a strictly higher utility in $\Sigma_{\lambda'}$.*

The proposition shows that the interaction frequency also exhibits positive externalities. As i and j meet more often, they each gets a higher utility from cooperation, which provides the incentive to not only contribute greater effort to their partnership ij, but also to partnerships with their common neighbors $k \in N_i \cap N_j$. Thus, k receives a higher utility from i and j as they meet more frequently.

However, the positive externalities found in Proposition 2 and Proposition 3 are local effects to agents in $\{i,j\} \cup (N_i \cap N_j)$. Other agents who only know either i or j do not know whether ij are connected or their frequency of interaction. Thus, they cannot choose their efforts based on the link ij, nor benefit from its existence or its increased frequency. This type of local knowledge seems particularly relevant when a person is considering migrating to a potential destination. Because an agent has not yet moved to the destination, it is unlikely that she knows much beyond her immediate neighbors.

To summarize from Proposition 2 and 3, the cooperation model has the following testable implications:

REMARK **1.** *In general, an agent is more likely to migrate if her network in the destination has (or she is less likely to migrate if her network in the hometown has):*

- *Higher degree;*
- *Higher support/clustering, when fixing degree;*
- *Higher own interaction frequency, when fixing the network;*
- *Higher interaction frequency between neighbors, when fixing the network.*

The number of indirect neighbors in the destination network has no effect on one's migration. However, if agents know their indirect neighbors within a certain distance at home, then the number of these indirect neighbors has a negative effect on one's migration.

2.3 The full model

We are now ready to discuss the migration decision. We say that an individual i receives utility $u_i(G)$ from a social network G. In deciding whether or not to migrate, the individual weighs the utility of i's home network G^H against the utility of the potential destination network G^D, and migrates if the difference is greater than some threshold τ plus an idiosyncratic error ε_i that can reflect, among other things, the extent to which i is unusually averse to migrating.

$$u_i(G^D) - u_i(G^H) > \tau + \varepsilon_i \tag{5}$$

For an arbitrary network G, we assume that the total utility an agent i receives from G can be expressed as

$$u_i = u_i^C + u_i^I \tag{6}$$

in other words, a linear combination of value from information and value from cooperation.[10]

[10] We do not imply that u^I and u^C are orthogonal or that other aspects of network do not weigh in the decision to migrate. However, this formulation allows us to contrast two archetypical properties of network structure that can be estimated with our data.

As a starting point, we use Corollary 1 as the equilibrium for value from cooperation. (Later on, we will extend the estimation to more complex equilibrium in Proposition 1.) That is, agent i gets a utility of

$$u_i^C = u_0 d_i^{NS} + u_1 d_i^S \tag{7}$$

from cooperating with her neighbors, in which $u_0 = \frac{\lambda}{r} v(a_0^*)$ is the utility of cooperating on an unsupported link, and $u_1 = \frac{\lambda}{r} v(a_1^*)$ is the utility of cooperation on a supported link. d_i^{NS} is the number of i's unsupported links, and d_i^S is the number of i's supported links. The amount of information agent i gets is simply her diffusion centrality, $u_i^I(G) = DC_i$, and an agent gets $\tilde{u}$ for each unit of information. So the overall utility is

$$u_i = \tilde{u} DC_i + u_0 d_i^{NS} + u_1 d_i^S \tag{8}$$

We want to contrast the value of information versus the value of cooperation, and contrast the value of unsupported links versus supported links. So we replace the parameters $(u_0, u_1, \tilde{u})$ by (π, α, ρ) and rewrite the overall utility:

$$u_i = \rho \left(\pi DC_i + (1-\pi) \left(\alpha d_i^{NS} + (1-\alpha) d_i^S \right) \right) \tag{9}$$

Then (5) becomes

$$\begin{aligned} &\rho^D \left(\pi^D DC_i(G^D) + (1-\pi^D) \left(\alpha^D d_i^{NS} G^D) + (1-\alpha^D) d_i^S G^D) \right) \right) \\ &\quad - \rho^H \left(\pi^H DC_i(G^H) + (1-\pi^H) \left(\alpha^H d_i^{NS}(G^H) + (1-\alpha^H) d_i^S(G^H) \right) \right) > \tau + \varepsilon_i \end{aligned} \tag{10}$$

We can divide ρ^D from both sides of the inequality, and let $\rho = \rho^H/\rho^D$, then

$$\pi^D DC_i + (1-\pi^D)\left(\alpha^D d_i^{NS} + (1-\alpha^D) d_i^S\right) - \rho\left(\pi^H DC_i + (1-\pi^H)\left(\alpha^H d_i^{NS} + (1-\alpha^H) d_i^S\right)\right) > \hat{\tau} + \hat{\varepsilon}_i \quad (11)$$

Notice that we allow agents to have different weights π for home and destination network, because it is possible that the relative value of information and cooperation is different in a home network than in a destination network.[11] Similarly, we also allow (α, ρ) to differ for home and destination network. When we do not differentiate home and destination, $\rho^D = \rho^H$, then $\rho = 1$. The basic formulation thus leaves three sets of structural parameters of interest: π^H (or π^D), the importance of information in a home (or destination) network relative to cooperation; α^H (or α^D), the value of cooperation from an unsupported link relative to a supported link; and $\hat{\tau}$ (and ρ), which we loosely interpret as the average cost of migrating.

3 Data

We exploit a novel source of data to test the predictions of our model. These data make it possible to observe rich information about the social network structure and migration histories of over a million individuals in Rwanda. The data were obtained from Rwanda's primary mobile phone operator, which held a near monopoly on mobile telephony until late 2009. We focus on an analysis of the operator's mobile phone Call Detail Records (CDR) covering a 4.5-year period from January 2005 until June 2009. The CDR contain detailed metadata on every event mediated by the mobile phone network. In total, we observe over 50 billion mobile phone calls and text messages. For each of these events, we observe a unique identifier for the caller (or sender, in the case of a text message), a unique identifier for the recipient, the date and time of the event, as well as the location of the cellular phone towers

[11] In ongoing work, we allow for different weights for rural and urban networks.

through which the call was routed. All personally identifying information is removed from the CDR prior to analysis.

We use these data to infer migration events, and to observe the social network structure, of each of the roughly 1.5 million unique subscribers who appear in the dataset. Summary statistics are presented in Table 2. Our methods for inferring migration and measuring social networks are described below. In Section 6.1, we address the fact that the mobile subscribers we observe are a non-random sample of the overall Rwandan population, and discuss the extent to which these issues might bias our empirical results.

3.1 Measuring migration with mobile phone data

We construct individual migration trajectories for each individual in three steps.

First, we extract from the CDR the approximate location of each individual at each time in which he or she is involved in a mobile phone event, such as a phone call or text message. This creates a set of tuples $\{ID, Timestamp, Location\}$ for each subscriber. We cannot directly observe the location of any individual in the time between events appearing in the CDR. The location is approximate because we can only resolve the location to the geocoordinates of the closest mobile phone tower (in standard GSM networks, the operator does not record the GPS location of the subscribers). The locations of all towers in Rwanda, circa 2008, are shown in Figure 1.

Second, we assign each subscriber to a "home" district in each month of the data in which she makes one or more transactions. Our intent is to identify the location at which the individual spends the majority of her time, and specifically, the majority of her evening hours.[12] We treat the three districts that comprise the capital of Kigali as a single urban district; the 27 other districts in Rwanda are treated as separate rural districts. Algorithm 1 describes the algorithm exactly. To summarize, we first assign all towers to a geographic

[12] A simpler approach simply uses the model tower observed for each individual in a given month as the "home" location for that person. While our later results do not change if home locations are chosen in this manner, we prefer the algorithm described in the text, as it is less susceptible to biases induced from bursty and irregular communication activities.

district, of which there are 30 in Rwanda (see Figure 1). Then, for a given month and a given individual, we separately compute the most frequently visited district in every hour of that month (e.g., a separate modal district-hour is calculated for each of the 24×30 different hours in a 30-day month). Focusing only on the hours between 6pm and 7am, we then determine the for each day in the month, that individual's monthly modal district-day – defined as the district that is observed with the largest number of modal district-hours for the following night. Finally, we determine the modal monthly district for that individual as the district in which the individual is observed for the largest number of modal district-days.[13] After this step, we have an unbalanced panel indicating the home loaction of each individual in each month.

Finally, we use the sequence of monthly home locations to determine whether or not each individual i migrated in each month t. As in Blumenstock (2012), we say that a migration occurs in month t if three conditions are met: (i) the individual's home location is observed in district d for at least k months prior to (and including) t; (ii) the home location d' in t is different from the home location in $t+1$; and (iii) the individual's new home location is observed in district d' for at least k months after (and including) $t+1$. Individuals whose home location is observed to be in d for at least k months both before and after t are considered residents, or stayers. Individuals who do not meet these conditions are treated as "other" (and are excluded from later analysis).[14] Complete details are given in Algorithm 2. Our preferred specifications use $k = 2$, i.e., we say a migration occurs if an individual stays in one location for at least 2 months, moves to a new location, and remains in that new

[13] At each level of aggregation (first across transactions within an hour, then across hours within a night, then across nights within a month), there may not be a single most frequent district. To resolve such ties, we use the most frequent district at the next highest level of aggregation. For instance, if individual i is observed four times in a particular hour h, twice in district p and twice in q, we assign to i_h whichever of p or q was observed more frequently across all hours in the same night as h. If the tie persists across all hours on that night, we look at all nights in that month. If a tie persists across all nights, we treat this individual as missing in that particular month.

[14] Individuals are treated as missing in month t if they are not assigned a home location in month any of the months $\{t-k, ..., t, t+k\}$, for instance if they do not use their phone in that month or if there is no single modal district for that month. Similarly, individuals are treated as missing in t if the home location changes between $t-k$ and t, or if the home location changes between $t+1$ and $t+k$.

location for at least 2 months. While the number of observed migrations is dependent on the value of k chosen, we show in Section 6 that our results are not sensitive to reasonable values of k.

Figure 3 shows the distribution of individuals by migration status, for a single month (January 2008). To construct this figure, and in the analysis that follows, we classify migrations into three types: rural-to-urban if the individual moved from outside the capital city of Kigali to inside Kigali; urban-to-rural if the move was from inside to outside kigali; and rural-to-rural if the migration was between districts outside of Kigali. As can be seen in the figure, of the 15,849 migrations observed in that month, the majority (10,059) the majority occured between rural areas; 2,795 people moved from rural to urban areas and 2,995 moved from urban to rural areas.

3.2 Inferring social network structure from mobile phone data

The mobile phone data allow us to observe all mobile phone calls placed, and all text messages sent, over a 4.5-year period in Rwanda. From these pairwise interactions, we can construct a very detailed picture of the social network of each individual in the dataset. To provide some intution, the network of a single migrant is shown in Figure 2. Nodes in this diagram represent individuals and edges between nodes indicate that those individuals were observed to communicate in the month prior to migration. The individual i of interest is shown as a green circle; red and blue circles denote i's direct contacts (blue for people who live the migrant's home district and red for people in the migrant's destination district); grey circles indicate i's "friends of friends", i.e., people who are not direct contacts of i, but who are direct contacts of i's contacts.[15]

To test the empirical predictions of the model described in Section 2, we collapse this network structure into a handful of descriptive characteristics, separately for each of the roughly 1 million individuals in our dataset, for each of the 24 months that we study. The

[15]Nodes are spaced using the force-directed algorithm described in Hu (2005).

characteristics of primary interest are:

- **Degree Centrality:** The number of unique individuals with whom i is observed to communicate.
- **"Information":** The number of friends of friends of i. Specifically, we count the unique 2nd-order connections of i, excluding i's direct connections.
- **"Support":** The number of i's neighbors who share a common neighbor with i.
- **"Weighted" degree, information, and support:** Accounts for the frequency of interaction between neighbors, following the discussion in Section 2.2. Specifically, *weighted degree* is the number of interactions between i and her immediate neighbors. *weighted information* is the count of all interactions between i's neighbors and their neighbors. *Weighted support* is the count of all interactions between i's neighbors and their common neighbors of i.

In genreal, we compare network characteristics derived from data in month $t-1$ to migration behavior observed in month t. Concerns of serial correlation are discussed in Section 6.

