Migration and the Value of Social Networks*

Joshua Blumenstock† — U.C. Berkeley
Guanghua Chi‡ — U.C. Berkeley
Xu Tan§ — University of Washington

June 2023

Abstract

How do social networks influence the decision to migrate? Prior work suggests two distinct mechanisms that have historically been difficult to differentiate: as a conduit of information, and as a source of social and economic support. We disentangle these mechanisms using a massive 'digital trace' dataset that allows us to observe the migration decisions made by millions of individuals over several years, as well as the complete social network of each person in the months before and after migration. These data allow us to establish a new set of stylized facts about the relationship between social networks and migration. Our main analysis indicates that the average migrant derives more social capital from 'interconnected' networks that provide social support than from 'extensive' networks that efficiently transmit information.

JEL classification: O15, R23, D85, Z13, O12, C55

Keywords: Networks, Migration, Social Networks, Social Capital, Big Data, Development

*We thank Oriana Bandiera, Lori Beaman, Francis Bloch, Thomas Chaney, Margherita Comola, Ben Faber, Marcel Fafchamps, Andrew Foster, Matthew Jackson, David Miller, Kaivan Munshi, Mark Rosenzweig, Sebastian Sotelo, Yves Zenou, as well as seminar participants at Stanford, Brown, Paris School of Economics, U.C. Berkeley, the BREAD/CEPR Conference on Development Economics, 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 GLM-LIC Grant C2-RA4-205 and UW RRF #65-7397. This research was approved by the U.C. Berkeley Committee for Protection of Human Subjects, protocol #042196-001.

†University of California, Berkeley, jblumenstock@berkeley.edu
‡University of California, Berkeley, guanghua@berkeley.edu
§University of Washington, tanxu@uw.edu

1 Introduction

The decision to migrate is one of the most important economic decisions an individual can make. Many factors influence this decision, from employment prospects and amenity differentials to life-cycle considerations and migration costs. In each of these factors, social networks play a prominent role. It is through social networks that migrants learn about opportunities and conditions in potential destinations; at home, the structure of migrants' social networks shapes their ability and desire to leave.

This paper uses a rich source of digital data to add considerable nuance to our understanding of *how* social networks influence an individual's decision to migrate. Here, prior work emphasizes two distinct mechanisms: first, that networks provide migrants with access to information, for instance about jobs and conditions in the destination (Borjas, 1992, Topa, 2001, Munshi, 2003, Dustmann et al., 2016); and second, that networks act as a safety net for migrants by providing material or social support (Carrington et al., 1996, Edin et al., 2003, Dolfin and Genicot, 2010, Munshi, 2014, Comola and Mendola, 2015). However, there is considerable ambiguity about the nature and relative importance of these two mechanisms. For instance, the prevailing view in the migration literature is that migrants tend to go to places where they have larger networks, but a handful of studies argue that larger networks may actually deter migration, for instance if migrants compete with one another over opportunities and resources.[1] Similarly, robust risk sharing networks can both facilitate migration by providing informal insurance against negative outcomes (Morten, 2019), and discourage migration if migrants fear those left behind will be sanctioned for their departure (Munshi and Rosenzweig, 2016, Banerjee and Newman, 1998).

These ambiguities arise in part because it is difficult to link social network structure to migration decisions using traditional data (Chuang and Schechter, 2015). Instead, most existing work relies on indirect proxies for a migrant's social network, such as the assumption that individuals from the same hometown, or with similar observable characteristics, are more likely to be connected than two dissimilar individuals.[2] Such proxies can provide a

[1]Classic papers documenting the 'prevailing' view include Rees (1966), Greenwood (1969), Granovetter (1973), Montgomery (1991), and Borjas et al. (1992). More recent examples include Munshi (2003), Winters et al. (2001), Dolfin and Genicot (2010), Patel and Vella (2012), Fafchamps and Shilpi (2013), Mahajan and Yang (2017), Giulietti et al. (2018), Bertoli and Ruyssen (2018). Papers that highlight the potential deterrent effect of larger networks include Calvó-Armengol (2004), Calvó-Armengol and Jackson (2004), Wahba and Zenou (2005) and Beaman (2012).

[2]For instance, Munshi (2003) uses rainfall shocks at origin to instrument for network size at destination. Beaman (2012) exploits exogenous variation in the size of the migrant's social network induced by the quasi-random assignment of political refugees to new communities. Kinnan et al. (2018) take advantage of a resettlement program in China that sent 18 million urban youth to rural areas. Related approaches are used

reasonable approximation of the size of a migrant's social network, but they do not reveal if and how other aspects of social network structure influence the migration decision. Higher-order network structure – i.e., the connections of an individual's connections – plays a critical role in decisions about employment, education, health, finance, product adoption, and the formation of strategic alliances.[3] Yet, the role of such network structure in migration has not been systematically studied.

We leverage a rich new source of 'digital trace' data to provide a detailed empirical perspective on how social networks influence the decision to migrate. These data capture the entire universe of mobile phone activity in Rwanda over a five-year period. Each of roughly one million individuals is uniquely identified throughout the dataset, and every time they make or receive a phone call, we observe their approximate location, as well as the identity of the person they are talking to. From these data, we can reconstruct each subscriber's 5-year migration trajectory, as well as a detailed picture of their social network before and after migration.

The empirical analysis links each individual's migration decisions over time to the evolving structure of their social network. For instance, we use these data to confirm the longstanding hypothesis that people move to places where they know more people; conversely, individuals are less likely to leave places where they have larger networks. While these results may be intuitive, our data make it possible to disaggregate this relationship in a way that has not been done previously. In particular, we observe migration decisions for every possible network size and structure; we thus can estimate, for instance, that roughly 4% of individuals with 10 contacts in a potential destination d eventually migrate to that location. More broadly, we observe that the relationship between migration and network size is positive, monotonic, and approximately linear with slope of unity, such that the probability of migration roughly doubles as the number of contacts in the destination doubles. Superficially, this result diverges from a series of studies that predict eventual negative externalities from network size, as when members compete for information and opportunities (Calvó-Armengol, 2004, Calvó-Armengol and Jackson, 2004, Beaman, 2012, Dagnelie et al., 2019).

by Card (2001), Hanson and Woodruff (2003) and Dinkelman and Mariotti (2016).