4 Characterizing migration and social networks

To study the relationship between social network structure and the decision to migrate, we compare characteristics of individual i's network in month $t-1$ with the migration decision made by i in month t. Our canonical specification requires that the individual remain in one district for $k = 2$ months, then move to another place for $k = 2$ months, to be considered a migrant. As a concrete example, when t is set to January 2008, the individual is considered a migrant if her home location is determined to be one district d in December 2007 and January 2008, and a different district $d' \neq d$ in both February 2008 and March 2008. The first column of Table 2 shows how the sample of 455,704 unique individuals is distributed across residents and migrants, for just the month of January 2008. To increase the power of our analysis, we then aggregate migration behavior over the 24 months between July 2006 and June 2008. Summary statistics for this aggregated person-month dataset are given in Table 2, column 2.[16]

We calculate properties of i's network in $t-1$ following the procedures described in Section 3.2 for both the individual's home and destination networks. This is a straightforward process for the home network: we determine i's home location d in $t-1$, consider all contacts of i whose home location in $t-1$ was also d, and then calculate the properties of that induced subnetwork. Calculating properties of the destination network is more subtle, since non-migrants do not have a destination. To address this, for every individual we consider all 27 districts other than the home district as a "potential" destination, and separately study each of i's 27 potential migrations. [17]

Our core empirical results compare the migration outcomes of people with differently structured social networks. Our analysis focuses focus on the key predictions of the model in Section 2: that individuals are more likely to migrate if their destination network has (i) higher degree; (ii) higher interaction frequency, which we calculate as weighted degree; (iii) higher support, fixing degree; and (iv) greater interaction between neighbors, fixing the network, which we calculate as weighted support; and (v) the extent of the neighbor's neighbors, which we label "information." We discuss each of these results in turn in the sections that follow, and summarize them in Tables 3-4

While it is tempting to interpret these relationships as indicative of the causal effect of network structure on migration decisions, the evidence we present simply shows the cor-

[16]Note that this process of aggregation means that a single individual will appear multiple times in our analysis. In later robustness tests we show that very little changes if we restrict our analysis to a single month.

[17]When estimating standard errors and confidence intervals, we cluster by individual-month. In robustness tests described in Section 6, we run a separate specification that allows each individual to have only one potential destination, defined as the district other than d to which the individual made the most phone calls in $t-1$.

relation between migration and network structure. We attempt to limit the endogenity of network structure in two principal ways. First, we measure network structure in the months before the migration event actually occurs, and in separate robustness checks, show that our results do not change even when the network is measured 6 or 12 months prior to migration. This mitigates the possibility that migrants shape their networks in immediate anticipation of migrating. Second, in the majority of the results that follow, we will control for lower-order network structure when analyzing higher-order network structure. For instance, we condition on the number of common neighbors when analyzing the effect of the frequency of communication between common neighbors. Thus, while having a large number of contacts in a destination may be endogenous to migration, and likely migrants may even select contacts who are connected to each other, we assume migrants will be less able to control the extent to which those contacts communicate amongst themselves. In spite of these attempts, we acknowledge that we are unable to eliminate concerns of the endogeneity of network structure.

4.1 Degree centrality and weighted degree

Figure 4a shows the relationship between the migration rates and degree centrality in the destination. A point on this figure can be interpreted as the average migration rate (y-axis) across individuals with a fixed number of contacts in the destination (x-axis). For instance, roughly 11% of individuals who have 30 contacts in a potential district d' in month $t-1$ are observed to migrate to d' in month t. The bottom panel of the figure shows the distribution of destination degree centrality, aggregated over individuals, months (24 total), and potential districts (27 per individual). We observe that in the vast majority of these (individual $\times$ month $\times$ potential destinations) observations, the destination degree centrality is less than 3; in roughly 500,000 cases the individual has 10 contacts in the potential destination. Figure 4b shows the corresponding relationship between migration rates and the degree centrality of the individual's home network.[18]

Figure 4a thus validates a central thesis of prior research on networks and migration. Individuals with more contacts in a destination community are more likely to migrate to that community. We also see that this relationship is positive, monotic, and approximately linear. In other words, individuals with k times as many contacts in a destination district are k times more likely to migrate to that district. Figure 4b conversely indicates that individuals with more contacts in their home community are less likely to leave that community, but that there marginal effect of additional contacts at home is lower for individuals with a large number of contacts.

We observe a similar relationship for *weighted* degree, which reflects the intensive margin of communication, i.e., the total number of calls between the individual and his or her first-degree contacts. As shown in Figure 5a, individuals with a higher weighted degree in the destination are more likely to migrate, whereas individuals with a higher weighted degree at home are less likely to migrate (Figure 5c).

Figures 5a and 5c show the unconditional relationship between migration rate and weighted degree. For this analysis and much of what follows, we also find it instructive to analyze the relationship dbetween our independent variable of interest x (in this case, weighted degree) and the migration rate, after controlling for degree. Thus, the right column of subfigures in Figure 5 (as well as the right column of subfigures in Figures 7 - 6) indicates the conditional effect of x for individuals of a fixed degree k. More precisely, we construct Figure 5b by plotting the 20 β_k coefficients estimated by running 20 regressions of the form

$$migration_i = \alpha + \beta_k x_i + \epsilon_i \tag{12}$$

where a separate regression is estimated for each value of degree between 0 and 20. Positive values in Figure 5b indicate that, holding degree fixed, individuals with a higher weighted

[18] Note that the degree centrality distribution in the bottom panel of Figure 4b does not match that in the bottom panel of Figure 4a, since each individual has only one home district, but 27 potential destination districts.

degree are more likely to migrate. To faciliate comparison of the different β_k within a figure, we use the z-score of x at each fixed k when estimating equation (12), so that the coefficient β_k can be interpreted as the increase in migration rate assocated with a one standard deviation increase in x.[19]

We obseve in Figure 5b that, conditional on degree, the effect of weighted degree at the destination is somewhat ambiguous. For the vast majority of individuals who only have a few contacts in the potential destination (degree between 1 and 5), there is a small positive correlation between the intensity of communication with those contacts and the likelihood of migration. However, for individuals with a larger number of contacts, there is a much weaker, and sometimes weakly negative, association between the intensity of commication and the migration rate. A simliar pattern is observed in Figure 5d with respect to weighted degree at home: individuals with a small number of contacts are less likely to migrate if they interact with those contacts frequently, but individuals with a large number of contacts are not significantly more likely to migrate if they interact with those contacts frequently than similarly situated individuals who interact with the same number of contacts less frequently.

4.2 Information

In later analysis, we will take a more principled approach to estimating the relative importance of what we have loosely termed "information" and "cooperation" in model (6). First, we show the nonparametric relationships observed in the data, which in part motivated the structure in the model.

As defined in Section 3.2, we quantify information as the size of i's second-order network, i.e., the number of friends of i's friends. Figure 6a shows the general positive relationship between migration rate and information in the destination, while Figure 6c shows the opposite relationship for information at home. The shape of these curves resemble the relationship

[19] This standardization does not affect the sign or significance of the coefficients we estimate, only the magnitude. It is not strictly necessary, but helps net out the mechanical correlation between degree and x (for instance, that individuals the number of friends of friends increases super-linearly with the number of friends.

between migration rate and degree shown earlier in Figure 4: the average migration rate increases roughly linearly with information in the destination, and decreases monotonically but with diminishing returns relative to information at home.

Of course, our definition of information is mechanically correlated with degree, in that individuals with more friends are also likely to have more friends of friends. Thus, Figures 6b and 6d show the relationship between migration rate and information, holding degree fixed and re-estimating regression (12) separately for each degree. This result is more interesting, as we see that the likelihood of migrating does not generally increase with information in the destination, after holding destination degree fixed. This result is difficult to reconcile with most standard models of information diffusion, such as those proposed in Banerjee et al. (2013) and **?**. Indeed, much of the literature on migration and social networks seems to imply that, all else equal, individuals would be more likely to migrate if they have friends with many friends, as such networks would provide more natural conduits for information about job opportunities and the like. Our data provide little empirical support for this prediction, particularly for the vast majority of individuals who have only a small number of contacts in destination communities.

4.3 Cooperation

Finally, Corollary 1 of the strategic cooperation model presented in Section 2.2 predicts that all else equal, individuals will receive more utility from friends who share common friends. As described in Section 3.2, we measure this in our data as "support", i.e., the fraction of i's contacts who are also contacts with another of i's contacts, as originally proposed in Jackson, Rodriguez-Barraquer and Tan (2012).

Both at home and in the destination, the unconditional relationship between support and migration is ambiguous (Figures 7a and Figure 7c). However, this apparent null relationship obscures the fact that support is generally decreasing in degree; in other words, the larger an individual's network, the harder it is to maintain a constant level of support. Indeed, holding degree fixed, the role of support is more evident. Figure 7b indicates that for individuals with a fixed number of contacts in the destination, those whose contacts are mutually supported are significantly more likely to migrate. The converse effect is found in Figure 7d for support at home: holding degree fixed, people are less likely to leave home if their home contacts are more supported. As was the case with weighted degree, this effect is mostly observed for individuals with a modest number of contacts in their home community; for individuals with large home networks, support is not significantly associated with migration.[20]

5 Estimation and Results

[This section is a work in progress. A sketch of the reuslts are described below.]

The results shown thus far illustrate the correlations we observe between social network structure and migration rates. Most interesting, we observe that after conditioning on the number of immediate contacts in the destination, individuals whose contacts have more contacts in the destination are no more likely to migrate. By contrast, individuals with more tightly clustered destination networks are more likely to migrate. This pair of results are not so easily reconciled with a model in which migration decisions are determined entirely by access to information in the destination network, and motivate the model described in Section 2.3, where migrants balance both access to information and the potential for repeated cooperation. We turn now to estimating that full model.

5.1 Discrete choice model of migration

Mixed logit **?**, Conditional logit **??**, multinomial logit **?**. [21]

[20] Appendix Figure A1 shows that the closely related concept of network clustering looks identical to what we measure as support. Appendix Figure A2 shows the relationship between migration and "weighted support", i.e., the frequency of interactoin between supported contacts.

[21] Another possibility is to model the decision to migrate with a nested logit model, where the individual makes two independent decision: the first is whether or not to migrate and the second is, given the decision to move, the choice of destination **??**. We believe this approach is less appropriate to our context, as the decision to migrate is closely related to the possible destination choices – (**?**) provides a more complete discussion of this point.

5.2 Calibration

We begin by estimating Model (11). The structural parameters of primary interest are π^H, π^D, α^H, and α^D. To summarize, π indicates the importance of "information" relative to "cooperation." To calculate the information value of a network we use equation (1), and follow Banerjee et al. (2013) by setting q equal to the inverse of the first eigenvalue of the adjacency matrix, $\lambda_1(G)$, with $T = 7$.[22] The cooperation value of a network is defined by equation (7) as the weighted sum of supported and unsupported links in the network, where α indicates the weight at home (α^H) and in the destination (α^D).

We use maximum likelihood estimation over all possible combinations of possible parameters, and report the results in Table 8.[23] Our estimates of $\pi < 0.5$ indicate that both at home and in the destination, the potential for an individual to receive information from distant friends of friends is less important to migration decisions than network structures that are conducive to repeated cooperation. Within the cooperation model, the estimate of $\alpha < 0.5$ further emphasizes the point that contacts who are friends with each other are more important in migration decisions than contacts who are not connected to other friends of the migrant.

We stress that these results are preliminary, and we are currently running repeated simulations on a larger number of individuals to calibrate these parameters, and to construct confidence bands around the estimates. In ongoing work, we are also interested in understanding how these parameters differ for different types of migrants, such as rural versus urban migrants, temporary versus permanent migrants, and short-distance versus long-distance

[22] Banerjee et al. (2013) show that this approach to measuring diffusion centrality closely approximates a structural property of "communication centrality." This latter property could not be practically estimated on a network as large as the one we study.

[23] The current version fixes a rural home district H (Rwamagana) and a rural potential destination district D (Kayonza), then randomly draws 500 individuals who migrate from H to D and 500 individuals who remain in H in a single month of the data. For each possible set of parameters $< \pi, \alpha, \tau, \rho >$, we calculate the utility of the home and destination network for each migrant, and the total utility of migration. If the total utility of migration is positive, we predict that individual would migrate. We choose the set of parameters that minimizes the number of incorrect predictions. We define "all possible parameters" through a grid search over an empirically determined range: $0 \leq \pi \leq 1$; $0 \leq \alpha \leq 1$; $-100 \leq \tau \leq 100$; $-1000 \leq \rho \leq 1000$.

migrants. Thus far, we have produced versions of the "reduced form" results separately for urban and rural migrants. These are shown in Appendex Figures A3 - A5.

6 Robustness

The empirical results described above are robust to a large number of alternative specifications. In results available upon request, we have verified that our results are not affected by any of the following:

- **How we define "migration" (choice of k):** Our main specifications set $k = 2$, i.e., we say an individual has migrated if she spends 2 or more months in d and then 2 or more months in $d' \neq d$. We observe qualitatively similar results for $k = 1$ and $k = 3$.

- **How we define "migration" (home location sensitivity):** Our assignment of individuals to home locations is based on the set of mobile phone towers through wihch their communication is routed. Since there is a degree of noise in this process, we take a more restrictive definition of migration that only considers migrants that move between non-adjacent districts.

- **Definition of social network (reciprocated edges):** In constructing the social network from the mobile phone data, we normally consider an edge to exist between i and j if we observe one or more phone call or text message between these individuals. As a robustness check, we take a more restrictive definition of social network and only include edges if i initiates a call or sends a text message to j and j initiaties a call or sends a text message to i.

- **Definition of social network (strong ties):** We separately consider a definition of the social network that only includes edges where more than 3 interactions are observed between i and j. This is intended to address the concern that our estimates might be influenced by infrequent events such as misdialed numbers, text message spam and the like.

- **Definition of social network (ignore business hours):** To address the concern that our estimates may be picking up primarily on business-related contacts, and not the kinship networks commonly discussed in the literature, we only consider edges that are observed between the hours of 5pm and 9am.

- **Treatment of outliers (removing low- and high-degree individuals):** We remove from our sample all individuals (and calls made by individuals) with fewer than 3 contacts, or more than 500 contacts. The former is intended to address concerns that the large number of individuals with just one or two friends could bias linear regression estimates; the latter is intended to remove potential calling centers and businesses.

- **Sample Definition (single month):** We perform the analysis separately for each of the 24 months in the dataset, and do not aggregate over months. This ensures that an individual is not double-counted across time.

- **Sample Definition (single potential destination):** Instead of allowing each individual to consider 27 potential migration destinations, we choose that individual's most likely destination, and consider that to be the only potential destination for the migrant. This ensures that an individual is not double-counted within a given month.

6.1 Population representativeness and external validity

Our data allow us to observe the movement patterns and social network structures of a large population of mobile phone owners in Rwanda. These mobile subscribers represent a non-random subset of the overall Rwandan population. Likewise, the social network connections we observe for any given subscriber are assumed to be a partial and non-random subset of that subscriber's true social network.[24]

[24] This section is under revision, contact authors for details.

7 Conclusion

This paper presents new theory and evidence on the role that social network play in the decision to migrate. Our approach highlights how new sources of large-scale digital data can be used to simultaneously observe migration histories and the dynamic structure of social networks at a level of detail and scale that has not been achieved in prior work. These data make it possible to directly validate several long-standing assumptions in the literature on migration, which have been hard to test with traditional sources of data. For instance, we show that individuals are more likely to migrate to destinations where they have a large number of contacts, and that the elasticity of this response is approximately one (e.g., someone with 20 contacts in the destination is roughly twice as likely to migrate as someone with 10 contacts).

We also document several novel properties of the relationship between social networks and migration, not all which can be explained by simple models of information diffusion. For instance, we find that migration rates are not positively correlated with the number of friends of friends that one has in the destination, but that the migration rate is negatively correlated with the number of friends of friends at home. Similarly, we find significant and positive effects of having denser destination networks where friends are friends with each other. To reconcile these results, we propose a model of strategic cooperation that characterizes how individuals obtain value from their social network, and which captures many of the stylized features of our data.

We can imagine several directions to extend this analysis. First, while we focus on how social networks affect the decision to migrate, it is also likely that migration in turn affects network structure. For instance, individuals may strategically form links in anticipation of migration, and post-migration, may form different types of friendships, and let prior friendships go. We do our best to work around this endogeneity using several techniques discussed above, but we believe this process is of independent interest, though beyond the scope of this paper. Separately, subsequent analysis could more directly explore peer effects in migration, and how migrations cascade within local communities. Our analysis uncovered suggestive evidence that individuals are likely to follow the paths of prior migrants from their home community; again, a full treatment of this effect is beyond the scope of our current analysis, but could shed new light on the role of social networks in migration.

References

Ali, S. Nageeb, and David A. Miller. 2016. "Ostracism and Forgiveness." *American Economic Review*, 106(8): 2329–2348.

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. 2013. "The Diffusion of Microfinance." *Science*, 341(6144): 1236498.

Banerjee, Abhijit, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. 2014. "Gossip: Identifying Central Individuals in a Social Network." National Bureau of Economic Research Working Paper 20422.

Beaman, Lori A. 2012. "Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the US." *The Review of Economic Studies*, 79(1): 128–161.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2010. "Migration and Economic Mobility in Tanzania: Evidence from a Tracking Survey." *Review of Economics and Statistics*, 93(3): 1010–1033.

Blumenstock, Joshua Evan. 2012. "Inferring Patterns of Internal Migration from Mobile Phone Call Records: Evidence from Rwanda." *Information Technology for Development*, 18(2): 107–125.

Borjas, George J. 1992. "Ethnic Capital and Intergenerational Mobility." *The Quarterly Journal of Economics*, 107(1): 123–150.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica*, 82(5): 1671–1748.

Calv-Armengol, Antoni. 2004. "Job contact networks." *Journal of Economic Theory*, 115(1): 191–206.

Calv-Armengol, Antoni, and Matthew O. Jackson. 2004. "The Effects of Social Networks on Employment and Inequality." *The American Economic Review*, 94(3): 426–454.

Carletto, Calogero, Alan de Brauw, and Raka Banerjee. 2012. "Measuring migration in multi-topic household surveys." *Handbook of Research Methods in Migration*, 207.

Casciaro, Tiziana. 1998. "Seeing things clearly: social structure, personality, and accuracy in social network perception." *Social Networks*, 20(4): 331–351.

Chandrasekhar, Arun, Emily Breza, and Alireza Tahbaz-Salehi. 2016. "Seeing the forest for the trees? an investigation of network knowledge." *Working Paper.*

Chandrasekhar, Arun G., Cynthia Kinnan, and Horacio Larreguy. 2014. "Social networks as contract enforcement: Evidence from a lab experiment in the field." National Bureau of Economic Research.

Deshingkar, Priya, Sven Grimm, and International Organization for Migration. 2005. *Internal Migration and Development: A Global Perspective.* United Nations Publications. Google-Books-ID: r1jHX1nyKzYC.

Ghosh, Parikshit, and Debraj Ray. 1996. "Cooperation in Community Interaction Without Information Flows." *The Review of Economic Studies*, 63(3): 491–519.

Heath, Rachel. 2016. "Why do firms hire using referrals? Evidence from Bangladeshi Garment Factories."

Hu, Yifan. 2005. "Efficient, high-quality force-directed graph drawing." *Mathematica Journal*, 10(1): 37–71.

Jackson, Matthew O. 2010. *Social and Economic Networks.* Princeton University Press.

Jackson, Matthew O., and Leeat Yariv. 2010. "Diffusion, strategic interaction, and social structure." *Handbook of Social Economics, edited by J. Benhabib, A. Bisin and M. Jackson.*

Jackson, Matthew O., Tomas Rodriguez-Barraquer, and Xu Tan. 2012. "Social capital and social quilts: Network patterns of favor exchange." *The American Economic Review*, 102(5): 1857–1897.

Jensen, Robert. 2012. "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India." *The Quarterly Journal of Economics*, 127(2): 753–792.

Karlan, Dean, Markus Mobius, Tanya Rosenblat, and Adam Szeidl. 2009. "Trust and Social Collateral." *The Quarterly Journal of Economics*, 124(3): 1307 –1361.

Krackhardt, David. 1990. "Assessing the political landscape: Structure, cognition, and power in organizations." *Administrative science quarterly*, 342–369.

Lucas, Robert EB. 2015. "Internal Migration in Developing Economies: An Overview." *KNOMAD WORKING PAPER 6*.

McKenzie, David J., and Marcin J. Sasin. 2007. "Migration, Remittances, Poverty, and Human Capital: Conceptual and Empirical Challenges." Social Science Research Network SSRN Scholarly Paper ID 999482, Rochester, NY.

Morten, Melanie. 2015. "Temporary migration and endogenous risk sharing in village india."

Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market." *The Quarterly Journal of Economics*, 118(2): 549–599.

Munshi, Kaivan. 2014. "Community Networks and the Process of Development." *The Journal of Economic Perspectives*, 28(4): 49–76.

Munshi, Kaivan, and Mark Rosenzweig. 2016. "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap." *American Economic Review*, 106(1): 46–98.

Topa, Giorgio. 2001. "Social Interactions, Local Spillovers and Unemployment." *The Review of Economic Studies*, 68(2): 261–295.