[3] For example: Granovetter (1973), Burt (1992), and Karlan et al. (2009) provide examples of how higher-order network structure affects employment prospects. Banerjee et al. (2013), Beaman et al. (2015), and Ugander et al. (2012) illustrate the importance of higher-order structure in the adoption of microfinance, new plant seeds, and Facebook, respectively. Ambrus et al. (2015) and Chandrasekhar et al. (2018) relate network structure to contract enforcement and informal insurance. Keeling and Eames (2005) review how network structure influences the spread of infectious diseases. König et al. (2017) and Jackson and Nei (2015) link political network structure to strategic alliance formation. See Jackson (2010) and Easley and Kleinberg (2010) for an overview.

We then focus on developing a systematic understanding of how *higher-order* network structure — i.e., the connections of the migrant's connections (and their connections' connections, and so forth) — influences the decision to migrate. This is again something that would be very difficult to study with traditional survey data, but which we observe in rich detail in the mobile phone records. The purpose of this analysis is to understand whether, ceteris paribus, individuals are more likely to migrate to places where their social networks have particular network topologies. A stylized version of our approach is shown in Figure 1: we are interested in understanding whether, for instance, individual A is more likely to migrate than individual B, where both A and B know exactly two people in the destination and three people at home, and the only observable difference between A and B is that B's contacts are connected to each other whereas A's contacts are from two disjoint communities.

Our ability to identify the effect of social networks on migration is complicated by the fact that network structure is not exogenous. We address this concern in three principal ways. First, as noted, we focus on the relationship between the *higher-order* structure of a migrant's social network and subsequent migration decision. While a migrant may easily influence their direct connections, we assume they have less ability to influence the exact manner in which their connections are connected to one another. Second, we relate migration decisions in each month to the higher-order structure of the network several months prior. This is meant to minimize the likelihood that the decision to migrate shaped the social network, rather than vice versa.[4] Finally, we use an extremely restrictive set of fixed effects to eliminate many likely sources of omitted variable bias. Our preferred specification includes fixed effects for each individual migrant (to control for individual heterogeneity, for instance that certain people are both more likely to migrate and to have certain types of networks), fixed effects for each possible origin-destination-month combination (to control for factors that are shared by all people facing the same migration decision, such as wage and amenity differentials), and fixed effects for each possible destination network size (such that comparisons are always between places where the migrant has the exact same number of direct contacts, as in Figure 1). Thus, in our preferred specification, the identifying variation comes from within-individual differences in network structure between destinations and over different months in the 5-year window, net the population-average differences that vary by home-destination-month, and net any effects that are common to all people with exactly the

[4]Our main specifications relate migration decisions to network structure 2 months prior, but results are unchanged if we use lags between 2 and 6 months. We also find qualitatively similar results when we adopt a "shift-share" approach that relates migration decisions to prior *changes* in higher-order network structure, holding fixed the direct connections of the migrant.

same number of friends in the destination. We would observe such variation if, for instance, an individual had been considering a move to a particular destination for several months, but only decided to migrate after his friends in the destination became friends with each other — and if that tightening of his social network exceeded the average tightening of networks in that destination (as might occur around the holidays, for instance).

This analysis helps establish a new set of stylized facts about the relationship between migration and social networks. Most notably, we show that migrants are more likely to migrate to destinations where their social networks are *interconnected* (i.e., where the migrant's friends are friends with each other), but that they are no more likely to migrate to destinations where their networks are *extensive* (i.e., where their distance-2 and distance-3 neighborhoods are larger). In fact, conditional on network size migrants are *less* likely to go to places where their networks are extensive — a result that surprised us initially, given the emphasis prior work has placed on the value of connections to socially distant nodes in a network (e.g., Granovetter, 1973). In other words, of the three potential migrants in Figure 1, B is most likely to migrate and C is least likely, with A somewhere in between.

To better understand this 'surprising' result, we document considerable heterogeneity in the migration response to social network structure. In particular, we find that the negative effect of extensive networks is strongest when a migrant's direct contacts have a large number of "strong ties" in the destination (where we define a strong tie as one with 5 or more calls per month, which is equivalent to the 90^{th} percentile of call frequency); when a migrant's destination contacts have many weak ties (i.e., ties that are not strong), migration is not deterred. Such evidence suggests that there may be rivalry in information sharing in networks, which leads migrants to value connections to people for whom there is less competition for attention (as in Dunbar, 1998, Banerjee et al., 2012). We also find that while the *average* migrant is not drawn to locations where her friends have more friends (as in G_3), such structure does attract several less common types of migrants. In particular, repeat migrants (who have previously migrated from their home to the destination) and long-term migrants — both of whom are presumably better informed about the structure of the destination network — are more likely to migrate to locations where their networks are more extensive.

To summarize, this paper makes two main contributions. First, it provides a new empirical perspective on the determinants of migration in developing countries (cf. Lucas, 2015). In this literature, many scholars have noted the important role that social networks play in facilitating migration. Early examples in the economics literature include Rees (1966) and Greenwood (1969); a large number of subsequent studies document the empirical relationship between network size and migration rates.[5] More recently, Munshi and Rosenzweig (2016) document that the fear of losing social network ties may prevent profitable migration, while Morten (2019) shows that the act of migration can change social relationships and risk sharing. Kinnan (2019) theorizes about the two-way inter-connections: migration of one individual can make other network members better off if that individual has a new source of income, but others may be worse off if the act of migration improves the outside opportunity for that person or makes it easier to hide income. This paper builds on this line of work by exploiting a new source of data to establish a more nuanced set of stylized facts about networks and migration — highlighting, in particular, the value migrants place on interconnected networks, and substantial heterogeneity in how different types of migrants value networks.

Second, through the study of migration, we shed light on the more general question of how social networks provide social capital to individuals embedded in those networks (cf. Jackson, 2010, Banerjee et al., 2013, 2019).[6] We contrast interconnected and extensive networks, just as network theory distinguishes between networks that provide cooperation capital and networks that provide information capital (Jackson, 2020). In that literature, cooperation capital is usually motivated by repeated game models of network interaction, where interconnected networks (e.g., cliques) best support social reinforcement and sanctioning.[7] Information capital, which reflects the network's ability to efficiently transmit information, is associated with extensive subnetworks (e.g., stars and trees) where an individual is linked to many others via short network paths.[8] We show that, at least to migrants, topologies associated with cooperation capital matter most.

[5]Examples include Montgomery (1991), Borjas et al. (1992), Munshi (2003), McKenzie and Rapoport (2010), Dolfin and Genicot (2010), Beaman (2012), Patel and Vella (2012), Bertoli et al. (2013), Bertoli and Ruyssen (2018). Two recent papers use phone data to link spatial mobility and social networks. Büchel et al. (2020) use data from a Swiss cellphone operator to link migration decisions to phone calls, and document a similar relationship between network size and migration as the one we note in Section 4.1. Barwick et al. (2019) show that migrant flows in a Chinese city correlate with call volume between regions, and link this information flow to improved labor market outcomes. Both papers focus primarily on how on network size relates to migration, whereas our focus is on the role of higher-order network structure, conditional on network size.