Figures

Figure 1: Location of all mobile phone towers in Rwanda, circa 2008

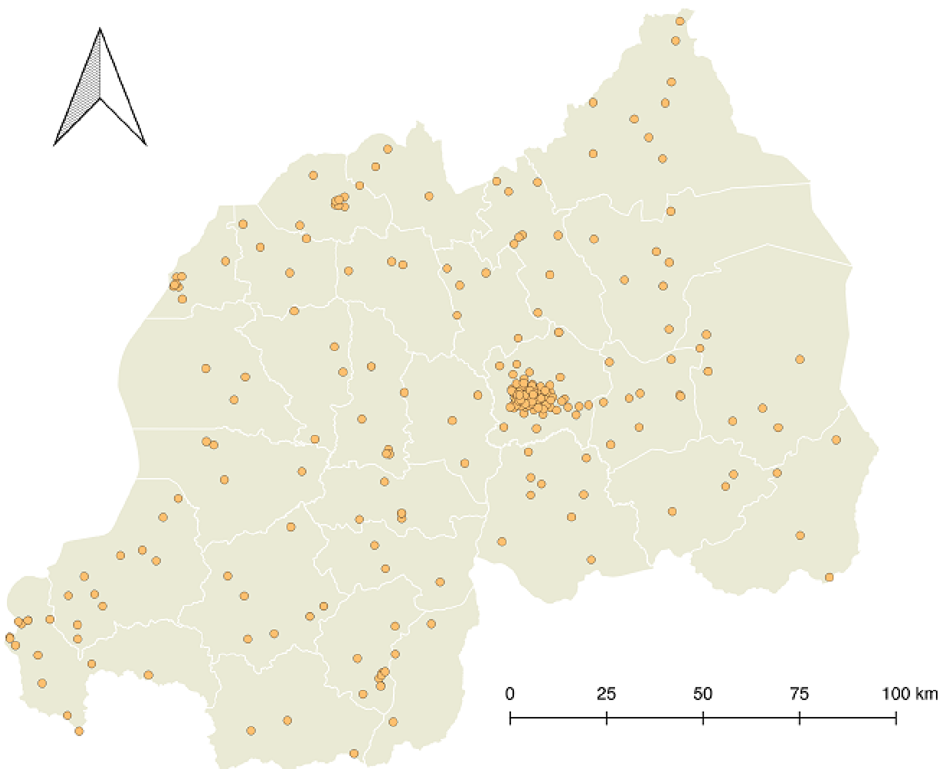


Figure 2: The social network of a single migrant

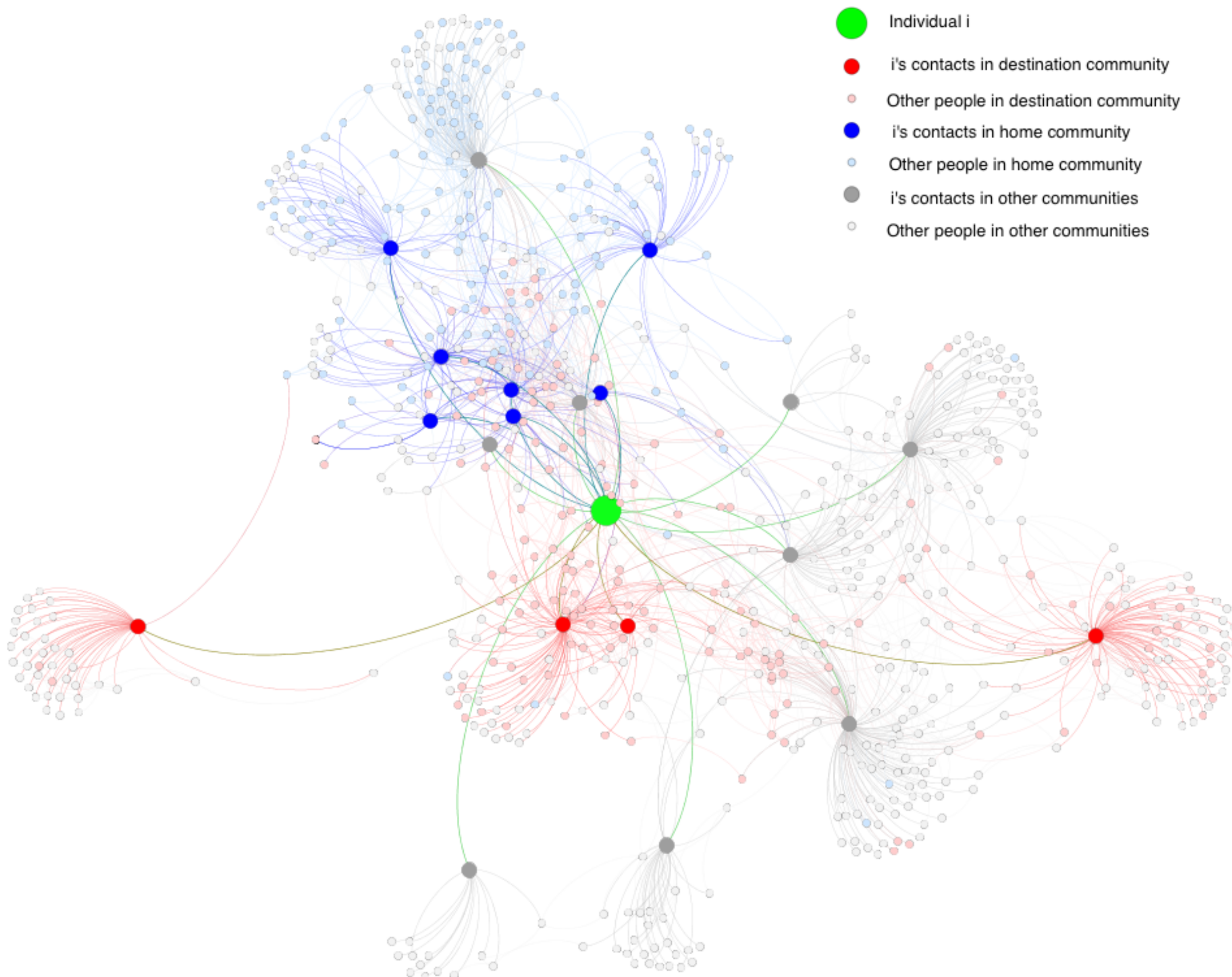


Figure 3: Population studied, by migration type

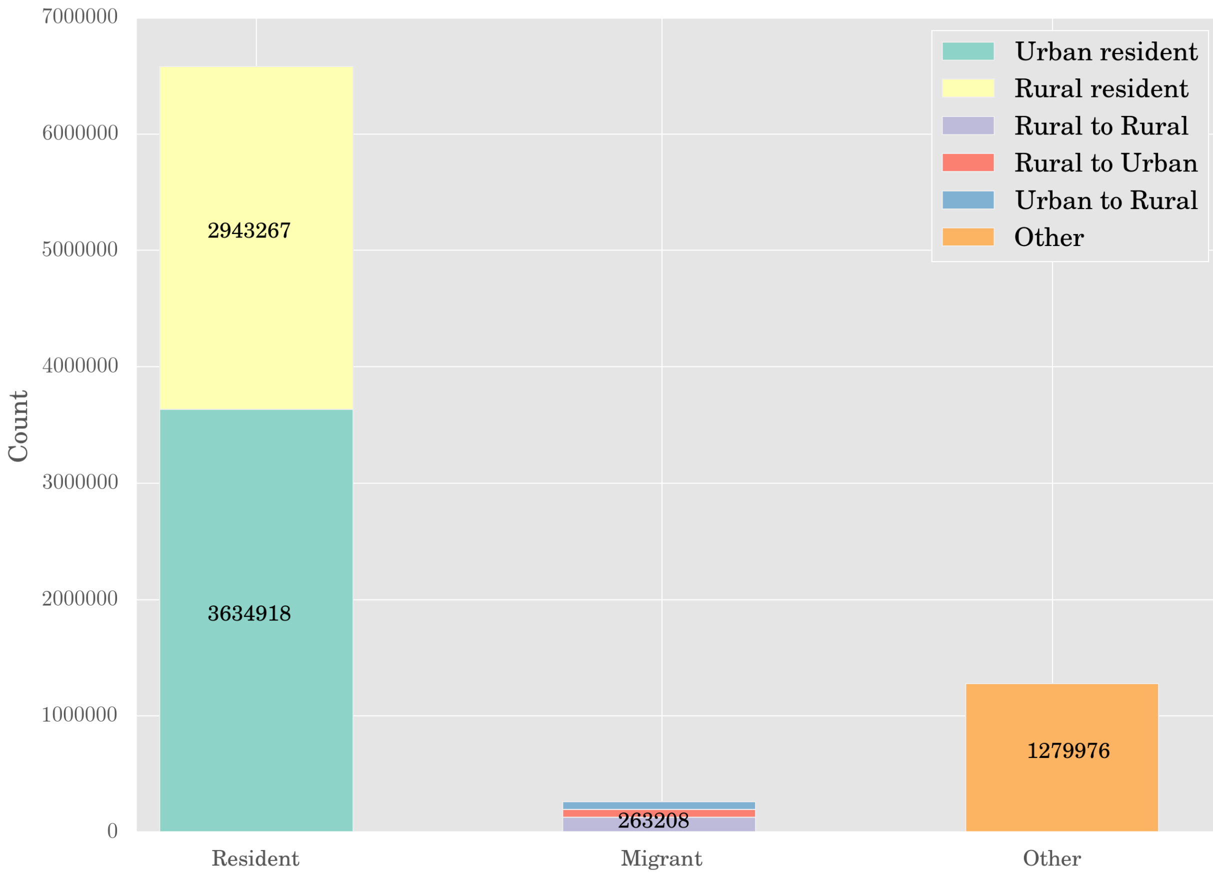
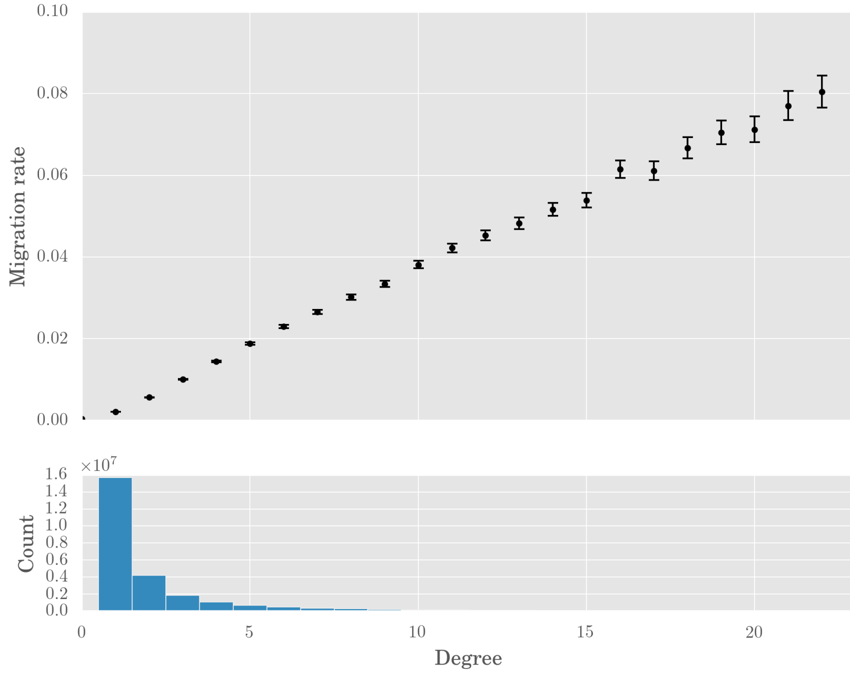
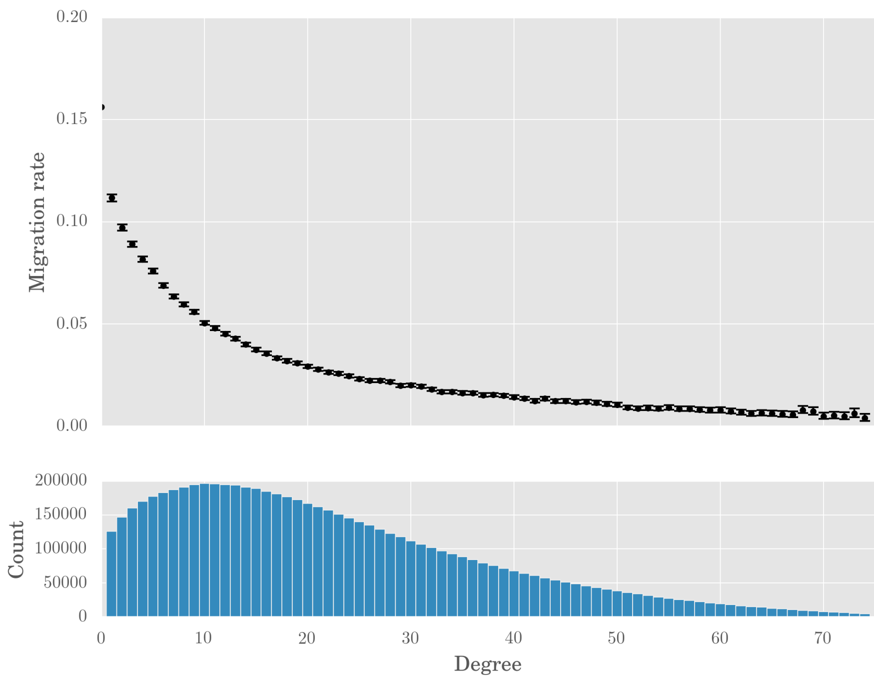


Figure 4: Migration rate and degree centrality (number of unique contacts in network)



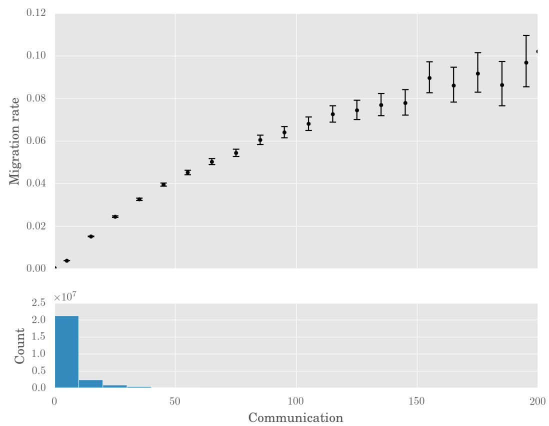
(a) Degree Centrality at Destination



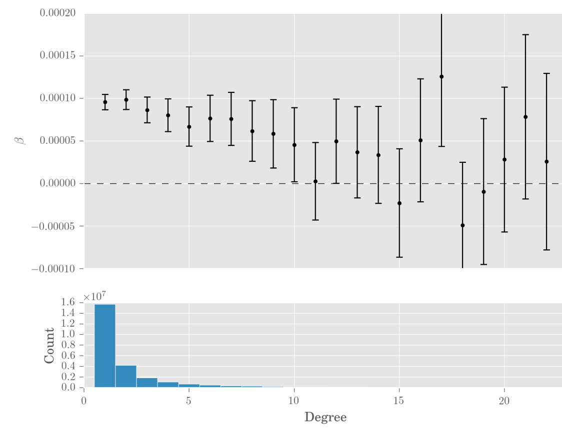
(b) Degree Centrality at Home

Notes: In both (a) and (b), the lower histogram shows the unconditional degree distribution, i.e., for each individual in each month, the total number of contacts in the (a) destination network and (b) home network. The upper figure shows, at each level of degree centrality, the average migration rate. Error bars indicate 95% confidence intervals, clustered by individual.

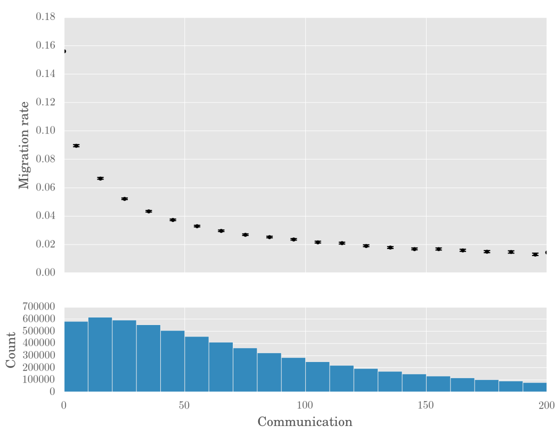
Figure 5: Relationship between migration rate and weighted degree



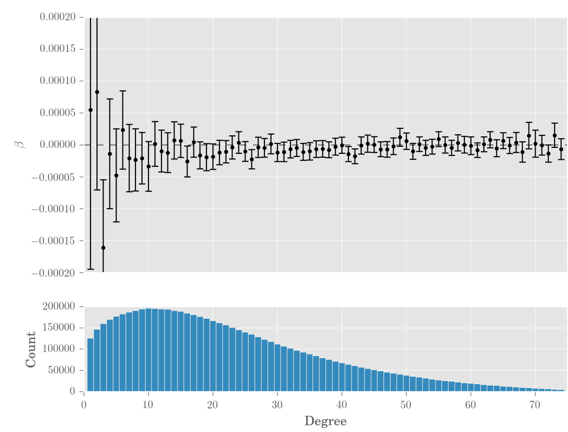
(a) Weighted Degree at Destination



(b) Weighted Degree at Destination, by Degree



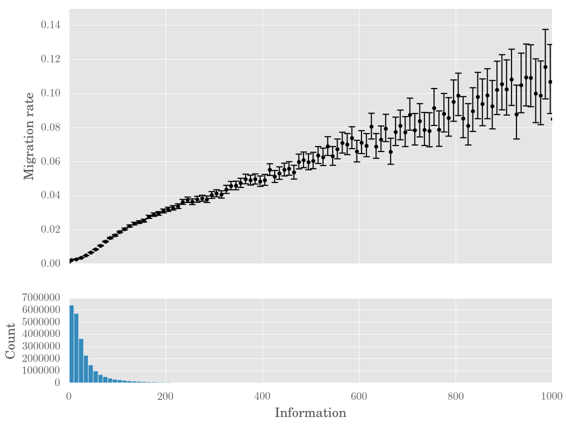
(c) Weighted Degree at Home



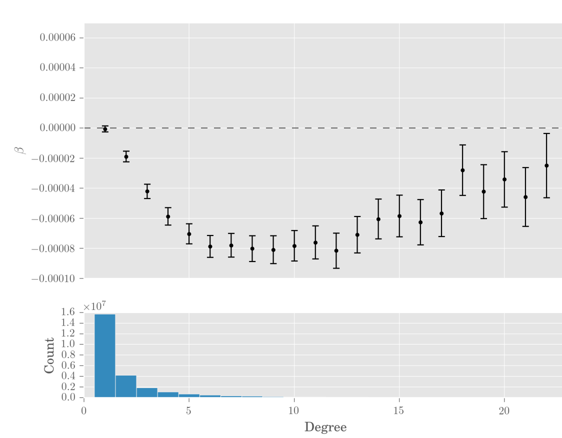
(d) Weighted Degree at Home, by Degree

Notes: "Weighted degree" denotes the number of calls between i and i's first-degree network. In all figures, the lower histogram shows the unconditional distribution of the x-variable. Top row (a and b) characterizes the destination network; bottom row (c and d) characterizes the home network. For the left column (a and c), the main figure indicates, at each level of weighted degree, the average migration rate. For the left column (b and d), the main figure indicates the correlation between the migration rate and weighted degree, holding degree fixed. In other words, each point represents the β_k coefficient estimated from a regression of $Migration_i = \alpha_k + \beta_k DegreeWeight_i$, estimated on the population of i who have degree equal to k. Error bars indicate 95% confidence intervals, clustered by individual.

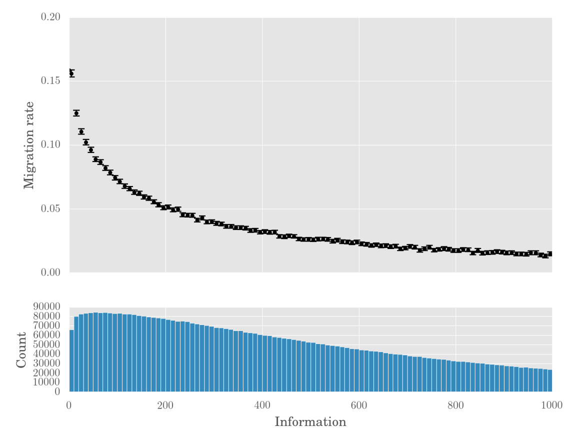
Figure 6: Relationship between migration rate and information



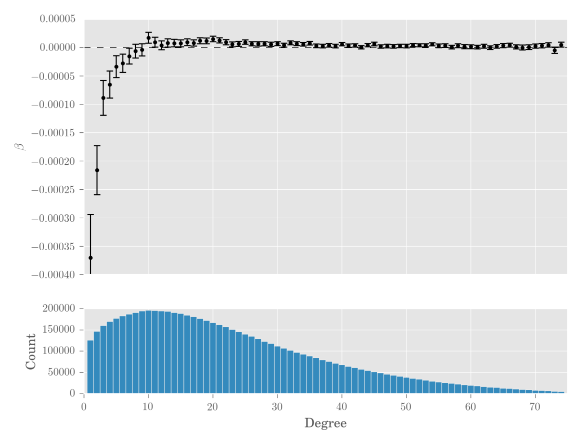
(a) Information at Destination



(b) Information at Destination, by Degree



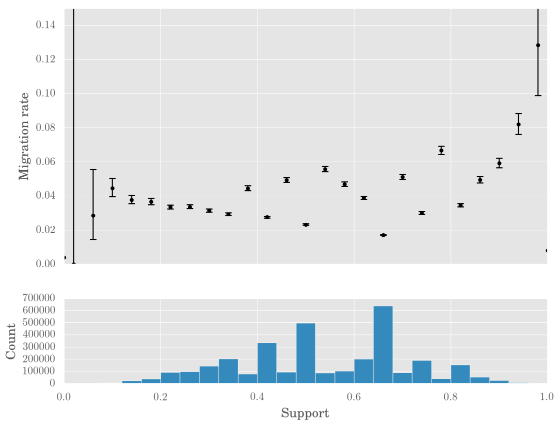
(c) Information at Home



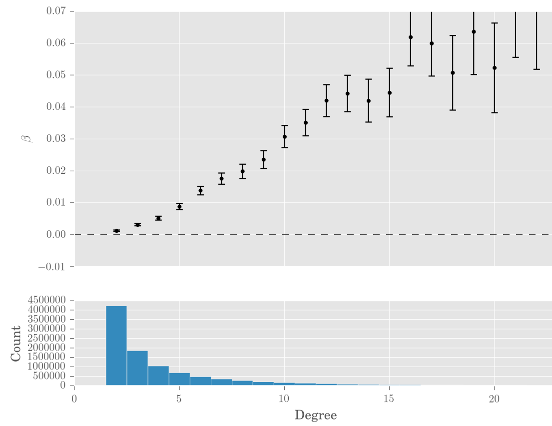
(d) Information at Home, by Degree

Notes: "Information" denotes the number of contacts if i's contacts. In all figures, the lower histogram shows the unconditional distribution of the x-variable. Top row (a and b) characterizes the destination network; bottom row (c and d) characterizes the home network. For the left column (a and c), the main figure indicates, at each level of weighted degree, the average migration rate. For the left column (b and d), the main figure indicates the correlation between the migration rate and information, holding degree fixed. In other words, each point represents the β_k coefficient estimated from a regression of $Migration_i = \alpha_k + \beta_k Information_i$, estimated on the population of i who have degree equal to k. Error bars indicate 95% confidence intervals, clustered by individual.

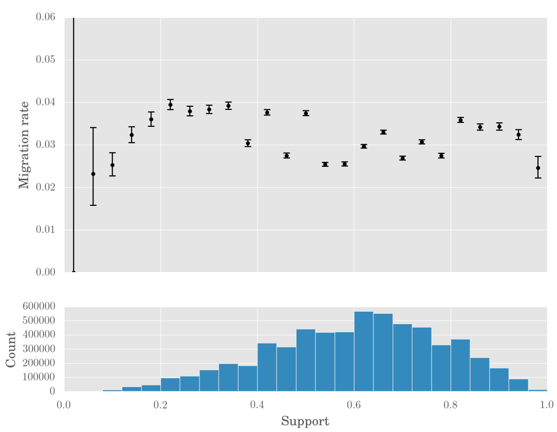
Figure 7: Relationship between migration rate and support



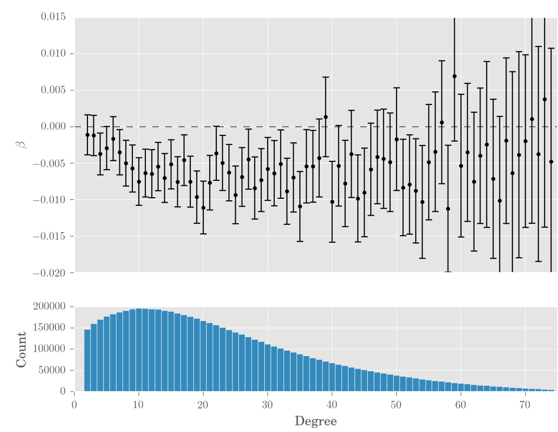
(a) Support at Destination



(b) Support at Destination, by Degree



(c) Support at Home



(d) Support at Home, by Degree

Notes: "Support" denotes the fraction of contacts supported by a common contact. In all figures, the lower histogram shows the unconditional distribution of the x-variable. Top row (a and b) characterizes the destination network; bottom row (c and d) characterizes the home network. For the left column (a and c), the main figure indicates, at each level of weighted degree, the average migration rate. For the left column (b and d), the main figure indicates the correlation between the migration rate and support, holding degree fixed. In other words, each point represents the β_k coefficient estimated from a regression of $Migration_i = \alpha_k + \beta_k Support_i$, estimated on the population of i who have degree equal to k. Error bars indicate 95% confidence intervals, clustered by individual.

Tables

Table 1: Summary statistics of mobile phone metadata

	In a single month (Jan 2008)	Over two years (Jul 2006 - Jun 2008)
Number of unique individuals	432,642	793,791
Number of person-months	432,642	8,121,369
Number of CDR transactions	50,738,365	868,709,410
Number of migrations	21,182	263,208
Number of rural-to-rural migrations	11,316	130,009
Number of rural-to-urban migrations	4,908	66,935
Number of urban-to-rural migrations	4,958	66,264

Table 2: Summary statistics: Migration

Migration Definition k	Total individuals	% Migrants	% Repeat migrants (to same district)	% Repeat migrants (to any district)	% I (n
1	1,087,229	29.751	9.615	18.870	
2	1,087,229	13.536	1.321	5.158	
3	1,087,229	6.653	0.196	1.379	
6	1,087,229	1.282	0.000	0.046	
12	1,087,229	0.014	0.000	0.000	

Notes: Table counts number of unique individuals meeting different definitions of a "migration event." Each row of the ta different k, such that an individual is considered a migrant if she spends k consecutive months in a district d and then k cons district $d' \neq d$ – see text for details.