[6]There is a large literature on social capital that studies how social structure fosters trust and cooperation in a society. In particular, the importance of social pressures on fostering cooperation has deep roots in the sociology literature (cf. seminal work by Simmel (1950) and Coleman (1988), among many others).

[7]Jackson et al. (2012) and Ali and Miller (2016) provide recent examples. See also Ligon and Schechter (2011), Jackson et al. (2012), Ambrus et al. (2015) and Chandrasekhar et al. (2018).

[8]Early models include Kermack and McKendrick (1927) and Jackson and Wolinsky (1996); more recent examples include Calvó-Armengol and Jackson (2004), Jackson and Yariv (2010), and Banerjee et al. (2013).

2 Data and Measurement

To study the empirical relationship between networks and migration, we exploit a novel source of data that contains detailed information on the migration histories and evolving social networks of roughly one million individuals in Rwanda. These data, obtained from Rwanda's near-monopoly telecommunications company, contain the Call Detail Records (CDR) for all mobile phone activity that occurred in Rwanda from January 2005 until June 2009. For each mobile phone call that occurs, the CDR contain a log of the (anonymized) phone numbers of the two parties involved in the call, a timestamp for when the call was placed, and the identifiers for the cell phone towers through which the call was routed, which in turn indicates the approximate geolocation of each party at the time of the call. In total, we observe roughly one billion mobile phone calls between roughly one million unique subscribers (Table 1).

By combining information on subscribers' locations (based on the cell towers they use) and social network structure (based on the people they speak to), we are able to study the relationship between migration and social networks. To provide intuition, the network of a single migrant, in the month before migration, is shown in Figure 2. This particular migrant (the green dot) had 20 unique contacts in the month prior to migration, 7 of whom were in his home district (blue dots), four of whom were in the destination district (red dots), and the remainder were in other districts (grey dots). Friends of friends are depicted as hollow grey circles.[9]

This section describes how we use these data to observe the structure of each individual's social network over time (Section 2.1) and to extract each individual's complete migration history (Section 2.2). Section 2.3 discusses limitations of these data. In our empirical analysis, we remove personally identifying information (including phone numbers) from the CDR we received from the mobile operator. In addition, to focus our analysis on individuals rather than businesses, and to remove the potential impact of spammers and call centers, we remove all transactions involving numbers with more than 200 unique contacts in a single month (this represents the 95th percentile). In later robustness checks we confirm that these thresholds for outlier removal do not affect our results.

[9] Throughout, we use the term 'friend' loosely, to refer to the contacts we observe in the mobile phone network. These contacts may be friends, family, business relations, or something else.

2.1 Modeling social network structure with mobile phone data

Our central goal is to understand how the structure of an individual's social network in different geographic locations affects their likelihood of migrating to and from those locations. The social network we observe is that of mobile phone communications. Specifically, we use the set of calls occurring within a specific time frame (typically a month) to define the (undirected) social network G at that time. Formally, let the network G be a matrix with $G_{ij} = G_{ji} = 1$ if i and j are observed to talk on the phone within a fixed time window and $G_{ij} = G_{ji} = 0$ otherwise (this includes $G_{ii} = 0$). A *path* between i and j is an ordered sequence of distinct agents $(ii_1i_2...i_hj)$ such that any two adjacent agents are connected in the network. The *distance* between i and j, denoted as $d(i,j)$, is the length of the shortest path between i and j.[10]

Since network structure can be quite complex (as in Figure 2), we focus on two archetypal ways that the *topology* of the social network influences the decision to migrate: through *information capital* and *cooperation capital*. Following Jackson (2020), we think of information capital as the potential for the social network to provide access to novel information — about jobs, new opportunities, and the like. By contrast, we consider the cooperation capital as the network's ability to facilitate interactions that benefit from cooperation and community enforcement, such as risk sharing and social insurance (similar to the notion of *favor capital* in Jackson (2020)).

Information capital. We construct a proxy measure of information capital for agent i in network G by measuring the size of the agent's second-degree neighborhood (or unique *friends of friends*, not counting direct contacts of the agent):

$$D_i^2(G) = \{j : d(i,j) = 2\}. \tag{1}$$

More generally, agent i's k^{th}-degree neighborhood is $D_i^k(G) = \{j : d(i,j) = k\}$.

In Appendix A1, we develop a model that provides microeconomic foundations for this particular measure of information capital. Broadly speaking, it is intended to capture the *extensiveness* of the agent's network, i.e., the extent to which one person is linked to many others via short network paths (cf. Granovetter, 1973). This measure relates closely to Jackson and Wolinsky's (1996) notion of decay centrality and Banerjee et al.'s (2013) measure of diffusion centrality. With both decay and diffusion centrality, information capital increases

[10] If there is no path between i and j, then $d(i,j) = \infty$.

with more friends, friends of friends, and so on.[11] As the number of friends is an agent's endogenous choice which we will examine separately, the number of unique friends of friends – the measure defined in equation (1) – is the key proxy to capture the individual's access to information in the outer network.

Cooperation capital. Our main proxy for an individual's cooperation capital is *network support*, i.e., the probability that an agent i's friend j has one or more friends in common with i. Formally, agent i's support in network G is 0 if i does not have any friend in G, otherwise

$$\text{Support}_i(G) \equiv \frac{\#\{j : G_{ij} = 1 \ \& \ (G^2)_{ij} \geq 1\}}{\#\{j : G_{ij} = 1\}}. \tag{2}$$

This proxy is designed to measure the *interconnectedness* of the agent's network, and relates closely to the notion of *favor capital* in Jackson (2020), defined as the network's "ability to exchange favors and transact with others through network position and repeated interaction and reciprocation" (p.315). Importantly, cooperation capital is facilitated by different network topologies than information capital: Appendix A1 shows that group enforcement is strong and cooperation is efficient when local subnetworks have high levels of support.[12] Network support is also correlated with *network clustering* (the probability that two friends are connected to each other), a metric we use in later tests of robustness.