Table 3: Single-variable OLS of migration rates on properties of destination network

	(1)	(2)	(3)	(4)	(5)
Destination degree	0.0035***				
	(0.0000)				
Destination weighted degree		0.0000334***			
		(0.0000052)			
Destination Support			0.0032073***		
			(0.0001008)		
Destination clustering				0.0011945***	
				(0.0000959)	
Destination information					-0.0000123
					(0.0000018)
N	184,717,611	10,087,878	10,087,878	10,087,878	10,087,878

Notes: Each column corresponds to a separate regression, where the dependent variable is a binary indicator of whethe month t. Column (1) treats every potential destination district (27 total) for each individual in each month as a separ (2)-(5), we restrict the analysis to individual-month-destination observations where the individual has one or more contact measures are undefined. In columns (2)-(5) we additionally standardize the independent variable separately for each le the text – this effectively controls for degree in the regression, and makes it possible to interpret the coefficient in st can be compared across degree. Independent variables are calculated using mobile phone data from month $t-1$, focu individual's (potential) destination network. Standard errors, clustered by individual, in parentheses. $^*p < 0.1$, $^{**}p < 0.$

Table 4: Single-variable OLS of migration rate on properties of home network

	(1)	(2)	(3)	(4)	(5)
Home degree	-0.0014***				
	(0.0000)				
Home weighted degree		-0.0000217***			
		(0.0000013)			
Home support			-0.0102658***		
			(0.0004360)		
Home clustering				-0.0145461***	
				(0.0010921)	
Home information					-0.0000062***
					(0.0000004)
N	6,841,393	6,192,588	6,192,588	6,192,588	6,192,588

Notes: Each column corresponds to a separate regression, where the dependent variable is a binary indicator of whethe month t. In columns (2)-(5) we normalize the independent variable separately for each level of degree – this effectivel regression, and makes it possible to interpret the coefficient in standard deviation units that can be compared across de are calculated using mobile phone data from month $t-1$, focusing on the structure of the individual's home network. S individual, in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table 5: OLS of migration rates on properties of destination network, controlling for degree and degree square

	(1)	(2)	(3)	(4)
Destination degree	4.489e-03***	4.502e-03***	4.582e-03***	4.726e-03***
	(2.162e-05)	(1.987e-05)	(1.953e-05)	(2.584e-05)
Destination degree square	-2.788e-05***	-2.611e-05***	-2.758e-05***	-2.692e-05***
	(6.928e-07)	(6.963e-07)	(6.932e-07)	(6.963e-07)
Destination weighted degree	2.904e-05***			
	(2.984e-06)			
Destination support		2.144e-03***		
		(1.124e-04)		
Destination clustering			3.458e-05	
			(1.385e-04)	
Destination information				-6.155e-06***
				(7.291e-07)
N	10,087,878	10,087,878	10,087,878	10,087,878

Notes:

Table 6: OLS of migration rates on properties of home network, controlling for degree and degree square

	(1)	(2)	(3)	(4)
Home degree	-2.411e-03***	-2.248e-03***	-2.513e-03***	-2.441e-03***
	(1.583e-05)	(1.553e-05)	(1.514e-05)	(1.736e-05)
Home degree square	2.532e-05***	2.281e-05***	2.549e-05***	2.541e-05***
	(2.218e-07)	(2.271e-07)	(2.245e-07)	(2.308e-07)
Home weighted degree	-1.883e-05***			
	(1.631e-06)			
Home support		-1.520e-02***		
		(3.333e-04)		
Home clustering			-6.564e-03***	
			(7.207e-04)	
Home information				-1.920e-06***
				(4.634e-07)
N	6,192,588	6,192,588	6,192,588	6,192,588

Notes:

Table 7: Parameter estimates of the structural model

Parameter (interpretation)	Estimate	(S.E.)
π (information, relative to cooperation)	0.1	(0.03)
α (unsupported links, relative to supported links)	0.4	(0.03)
τ (average fixed cost of migrating)	2.0	(0.23)

Notes: Maximum likelihood estimation of model (11) minimizes prediction error over a grid search of possible paramet the model parameters from k random draws of N migrants and N non-migrants are used to compute the standard error (in parentheses).

Table 8: Parameter estimates of the structural model

Parameter (interpretation)	Estimate	S.E.
π^H (home information, relative to support)	0.63	(0.07)
π^D (destination information, relative to support)	0.21	(0.06)
α^H (home unsupported links, relative to home supported links)	0.503	(0.05)
α^D (destination unsupported links, relative to destination supported links)	0.49	(0.03)
q (probability of passing information)	0.0085	(0.019)
τ (average fixed cost of migrating)	-0.132	(0.08)
ρ (importance of home network, relative to destination network)	1.03	(0.15)

Notes: Maximum likelihood estimation of model (11) minimizes prediction error over a grid search of possible parameters. Bootstrap estimates of the model parameters from k random draws of N migrants and N non-migrants are used to compute the standard errors of the parameter estimates (in parentheses).

A1 Proofs

Proof of Proposition 1: We will show that an agent, say i, has no profitable deviations under each of the following two cases when facing $j \in N_i$.

(Case 1) k is innocent for all $k \in N_i$

Consider the most profitable deviation for i: choose $a_{ij} = 0$ and $a_{ik} \in \{0, a_{ik}^G\}$ in any penalty stage game facing with any $k \in N_i$. However, from (3) we have

$$\begin{aligned} b(a^*_{m(ij)}) + 0 &= v(a^*_{m(ij)}) + \int_0^\infty e^{-rt}\lambda[v(a^*_{m(ij)}) + m(ij)v(a^*_1)]dt \\ &= v(a^*_{m(ij)}) + \frac{\lambda}{r}[v(a^*_{m(ij)}) + m(ij)v(a^*_1)] \\ &\leq v(a^*_{m(ij)}) + \frac{\lambda}{r}\sum_{k\in N_i} v(a^*_{m(ik)}), \end{aligned}$$

where the last inequality holds since $m(ik) \geq 1$ for all $k \in N_i \cap N_j$, and $|N_i| \geq m(ij) + 1$.

(Case 2) k is guilty for some $k \in N_i$

If i has profitable deviations in this case, then i will perform any of them after all of her neighbors become back to innocent, contradicting to what we have proved in previous case.

∎

Proof of Proposition 2: Let $a_{hl} \in \Sigma_G$ be the effort level chosen by h when meeting l. We construct $\Sigma_{G'}$ as follows. Let $a'_{ij} = a'_{ji} = a^* < a^*_0$ where $a^* > 0$ satisfies $c(a^*) < \int_0^\infty e^{-rt}\lambda v(a^*)dt$ (such a^* exists from assumption 1. and 2.), so the effort level is sustainable with positive net utility by the link itself. Also let $a'_{hl} = a_{hl}$ for all h and l such that $G(hl) = 1$ and $(h,l) \notin \{(i,k),(j,k) : k \in N_i \cap N_j\}$. For a'_{ik} and a'_{jk}, first note that

$$c(a_{ik}) \leq \sum_{h\in N_i(G)} \int_0^\infty e^{-rt}\lambda(b(a_{hi}) - c(a_{ih}))dt,$$

$$c(a_{jk}) \leq \sum_{h\in N_j(G)} \int_0^\infty e^{-rt}\lambda(b(a_{hj}) - c(a_{jh}))dt,$$

$$c(a_{kl}) \leq \sum_{h \in N_k(G)} \int_0^\infty e^{-rt}\lambda(b(a_{hk}) - c(a_{kh}))dt \ \text{ for } l \in \{i,j\}.$$

Consider $\varepsilon > 0$ such that $C_l < \int_0^\infty e^{-rt}\lambda\Big(v(a^*) + \sum_{h \in N_i \cap N_j} \big(c(a_{lh}) - c(a_{lh}+\varepsilon)\big)\Big)dt$ where $C_l \in \{c(a^*), c(a_{lk}+\varepsilon) - c(a_{lk})\}$ and $l \in \{i,j\}$. Choosing $a'_{ik} = a_{ik} + \varepsilon$, $a'_{jk} = a_{jk} + \varepsilon$, $a'_{ki} = a_{ki}$, and $a'_{kj} = a_{ki}$ satisfies the incentive constraints among i, j, and those k's with $cl_k(G') > cl_k(G)$:

$$\begin{aligned}
c(a'_{ik}) &< c(a_{ik}) + \int_0^\infty e^{-rt}\lambda\Big(v(a^*) + \sum_{h \in N_i \cap N_j} \big(c(a_{ih}) - c(a'_{ih})\big)\Big)dt \\
&\leq \int_0^\infty e^{-rt}\lambda\Big(\sum_{h \in N_i(G)} \big(b(a_{hi}) - c(a_{ih})\big) + v(a^*) + \sum_{h \in N_i \cap N_j} \big(c(a_{ih}) - c(a'_{ih})\big)\Big)dt \\
&= \sum_{h \in N_i(G')} \int_0^\infty e^{-rt}\lambda(b(a'_{hi}) - c(a'_{ih}))dt, \\
c(a'_{jk}) &\leq \sum_{h \in N_j(G')} \int_0^\infty e^{-rt}\lambda(b(a'_{hj}) - c(a'_{jh}))dt, \\
c(a'_{ij}) &= c(a'_{ji}) = c(a^*) < \int_0^\infty e^{-rt}\lambda\Big(v(a^*) + \sum_{k \in N_i \cap N_j} \big(c(a_{ik}) - c(a'_{ik})\big)\Big)dt \\
&\leq \int_0^\infty e^{-rt}\lambda\Big(v(a^*) + \sum_{k \in N_i \cap N_j} \big(c(a_{ik}) - c(a'_{ik})\big) + \sum_{h \in N_i(G)} (b(a_{hi}) - c(a_{ih}))\Big)dt \\
&= \int_0^\infty e^{-rt}\lambda v(a^*)dt + \sum_{h \in N_i(G)} \int_0^\infty e^{-rt}\lambda(b(a'_{hi}) - c(a'_{ih}))dt,
\end{aligned}$$

$$\begin{aligned}
c(a'_{kl}) = c(a_{kl}) &\leq \sum_{h \in N_k(G)} \int_0^\infty e^{-rt}\lambda(b(a_{hk}) - c(a_{kh}))dt \\
&< \sum_{h \in N_k(G')} \int_0^\infty e^{-rt}\lambda(b(a'_{hk}) - c(a'_{kh}))dt,
\end{aligned}$$

where $l \in \{i,j\}$ and $a'_{ik} > a_{ik}$ and $a'_{jk} > a_{jk}$ yields the last (strict) inequality. For all other remaining incentive constraints, they are satisfied since their counterparts in Σ_G hold by definition. To complete the construction of $\Sigma_{G'}$, if someone deviated, then all her neighbors use the social norm with sequential move described before Proposition 1 to punish the deviator.

By the construction of this new equilibrium $\Sigma_{G'}$, we have

$$\begin{aligned} u_k(\Sigma_{G'}) &= \sum_{h\in N_k(G')} \int_0^\infty e^{-rt}\lambda(b(a'_{hk}) - c(a'_{kh})))dt \\ &> \sum_{h\in N_k(G)} \int_0^\infty e^{-rt}\lambda(b(a_{hk}) - c(a_{kh})))dt = u_k(\Sigma_G). \end{aligned}$$

for all $k \in N_i \cap N_j$. ▮

Proof of Proposition 3: For notational convenience, we define $\Delta = \lambda'_{ij} - \lambda_{ij}$, $N_{ij} = N_i \cap N_j$, $v_i = b(a_{ji}) - c(a_{ij})$, and $v_j = b(a_{ij}) - c(a_{ji})$. Let $k \in N_{ij}$ and $a_{hl} \in \Sigma_\lambda$ be the effort level chosen by h when meeting l. We construct $\Sigma_{\lambda'}$ as follows. Without loss of generality, assume $v_i > 0$. (If $v_i = 0$, we can sustain an equilibrium by increasing both a_{ij} and a_{ji} to make the assumption hold without decreasing any agent's utility). Also let $a'_{hl} = a_{hl}$ for all h and l such that $G(hl) = 1$ and $(h,l) \notin \{(i,k),(j,k) : k \in N_i \cap N_j\}$. For a'_{ik} and a'_{jk}, first note that

$$c(a_{ik}) \leq \sum_{h\in N_i(G)} \int_0^\infty e^{-rt}\lambda_{ih}(b(a_{hi}) - c(a_{ih}))dt,$$

$$c(a_{jk}) \leq \sum_{h\in N_j(G)} \int_0^\infty e^{-rt}\lambda_{jh}(b(a_{hj}) - c(a_{jh}))dt,$$

$$c(a_{kl}) \leq \sum_{h\in N_k(G)} \int_0^\infty e^{-rt}\lambda_{kh}(b(a_{hk}) - c(a_{kh}))dt \quad \text{for } l \in \{i,j\}.$$

If $v_j \leq 0$, then choose $\eta \geq 0$ such that $a'_{ij} \triangleq a_{ij} + \eta = a_{ji}$ with $v'_i = b(a_{ji}) - c(a'_{ij}) > 0$ (Such η exists since $v_i + v_j > 0$); if not, then let $\eta = 0$. Given η, there exists $\varepsilon > 0$ such that

$$\int_0^\infty e^{-rt}\Big(\Delta v'_i + \sum_{h\in N_{ij}} \lambda_{ih}\big(c(a_{ih}) - c(a_{ih}+\varepsilon)\big) + \lambda_{ij}(c(a_{ij}) - c(a'_{ij}))\Big)dt \geq c(a_{ij}+\eta) - c(a_{ij}),$$

$$\int_0^\infty e^{-rt}\Big(\Delta v'_i + \sum_{h\in N_{ij}} \lambda_{ih}\big(c(a_{ih}) - c(a_{ih}+\varepsilon)\big) + \lambda_{ij}(c(a_{ij}) - c(a'_{ij}))\Big)dt \geq c(a_{ik}+\varepsilon) - c(a_{ik}),$$

for all $k \in N_{ij}$. Define $a'_{ik} = a_{ik} + \varepsilon$, $a'_{jk} = a_{jk}$, $a'_{ki} = a_{ki}$, $a'_{kj} = a_{kj}$, and $a'_{ji} = a_{ji}$, then the incentive constraints among i, j, and those k's in N_{ij} are satisfied:

$$\begin{aligned}
c(a'_{ik}) &\leq c(a_{ik}) + \int_0^\infty e^{-rt}\Big(\Delta v'_i + \sum_{h \in N_{ij}} \lambda_{ih}\big(c(a_{ih}) - c(a'_{ih})\big) + \lambda_{ij}(c(a_{ij}) - c(a'_{ij}))\Big)dt \\
&\leq \int_0^\infty e^{-rt}\Big(\sum_{h \in N_i} \lambda_{ih}\big(b(a_{hi}) - c(a_{ih})\big) + \Delta v'_i + \sum_{h \in N_{ij} \cup \{j\}} \lambda_{ih}\big(c(a_{ih}) - c(a'_{ih})\big)\Big)dt \\
&= \sum_{h \in N_i} \int_0^\infty e^{-rt} \lambda'_{ih}(b(a'_{hi}) - c(a'_{ih}))dt, \\
c(a'_{ij}) &\leq \sum_{h \in N_i} \int_0^\infty e^{-rt} \lambda'_{ih}(b(a'_{hi}) - c(a'_{ih}))dt, \\
c(a'_{jk}) &= c(a_{jk}) \leq \sum_{h \in N_j} \int_0^\infty e^{-rt} \lambda_{jh}(b(a_{hj}) - c(a_{jh}))dt \leq \sum_{h \in N_j} \int_0^\infty e^{-rt} \lambda'_{jh}(b(a'_{hj}) - c(a'_{jh}))dt, \\
c(a'_{ji}) &= c(a_{ji}) \leq \sum_{h \in N_j} \int_0^\infty e^{-rt} \lambda'_{jh}(b(a'_{hj}) - c(a'_{jh}))dt, \\
c(a'_{kl}) &= c(a_{kl}) \leq \sum_{h \in N_k} \int_0^\infty e^{-rt} \lambda_{kh}(b(a_{hk}) - c(a_{kh}))dt \\
&\leq \sum_{h \in N_k} \int_0^\infty e^{-rt} \lambda_{kh}(b(a'_{hk}) - c(a'_{kh}))dt \\
&= \sum_{h \in N_k} \int_0^\infty e^{-rt} \lambda'_{kh}(b(a'_{hk}) - c(a'_{kh}))dt,
\end{aligned}$$

where $l \in \{i, j\}$. For all other remaining incentive constraints, they are satisfied since their counterparts in Σ_λ hold by definition. From this new equilibrium $\Sigma_{\lambda'}$, we have

$$\begin{aligned}
u_k(\Sigma_{\lambda'}) &= \sum_{h \in N_k} \int_0^\infty e^{-rt} \lambda'_{kh}(b(a'_{hk}) - c(a'_{kh}))dt \\
&> \sum_{h \in N_k} \int_0^\infty e^{-rt} \lambda_{kh}(b(a_{hk}) - c(a_{kh}))dt = u_k(\Sigma_\lambda).
\end{aligned}$$

for all $k \in N_{ij} \cup \{j\}$; particularly for i, we have

$$\begin{aligned} u_i(\Sigma_{\lambda'}) &= \sum_{h \in N_i} \int_0^\infty e^{-rt} \lambda'_{ih}(b(a'_{hi}) - c(a'_{ih}))dt \\ &= \int_0^\infty e^{-rt} \Delta v'_i dt + \sum_{h \in N_i} \int_0^\infty e^{-rt} \lambda_{ih}(b(a'_{hi}) - c(a'_{ih}))dt \\ &= \int_0^\infty e^{-rt} \Delta v'_i dt + \sum_{h \in N_{ij} \cup \{j\}} \int_0^\infty e^{-rt} \lambda_{ih}(c(a_{ih}) - c(a'_{ih}))dt + u_i(\Sigma_\lambda) \\ &> u_i(\Sigma_\lambda). \end{aligned}$$

∎

A2 Algorithms

Data: $< ID, datetime, location >$ tuples for each mobile phone interaction
Result: $< ID, month, district >$ tuples indicating monthly modal district
Step 1 Find each subscriber's most frequently visited tower;
$\rightarrow$ Calculate *overall daily modal districts*;
$\rightarrow$ Calculate *overall monthly modal districts*;
Step 2 calculate the *hourly modal districts*;
if *tie districts exit* **then**
    if *overall daily modal districts can resolve* **then**
        return the district with larger occurance number;
        else
            if *overall monthly modal districts can resolve* **then**
                return the district with larger occurance number
            end
        end
    end
end
Step 3 calculate the *daily modal districts*;
if *tie districts exit* **then**
    if *overall daily modal districts can resolve* **then**
        return the district with larger occurance number;
        else
            if *overall monthly modal districts can resolve* **then**
                return the district with larger occurance number
            end
        end
    end
end
Step 4 calculate the *monthly modal districts*;
if *tie districts exit* **then**
    if *overall monthly modal districts can resolve* **then**
        return the district with larger occurance number;
    end
end

Algorithm 1: Home location assignment

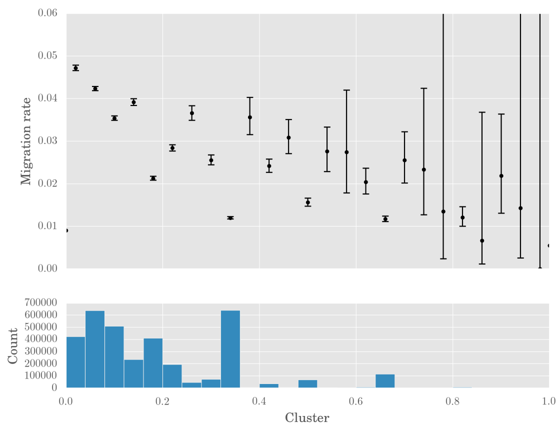
Data: Monthly modal district for four consecutive months: D_1, D_2, D_3, D_4
Result: Migration type

```
if D1 == D2 AND D3 == D4 then
    if D2 == D3 then
        if D4 == Kigali then
            migration type is urban resident
        end
        else
            migration type is rural resident
        end
    end
    else
        if D4 == Kigali then
            migration type is rural to urban
        end
        else
            if D1 == Kigali then
                migration type is urban to rural
            end
            else
                migration type is rural to rural
            end
        end
    end
end
else
    migration type is other
end
```

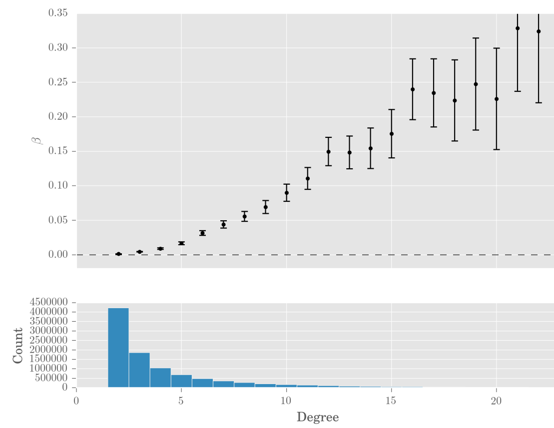
Algorithm 2: Classifying individuals by migrant type for k=2

A3 Appendix Figures and Tables

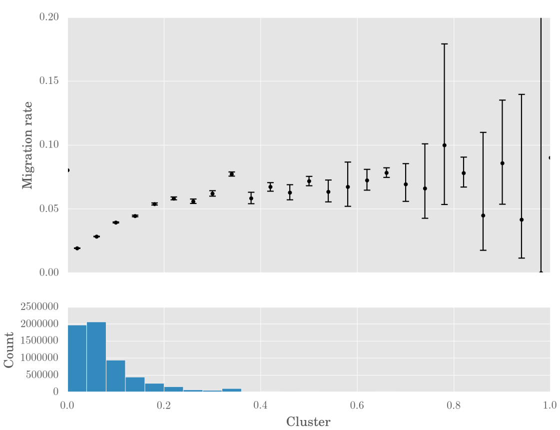
Figure A1: Relationship between migration rate and clustering



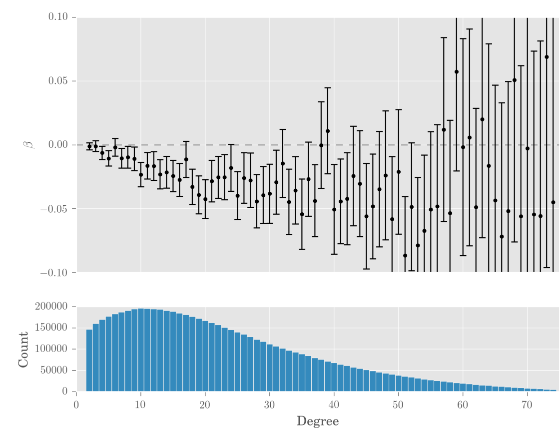
(a) Clustering at Destination



(b) Clustering at Destination, by Degree



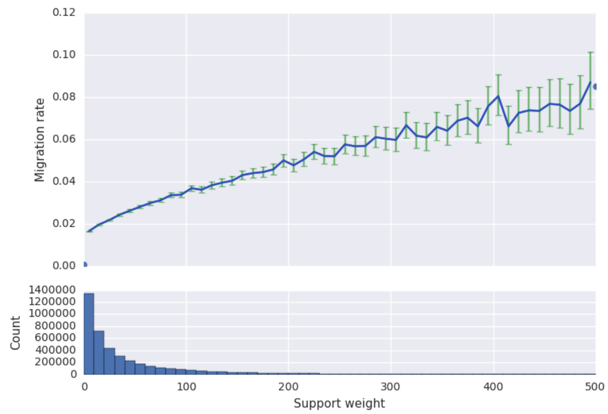
(c) Clustering at Home



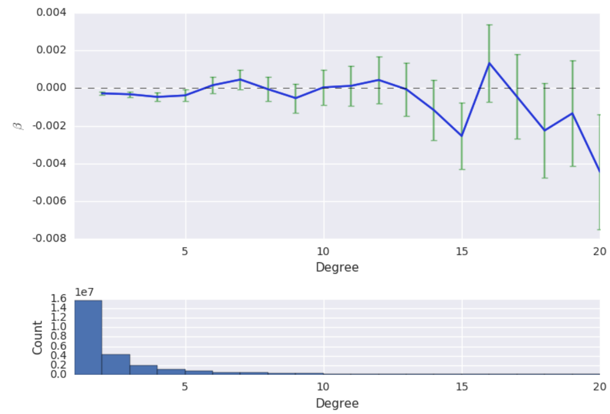
(d) Clustering at Home, by Degree

Notes: "Clustering" denotes the proportion of potential links between i's friends that exist. In all figures, the lower histogram shows the unconditional distribution of the x-variable. Top row (a and b) characterizes the destination network; bottom row (c and d) characterizes the home network. For the left column (a and c), the main figure indicates, at each level of weighted degree, the average migration rate. For the left column (b and d), the main figure indicates the correlation between the migration rate and clustering, holding degree fixed. In other words, each point represents the β_k coefficient estimated from a regression of $Migration_i = \alpha_k + \beta_k Clustering_i$, estimated on the population of i who have degree equal to k. Error bars indicate 95% confidence intervals, clustered by individual.