Summary. Our empirical analysis distills the complex structure of social networks into two stylized network statistics: *Unique friends of friends*, a proxy for information capital, defined by equation (1); and *% Friends with common support*, a proxy for cooperation capital, defined by equation (2). We will also show results pertaining to i's *degree centrality*, $|D_i^1(G)|$, which simply counts the number of unique individuals with whom each person communicates. Ancillary results will separately account for the *strength* of a social tie, which we measure

[11] Jackson (2020) describes information capital as "ability to acquire valuable information and/or spread it to other people through social connections" (p.4). Decay centrality (Jackson and Wolinsky, 1996) assumes each agent receives a value $q < 1$ (the probability of information transmission) from each direct friend, a discounted value of q^2 from each friend of friend, and so on. Diffusion centrality (Banerjee et al., 2013) further accounts for the fact that multiple paths could increase the chance that information makes it from one individual to other.

[12] See also the references in footnote 7. In particular, Ali and Miller (2016) model a dynamic game of repeated cooperation and find that a clique network (a completely connected network) generates more cooperation and higher average cooperation capital than any other networks; Jackson et al. (2012) model a game of repeated favor exchanges and highlight the importance of *supported* relationships, where a link is supported if the two agents of the link share at least one common friend.

as the number of (undirected) calls between two individuals; when we compare strong and weak ties, we consider "strong" ties to be those ties in the 90^{th} percentile of the tie strength distribution (equivalent to 5 or more calls per month).

For most of the analysis that follows, we partition the full social network of Rwandan mobile subscribers (containing approximately 800,000 individuals) into 27 location-specific subnetworks, each of which is defined by the administrative districts of Rwanda.[13] Thus, we calculate (1) and (2) separately for each subnetwork G_d, which only has entries for individuals who reside in d. This simplifying assumption dramatically expedites our computational analysis, but assumes that agent i cannot derive social capital from a given district d via people residing outside d. We therefore include analysis that shows how our main results are affected by relaxing this assumption (see Section 4.3).

2.2 Modeling migration

Internal migration in Rwanda. Internal migration is a prominent feature of most developing countries. According to the United Nations Population Division (2013), there are an estimated 762 million internal migrants in the world. Yet, survey-based data on internal migration are notoriously unreliable, particularly in developing countries where many migrations are temporary (Deshingkar and Grimm, 2005, McKenzie and Sasin, 2007, Carletto et al., 2012, Lucas, 2015).

Our empirical analysis focuses on internal migration in Rwanda, a small agricultural economy in East Africa. Rwanda has high rates of poverty, estimated by the National Institute of Statistics of Rwanda (2012) to be 56.7% in 2005 (the beginning of the period we study). While fewer than 4% of Rwandan residents are born abroad, internal migration in Rwanda is common. According to the National Institute of Statistics of Rwanda (2014), roughly 20% of the resident population has experienced a lifetime migration, with similar migration rates for men and women. As with many predominantly agricultural societies, the most frequent type of internal migration in Rwanda is from one rural location to another (Lucas, 2015). For instance, the World Bank estimates that between 2005-2011, roughly two thirds of all migrants went to rural destinations; less than 20% of migrants were from rural to urban areas (World Bank Group, 2017).

The push and pull factors driving internal migration in Rwanda have varied over the last few decades. The 1994 genocide and surrounding conflict were major drivers of internal

[13] Our analysis groups the three smaller and contiguous districts that comprise the capital of Kigali into one "district".

migration in the 1990s, but conflict has been far less common since 2000. While the National Institute of Statistics of Rwanda (2014) did not collect data on migration motives, their analysis of patterns of urban and rural migration by gender "suggests that males mainly migrate toward urban areas for employment purposes while women tend to move shorter distances, either for marriage or agricultural purposes" (p. 7). Likewise, a series of reports from the Famine Early Warning System highlights the role that agriculture and construction play in driving labor migration, but also emphasizes the unpredictability of this demand (FEWS NET Rwanda, 2014). A more comprehensive study of internal migration in Rwanda, conducted by the World Bank Group (2017) and based on nationally representative household survey data from 2014 (EICV4), notes other factors common to many African countries: Rural-to-urban migrations are driven by urban employment opportunities and rural land shortages, and urban-to-rural migrations are frequently motivated by the high cost of living in the city and a desire for lower density areas where farmland may be available (especially in the Eastern Province).

Measuring migration with mobile phone data. We use mobile phone metadata to provide a detailed quantitative perspective on the migration patterns of mobile phone owners in Rwanda. This is possible because each time a mobile phone call is placed, the operator logs the cell phone towers through which the caller and receiver were connected; typically, these are the towers closest to the subscribers at the time of the call. As can be seen in Figure 3, this allows us to approximately and intermittently locate each subscriber, with a geographic precision of a few hundred meters in urban areas and several kilometers in rural areas.

We use the sequence of mobile phone towers to reconstruct each individual's history of migration. Our approach, described in more detail in Appendix A3, builds on prior methodological work to infer migration from mobile phone data (cf. Blumenstock, 2012, Lai et al., 2019, Chi et al., 2020). To summarize, we first calculate the district of residence d of every individual i in each calendar month t (districts are shown with black lines in Figure 3). We do this using Algorithm 1, which determines the district in which i spent the majority of evenings during t.[14] From this sequence of monthly residential locations, we then determine whether or not each individual migrates in each month. Following Blumenstock (2012), we say that a migration occurs in month $t+1$ if three conditions are met: (i) the individual's

[14]For each evening hour, we infer i's location as the district in which a majority of their calls occurred. We then infer i's location for the entire evening as the district in which the evening hours were spent. Ties are resolved with a coin toss.

home location is observed in district d for at least k months prior to (and including) t; (ii) the home location d' in $t+1$ is different from d; and (iii) the individual's new home location is observed to be 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 (such as individuals who do not use their phone for an entire month, or individuals who do not remain in one district for k consecutive months) are treated as "other" (and are excluded from later analysis).

Using this approach, we are able to provide granular detail on patterns of internal migration in Rwanda. Table 1 provides summary statistics on internal migration events observed in the mobile phone data, where a migration is defined as an instance where an individual stays in one district for at least 2 months, then moves to a new district and remains in that new district for at least 2 months (i.e., $k = 2$). The first column highlights data from a single month; the second column aggregates over a two-year period. Table A1 includes summary statistics for alternate definitions of migration – including different values of k and specific types of migration that are prominent in the literature on internal migration in developing countries (cf. Todaro, 1980, Lucas, 1997, 2015). Broadly, we observe a large number of repeat and circular migrants, with a majority of migrants traveling long distances. We also note that, comparing the rows of Table A1, the implied migration rate decreases as the minimum stay length k is increased. Such comparisons would be difficult with traditional survey data, which typically capture a single definition of migration. In later analysis, we use a minimum stay length of $k = 2$ as our base definition of migration, as the implied migration rate roughly matches official statistics on internal migration provided by the Rwandan government (National Institute of Statistics of Rwanda, 2014).