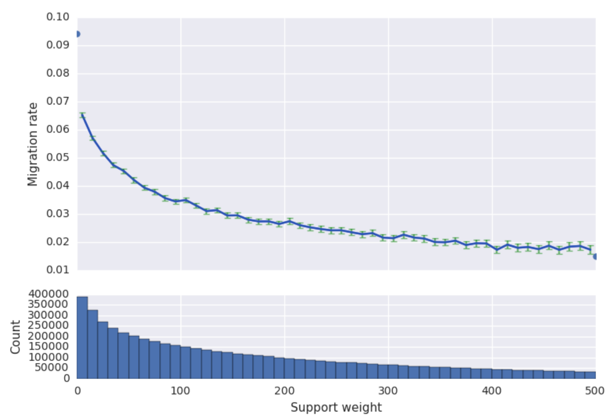
Figure A2: Relationship between migration rate and weighted support



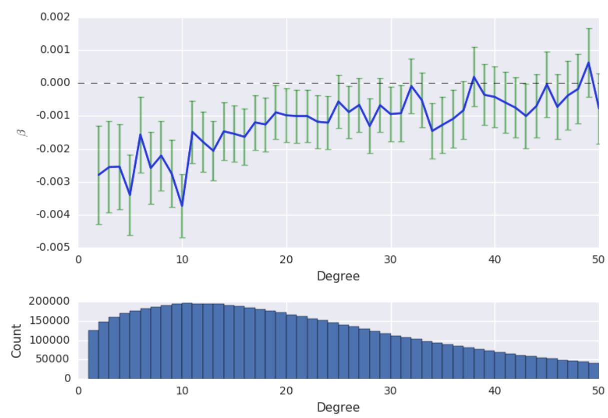
(a) Weighted Support at Destination



(b) Weighted Support at Destination, by Degree



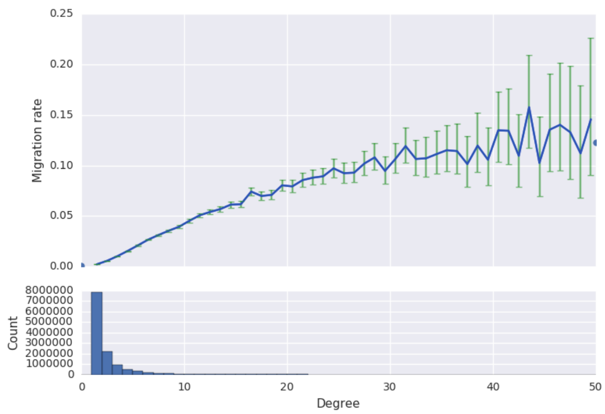
(c) Weighted Support at Home



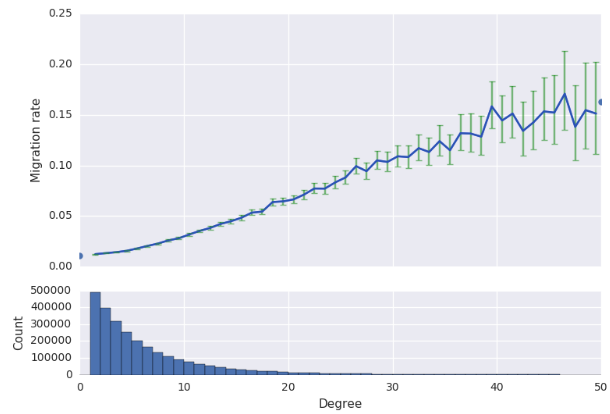
(d) Weighted Support at Home, by Degree

Notes: “Support Weight” denotes the frequency of interaction between supported contacts. In all figures, the lower histogram shows the unconditional distribution of the x-variable. Top row (a and b) characterizes the destination network; bottom row (c and d) characterizes the home network. For the left column (a and c), the main figure indicates, at each level of weighted degree, the average migration rate. For the left column (b and d), the main figure indicates the correlation between the migration rate and support weight, holding degree fixed. In other words, each point represents the β_k coefficient estimated from a regression of $Migration_i = \alpha_k + \beta_k SupportWeight_i$, estimated on the population of i who have degree equal to k. Error bars indicate 95% confidence intervals, clustered by individual.

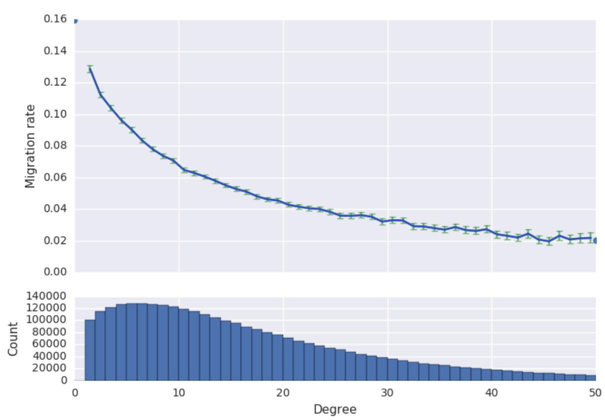
Figure A3: Urban vs. rural relationship between migration rate and degree centrality



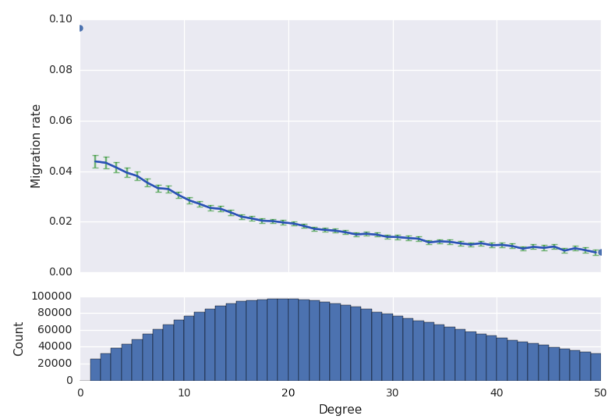
(a) Destination Degree (Rural Destinations)



(b) Destination Degree (Urban Destinations)

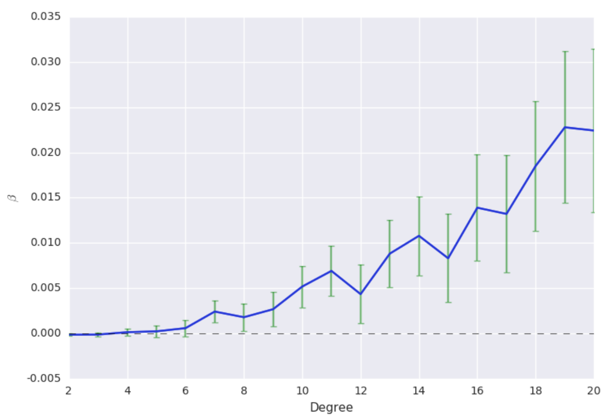


(c) Home Degree (Rural Home)

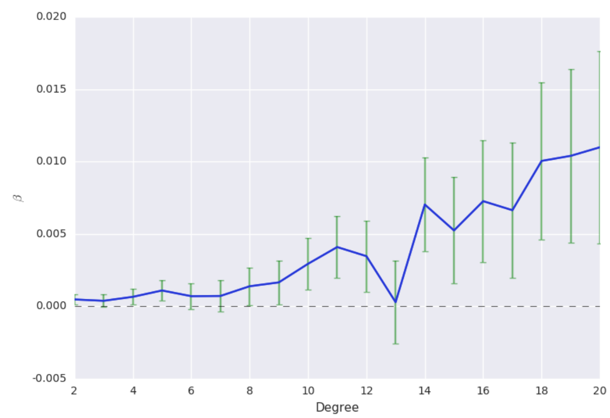


(d) Home Degree (Urban Home)

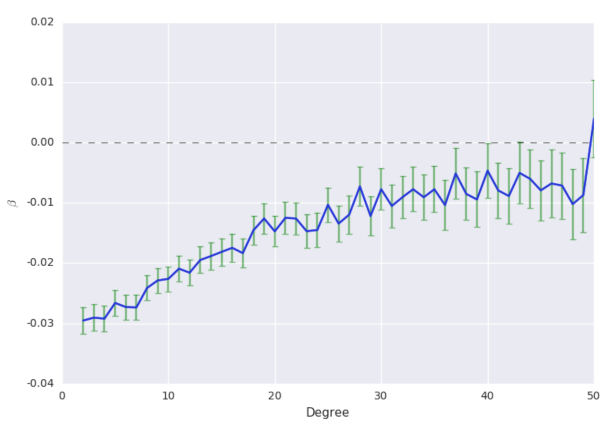
Figure A4: Urban vs. rural relationship between migration rate and information, holding degree fixed



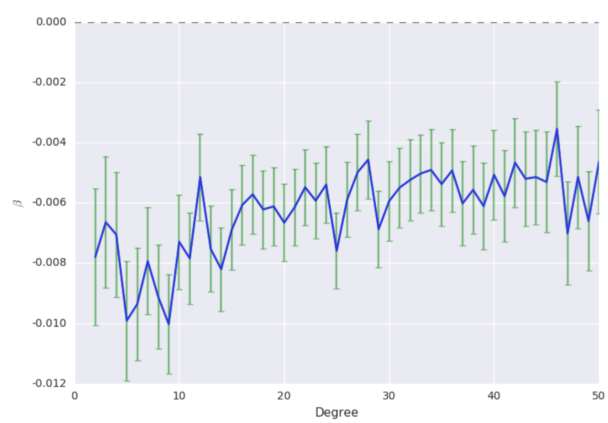
(a) Destination Information (Rural Destinations)



(b) Destination Information (Urban Destinations)

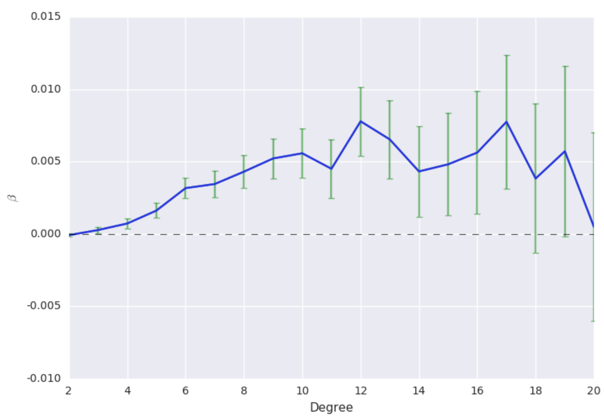


(c) Home Information (Rural Home)

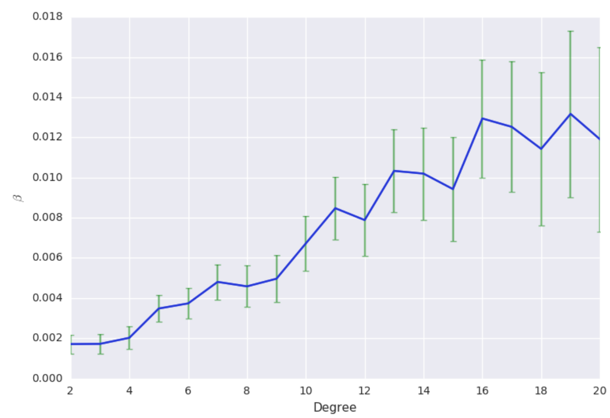


(d) Home Information (Urban Home)

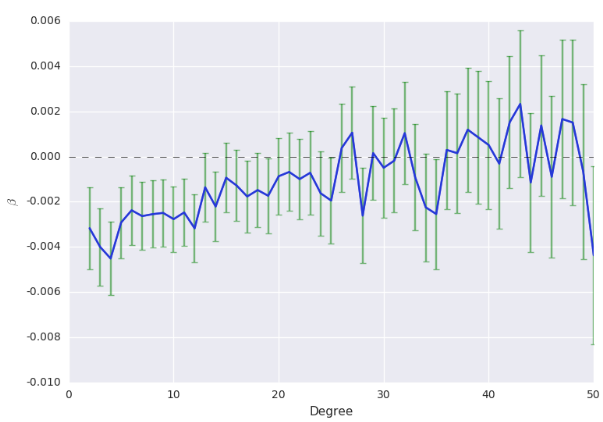
Figure A5: Urban vs. rural relationship between migration rate and support, holding degree fixed



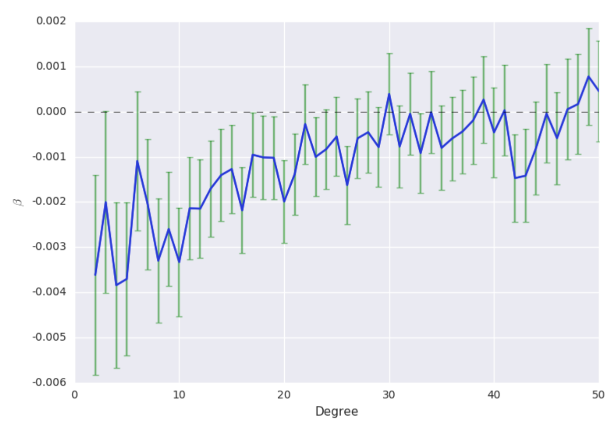
(a) Destination Support (Rural Destinations)



(b) Destination Support(Urban Destinations)



(c) Home Support (Rural Home)



(d) Home Support (Urban Home)