2.3 Data limitations and validation

While mobile phone data provide uniquely granular insight into the social networks and migration decisions of a large population, they also have several important limitations.

Non-representative population. During the period under study (early 2005 to early 2009), mobile phone penetration rose from roughly 5% to 22% (estimates based on the number of subscribers who appear in our dataset). During this time, mobile phone subscribers in Rwanda were not representative of the larger Rwandan population; survey evidence suggests they were significantly wealthier, older, better educated, and are more likely to be male (Blumenstock and Eagle, 2012). While this non-representativeness limits the external validity of our analysis, survey evidence suggests that the population of phone owners and the population of migrants have similar demographic characteristics.[15] More importantly, our empirical specifications are designed to limit the scope for patterns of phone ownership, including trends in mobile penetration over time, to bias our results — see the discussion of omitted variable bias and shift-share analysis in Section 3.3.

Phones are not people. The unique identifiers we observe are for mobile phone numbers, not individuals. As noted above, we attempt to limit the extent to which firms and organizations influence our analysis by removing phones with unusually high activity (as well as any traffic associated with those phones). Still, when multiple people share the same phone number, we may overestimate the size of an individual's network. It is also possible that a single individual might use multiple phone numbers, but we believe this was less common since there was only one dominant phone operator at this time. In principle, our data make it possible to uniquely identify devices and SIM cards, in addition to phone numbers. Of these, we believe that phone numbers (which is portable across devices and SIM cards) most closely correspond to unique individuals.

Construct validity. The social network we observe is the network of mobile phone relations, which is a subset of all true social relations in Rwanda. This subset is non-random: it is biased toward certain demographic groups; it systematically understates certain types of relationships (such as those that are primarily face-to-face); and may overstate other more transient or functional relationships (such as with a shopkeeper). We address some of these concerns through robustness tests that vary the definition of "social tie," for instance by only considering edges with several observed communication events (see Section 4.4). Other concerns are ameliorated by the fact that much of our analysis focuses on long-distance relationships, and during this period in Rwanda the mobile phone was the primary means of communicating over distance.

Related, we measure migration based on the movement of phones, rather than with traditional survey-based instruments. Prior work suggests that patterns of migration inferred from mobile phone data broadly match inferences drawn from other sources — this includes work in Rwanda using the same dataset as in this paper (Blumenstock, 2012, Williams et

[15] For instance, the age distribution of migrants estimated from 2012 government census data (National Institute of Statistics of Rwanda, 2014, Figure 11) is similar to the age distribution of a representative survey of mobile phone owners in 2009 (Blumenstock and Eagle, 2010, Figure 2).

al., 2013), neighboring countries in East Africa (Wesolowski et al., 2013a,b, Pindolia et al., 2014), as well as other low-income (Bengtsson et al., 2011, Lu et al., 2016, Lai et al., 2019) and wealthy nations (Lenormand et al., 2014). In our context, the aggregate patterns of population flows that we calculate from the mobile phone data between 2005 and 2009 are broadly similar to those reported in the 2012 Rwandan census, but there are discrepancies between the two measurements. For instance, Figure A1 compares estimates of internal migration from the phone data (red bars) to those from the census (blue bars), as reported by National Institute of Statistics of Rwanda (2014, p.29). These inconsistencies could be due to the non-representativeness of phone owners, to differences in how the two instruments define migration,[16] or to the fact that mobile phone data is a more sensitive instrument for detecting human mobility than the typical census questionnaire.

3 Identification and estimation

The focus of this paper is on understanding how social networks influence the decision to migrate. While a host of other factors also influence that decision — from wage and amenity differentials to physical distance and associated migration costs — we study how, holding all such factors fixed, variation in social network structure systematically correlates with migration decisions. In the stylized example of Figure 1, we ask whether a person with network G_1 is more likely to migrate than someone with network G_2, whose network is marginally more interconnected and would be expected to provide marginally more cooperation capital. We similarly compare the migration decisions of such individuals to individuals with network G_3, which is slightly more extensive and would be expected to provide slightly more information capital. In practice, of course, the actual network structures are much more complex (as in Figure 2). We therefore use statistical models to estimate the effect of marginal changes in complex network structure on subsequent migration decisions.

The central difficulty in identifying the causal effect of social networks on migration is that the social networks we observe are not exogenous: people migrate to places where their networks have certain characteristics, but this does not imply that the network caused them to go there. Here, we describe our estimation strategy, and the identifying assumptions required to interpret our estimates as causal.

[16] Our algorithm defines a migrant as someone who remains in one district for 2 or more months and then moves to another district for 2 or more months. The Rwandan census does not capture this type of short-term migration; we instead show the census estimates of internal recent migrants, which are defined as "a person who moved to his/her current district of residence five years or less prior to the census."

3.1 Simultaneity

An obstacle to understanding the causal effect of networks on migration is that migration decisions may also shape networks. This would be expected if, for instance, migrants strategically formed links to destination communities in anticipation of migration, or simply made a large number of phone calls to their destination before migrating.

We address this concern in three principal ways. First, we study the effect of lagged network characteristics on the current decisions of migrants. Specifically, we relate the migration decision made by individual i in month t to the structure of i's social network s months prior. As a concrete example, when t = August 2008 and $s = 2$, we relate the August 2008 migration decision to the structure of the individual's social network in June 2008. Our main specifications use $s = 2$, but we later show that our results are unchanged when with longer lags. Second, rather than focus on the *number* of direct contacts a migrant has at home and in the destination, we focus on the *connections* of those contacts, holding the number of contacts fixed (as in Figure 1). This is because it seems easier for a migrant to directly control the number of contacts they have in the destination and at home than it is for them to alter the higher-order structure of their social network. Third, in tests described in Section 4.4, we adopt a "shift-share" specification that relates migration decisions to *changes* in an individual's higher-order network (for instance, between $t - 12$ and $t - 2$), holding lower-order network structure fixed, in order to further limit the extent to which the individual could endogenously shape their network.

These techniques reduce, but do not eliminate, the potential for simultaneity. In particular, a migrant might plan her migration many months in advance of migration, and in that process could change her higher-order network structure — for instance by asking a friend to make new friends on her behalf, or by encouraging two friends to talk to each other. To gauge the extent to which this might bias our results, we run several empirical tests, and find little evidence of such anticipatory behavior. For instance, Figure 4a shows, for a random sample of migrants, how the geographic distribution of migrants' social networks changes over time. Prior to migration, roughly 40% of the average migrant's contacts are in the origin and 25% are in the destination; three months after migration, these proportions have switched, reflecting how the migrant has adapted to her new community. Notably, however, migrants do not appear to strategically form contacts in the destination immediately prior to migrating; if anything, migrants shift their focus to the people in the community they are leaving – and any deviations from the long-term trend don't appear until the month of migration. These compositional changes do not mask a systematic increase in the *number* of contacts in the destination, or the number of total calls to the destination: Figure A2 indicates that the total number of contacts increases over time, but there is no sudden spike in the destination district in the months before migration; Figure A3 shows analogous results for total call volume. As a sort of 'placebo' test, Figure 4b shows how networks evolve over time for non-migrants, where we draw a sample non-migrants that matches the temporal distribution of migrants from Figure 4a.[17] While non-migrants have different network structure than migrants (in particular, non-migrants have a higher fraction of contacts in their home district than migrants do), there are no sudden changes in the composition of network structure of non-migrants. With non-migrants, as was the case with migrants, we observe a slow long-term trend toward a larger share of communication being within the home district.

What matters most to our identification strategy is that we similarly find no evidence that migrants are systematically altering the *higher-order* structure of their social networks in the months prior to migration. In particular, Figure A4 indicates that while the number of migrants' friends of friends slightly increases over time, there is no noticeable shift in the months prior to migration. Figure A5 shows similar results for the level of common support in the network.[18]

3.2 Omitted Variables

The second main threat to identification is the fact that network structure may be a proxy for other characteristics of the individual (e.g., wealth, ethnicity) and location (e.g., population density, wages) that also influence migration. Our main strategy for dealing with such omitted variables is to include an extremely restrictive set of fixed effects that control for many of the most concerning sources of endogeneity. This strategy is possible because of the sheer volume of data at our disposal, which allow us to condition on factors that would be impossible in regressions using traditional survey-based migration data.

Our preferred specification includes fixed effects for each individual (roughly 800,000 fixed effects), for each origin-destination-month tuple (roughly 18,000 fixed effects), and for the number of direct contacts in the destination (roughly 100 fixed effects). The individual fixed effects absorb all time-invariant individual heterogeneity (such as wealth, gender, ethnicity, personality type, family structure, and so forth), and addresses the fact that some people are inherently more likely to migrate than others (and have inherently different social

[17] Specifically, for each migrant who appears in Figure 4a and is observed to migrate in month t, we randomly sample a non-migrant from the same month t to include in Figure 4b.

[18] In Section 4.4, we further show that there are no sudden changes in higher-order structure even after 'freezing' the migrant's contacts at $t-12$.

networks). The origin-destination-month fixed effects control for any factor that similarly affects all individuals considering the same origin-destination migration in the same month. This includes factors such as physical distance, the cost of a bus ticket, location-specific amenities that all migrants value equally, average wage differentials, and many of the other key determinants of migration documented in the literature (including the usual "gravity" effects in a standard trade or migration model).[19] Finally, we include fixed effects for the number of first-degree contacts in the destination in order to isolate the effect of differences in higher-order network structure on migration.

3.3 Identification

To summarize, the identifying variation in our main specification is (i) within-individual over time and (ii) within-individual over potential destinations — in both instances, after controlling for any factors that are shared by all people considering the same destination in the same month, and for any effects that are common to all people with the same number of direct contacts in the destination. An example of (i) could occur if, for instance, an individual had been considering a move to a specific destination for several months, but only decided to migrate after his friends in the destination became friends with each other (the G_2 vs. G_1 comparison of Figure 1) — and if the increased interconnectedness exceeded the average increase of networks in that destination (as might occur around the holidays, for instance). An example of (ii) could occur if, in a given month, a single migrant were choosing between two destination districts, had the same number of contacts in each district, and then decided to migrate to the district where his contacts were more interconnected. Prima facie, it may seem unlikely that such small differences would shape the decision to migrate, but our data allow us to ascertain whether, across millions of individual migration decisions, such a general tendency exists.

The fixed effects we include significantly reduce the scope for omitted variables to bias our estimates of the effect of network structure on migration, but they do not eliminate such bias entirely. If, for instance, origin-destination wage differentials are individual-specific, our fixed effects will not help. This might occur if carpenters' networks in a particular district grew more interconnected over time (relative to carpenter network growth in other districts) than

[19]For instance, we know that rates of migration are higher to urban centers, and that social networks in urban centers look different from rural networks. Including a destination fixed effect removes all such variation from the identifying variation used to estimate the effect of networks on migration. The origin-destination-month fixed effects remove destination-specific variation, as well as more complex confounding factors that vary by destination and origin and time, such as the possibility that the seasonal wage differential between two districts correlates with (lagged) fluctuations in social network structure.

farmers' networks in that district (again relative to farmers' networks in other locations), and if migration rates of carpenters to that district are higher for reasons unrelated to network structure. Likewise, our use of lagged network structure reduces, but does not eliminate, the likelihood that a migrant would first decide to migrate and then modify his network accordingly.

We revisit these concerns, and other possible threats to identification, in Section 4.4, after introducing the estimation strategy and presenting the main results. In Section 4.4, we precisely state the identifying assumption, discuss the most likely threats to identification, and perform a number of tests to assess the plausibility of this identification strategy.

3.4 Estimation

Formally, for an individual i considering a move from home district h to destination district d in month t, we wish to estimate the effect of a vector of (s-lagged) network characteristics $Z_{id(t-s)}$ on the migration decision. This is a discrete choice setting in which i faces 27 mutually exclusive choices in month t, one for each district d in Rwanda (including the home district h). We assume the indirect social capital i would receive from being in d is a function of individual characteristics (μ_i), fixed characteristics of d in the month the choice is being made (π_{dt} and ν_{dt} for destination and home districts, respectively), and a vector of choice-specific attributes that may also differ across individuals ($Z_{id(t-s)}$):

$$U_{idt} = \mathbb{1}(d \neq h)\left[\beta_d' Z_{id(t-s)} + \pi_{dt}\right] + \mathbb{1}(d = h)\left[\gamma + \beta_h' Z_{id(t-s)} + \nu_{dt}\right] + \mu_i + \epsilon_{idt} \tag{3}$$

Our focus is on the influence of i's network $Z_{id(t-s)}$ (measured with s lags), which the above formulation allows to differ for home networks ($d = h$) and destination networks ($d \neq h$). The vectors β_d and β_h contain the coefficients of interest, which indicate the average effect of destination and home network properties, respectively, on the probability of migration. The parameter γ captures the average tendency for individuals to not migrate.

Assuming that ϵ_{idt} is drawn from an extreme value distribution, i will choose d at time t with probability:

$$P(M_{idt} = 1) = \frac{\exp(\tilde{U}_{idt})}{\Sigma_{d'} \exp(\tilde{U}_{id't})}$$

which can be estimated with a conditional logit model (using $\tilde{U}$ to denote U without the disturbance term ϵ).[20] We omit μ_i from $\tilde{U}$ because it does not vary across the set of choices

[20] We use the approach described by Eaton et al. (2012) and Sotelo (2019) to estimate the multinomial

faced by i in month t.

Since much of our focus will be on understanding how the shape of an individual's *higher order* network structure relates to their decision to migrate, many specifications will additionally control flexibly for the size of i's first-degree network:

$$U_{idt} = \mathbb{1}(d \neq h)\Big[\beta_d' Z_{id(t-s)} + \pi_{dt} + \sum_k \eta_k \mathbb{1}(D_{id(t-s)} = k)\Big] + \mathbb{1}(d = h)\Big[\gamma + \beta_h' Z_{id(t-s)} + \nu_{dt} + \sum_k \zeta_k \mathbb{1}(D_{id(t-s)} = k)\Big] + \mu_i + \epsilon_{idt} \quad (4)$$

In the above specification, the vectors of fixed effects η_k and ζ_k allow for migration probabilities to differ for people with different numbers of unique contacts k both at home and in the destination.

When estimating (3) and (4), individuals are only considered in months where they can be classified as a migrant or a non-migrant in that month. When an individual is classified as "other" (see Section 2.2), those observations are excluded from the regression. Except as noted, specifications use cluster robust standard errors, clustered by individual. Alternative treatments of the standard errors are discussed in Section 4.4.

4 Results

4.1 The effect of network size, in the destination and at home

Table 2 summarizes the main results from estimating the multinomial logit model (3) of the migration decision on lagged network structure. We find that on average, each additional contact in the destination is associated with a 0.316% increase in the likelihood of migration to that destination (column 2), and each contact at home is associated with a 0.081% decrease in the likelihood of migration. As discussed in Section 3.3, these coefficients are identified by changes within the individual's network over time, and across destinations in a single period. Comparing the coefficients on Destination Degree and Home Degree in the first two columns, we can compare the "push" and "pull" forces of networks on migration (cf. Hare, 1999): the effect of adding one additional contact in the destination is roughly 4 to 6 times as important as an additional contact at home.

The coefficients in the first row of Table 2 validate a central thesis of prior research on networks and migration, which is that individuals are more likely to migrate to places

model by Poisson pseudo maximum likelihood – see also Correia et al. (2019).

where they have more connections. But the richness of our data allow us to do much more than simply look at these average effects. For instance, Figure 5a shows how the average migration rate varies by degree centrality at destination (i.e., the number of unique contacts of the individual). We observe that, for instance, roughly 4% of individuals with 10 contacts in a potential destination d in month $t-2$ migrate to that location at t. More broadly, we observe that the relationship between migration and network size is positive, monotonic, and approximately linear with slope of unity.

Just as migrants appear drawn to destinations where they have a large number of contacts, migrants are less likely to leave origins where they have a large number of contacts. Figure 5b shows the monotonically decreasing relationship between migration rates and the individual's degree centrality at home, where the probability of leaving home decreases in proportion to the size of the home network.

4.2 Higher-order network structure

We next examine the role of *higher-order* network structure — i.e., the connections of the individual's contacts — in migration decisions. This analysis uses specification (4), which includes fixed effects for each possible network size, so that identification now comes from changes (over time and across destinations) in the *inter*-connections of the migrant's network (i.e., the connections of i's connections), holding the number of direct contacts fixed. Results in the second and third rows of Table 2 highlight the two main results that we explore in greater detail below: migrants are more likely to go to places where their destination networks are more interconnected (row 2); but they are in fact *less* likely to migrate to destinations where their contacts have a large number of contacts (row 3).[21]

Network 'interconnectedness'

The results in Table 2 indicate that, on average, migrants are more likely to go do destinations that are more interconnected. In other words, networks like G_2 in Figure 1 are more attractive than networks like G_1.[22] Our data allow us to disaggregate this effect, and unpack how migration rates vary at different levels of network interconnectedness. In particular, the

[21] Similar results are obtained when focusing on the choice of destination among individuals who have already decided to migrate: Table A2 presents multinomial logit results estimated just on the subset of individual-months where the individual is observed to migrate. The results across columns are qualitatively unchanged from the main results in Table 2.

[22] It is worth noting that in other settings, more network interconnections are not necessarily attractive. For instance, Ugander et al. (2012) show that people are less likely to sign up for Facebook when their pre-existing Facebook friend network is more interconnected.

left panels (a and c) of Figure 6 show how the average migration rates varies with network support, the measure of interconnectedness defined by equation (2). The lack of a clear trend in the left panels is difficult to interpret in part because network support can vary systematically with network size. For this reason, the right panels (b and d) of Figure 6 show how migration varies with network support, *holding network size fixed.*

Specifically, the right panels of Figure 6 are produced by plotting the β_{kd} and β_{kh} coefficients from:

$$U_{idt} = \mathbb{1}(d \neq h)\Big[\sum_k \mathbb{1}(D_{id(t-s)} = k)[\eta_k + \beta'_{kd} Z_{id(t-s)}] + \pi_{dt}\Big] + \\ \mathbb{1}(d = h)\Big[\gamma + \sum_k \mathbb{1}(D_{id(t-s)} = k)[\zeta_k + \beta'_{kh} Z_{id(t-s)}] + \nu_{dt}\Big] + \mu_i + \epsilon_{idt} \quad (5)$$

This specification is directly analogous to specification (4), which was used to estimate column 3 of Table 2. The key difference is that where specification (4) provided a single estimate of the average effect of network support on migration, specification (5) estimates the effect of network support $Z_{id(t-s)}$ separately for each unique value of network size k.

Panels b and d of Figure 6 reinforce the prior finding that people are systematically drawn to places where their networks are more interconnected: most coefficients in Figure 6b and Figure 6d are (weakly) positive, indicating that migrants with a wide variety of network sizes are drawn to places where those networks are more interconnected. The figure also adds a level of nuance that would not be possible with traditional survey-based data. For instance, the fact that the β_{kd} coefficients in Figure 6b are generally increasing indicates that as the potential migrant has more direct contacts in the destination, the value of interconnections between those contacts increases. Appendix Figure A6 finds qualitatively similar results when using network clustering, instead of network support, as a measure of interconnectedness.[23]

To provide further intuition for these results, and the variation that identifies them, we conduct the following test: We pull a random sample of 20,000 individuals who have exactly two contacts in a specific district for 4 consecutive months. We then calculate, for each person, whether those two contacts are became more connected or disconnected at the end of the 4-month period (by regressing a dummy for triadic closure on a linear time trend); we then compare the migration rate in month 5 among the population whose two contacts

[23]The distinction between support and clustering is that the former counts the proportion of i's friends with one or more friends in common, the latter counts the proportion of all possible common friendships that exist – see Jackson (2010).

became more connected relative to the migration rate in month 5 of the population whose two contacts became less connected. The migration rate is 2.2% in the former group, and 1.3% in the latter. In other words, when focusing on a sample who consistently have exactly two contacts in the destination, rates of migration are higher when a given individual's two contacts become more connected (over the 4-month period) than when they become more disconnected (over the 4-month period).

Network 'extensiveness'

The relationship between migration and network extensiveness is more surprising and subtle. Here, we are interested in the generalized comparison between G_1 and G_3 in Figure 1, and use the size of an individual's second-degree network $|D_i^2(G)|$ (i.e., their unique friends of friends) as a measure of extensiveness. Without controlling for the size of an individual's network, there is a strong positive relationship between migration and extensiveness in the destination (Figure 7a), and a strong negative relationship with extensiveness in the origin (Figure 7c). The shape of these curves resemble the relationship between migration rate and degree centrality shown earlier in Figure 5: average migration rates increase roughly linearly with the number of friends of friends in the destination, and decrease monotonically with the number of friends of friends at home.

Of course, the number of friends of friends a person has is heavily influenced by the number of friends that person has. Thus, Figures 7b and 7d show how the number of friends of friends relates to migration, using specification (5) to hold fixed the number of friends. For the home network, Figure 7d indicates the expected pattern: the fact that all of the coefficients are positive suggests that given a fixed number of friends at home, people are less likely to leave when those friends have more friends. We also observe that the number of friends of friends at home matters more for people with fewer direct contacts — by the time an individual has a very large number of direct home contacts, their contacts' contacts matter less.

The surprising result is Figure 7b, which indicates that the likelihood of migrating does not generally increase with the number of friends of friends in the destination, after conditioning on the number of friends. The friend of friend effect is positive for people with just one contact, but negative for people with three or more destination contacts. Averaged over all migrants, this effect is small but negative and statistically significant (row 3 of Table 2). This result is difficult to reconcile with standard models of information diffusion (e.g., Kempe et al., 2003, Banerjee et al., 2013). Indeed, much of the literature on migration and social networks suggests 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.

We run a large number of empirical tests to convince ourselves that this pattern is not an artifact of our estimation or measurement strategy — several of these are described in Section 4.4. However, the data consistently indicate that the average migrant is no more likely to go to places where she has a large number of friends of friends. This is perhaps most transparent in Figure A7, which shows the distribution of the count of friends of friends for all migrants and non-migrants with exactly 10 friends in the potential destination. Among this sample, it is apparent that, on average, non-migrants have more friends of friends in the destination networks than migrants.

4.3 Heterogeneity and the 'friend of friend' effect

The effect that networks have on the "average migrant" masks considerable heterogeneity in how different types of migrants are influenced by their social networks. Tables A3-A5 disaggregate the results from Table 2 along several dimensions that are salient in the migration literature: whether the migrant has previously migrated to the destination (Table A3); whether the migrant stays in the destination for a long period of time (Table A4); and whether the migration is between adjacent districts or over longer distances (Table A5).

Heterogeneity and unawareness of the broader network

Several patterns can be discerned from these tables, but we focus our attention on how the network "extensiveness" effect changes with these different subgroups, as that was the most unexpected of the above results. Here, we find that for certain types of migration — repeat migrations, long-term migrations, and short-distance migrations — the number of friends of friends is positively or insignificantly correlated with migration rates. Each of these types of migration are significantly less common than the typical migration event (a first-time, short-term, and long-distance migration) — hence the negative average effect observed in Table 2.

This heterogeneity suggests one possible explanation for the unexpected 'friend of friend' result: the average migrant may simply be unaware of the higher-order structure of their destination network. Such an explanation is supported by several other studies that find that people have incomplete information about the friends of their friends (Friedkin, 1983, Casciaro, 1998, Chandrasekhar et al., 2016). This lack of information is likely to be most pronounced when the would-be migrant lives far from, or has less experience with, the destination friend's community. And indeed, this is what the heterogeneity suggests: the migrants who are positively influenced by extensive destination networks are the migrants who seem likely to be more familiar with the structure of those networks. When the destination is more familiar, it begins to resemble the home network, where people have good information on (proxies for) their friends' centrality (Banerjee et al., 2019).

Recent migrants, recent visits, and strong and weak ties

Appendix Tables A6-A8 indicate that the "extensiveness" of a migrant's destination network does not increase their probability of migration, even after accounting for several other factors that likely matter in the decision to migrate. For instance, we observe that people are more likely to go to places where their contacts have recently migrated. Coefficient estimates in column 3 of Table A6 indicate that knowing a contact who previously made a specific origin-destination migration increases the likelihood of the migrant choosing that destination by 2-2.5 times the amount as knowing anyone else in the destination. The effect is similar for recent migrants who arrived in the destination very recently (last month) as for recent migrants who arrived at any point prior.

Likewise, Table A7 controls for a binary variable indicating whether i ever appeared in district d in the month prior to t. In column 2, a "prior visit" is defined as making or receiving a call or text message from a tower in d; in column 3, we only consider activity that occurs between 6pm and 7am, in an effort to capture overnight visits. There is a strong correlation between such visits and migration (the effect is roughly 12 times as large as the effect of an additional direct contact in the destination). Controlling for these in-person visits does not change the qualitative role that networks play in shaping migration, but it does noticeably attenuate the effect of destination network structure (i.e., the effect of direct contacts decreases by roughly 30% and the effect of support decreases by roughly 40%), suggesting the in-person experience might substitute for network connections. Controlling for in-person visits has little effect on the influence of home network structure.

Table A8 disaggregates the effect of social network connections by the 'strength' of the social tie, where we define a 'strong' tie as a contact with whom the individual communicates five or more times in the reference month (this is equivalent to the 90^{th} percentile of communication frequency). Here, and consistent with recent work by Giulietti et al. (2018), we find that both strong and weak ties matter in migration: the effect of a strong destination tie is 0-34% larger than that of a weak destination tie; at home, the effect of a strong tie is

150 - 200% larger than a weak tie.

Also interesting is the effect of *higher order* tie strength on migration decisions. In particular, our main results suggest that a migrant i is drawn to locations where i's contact j has a friend in common k, but that i is indifferent or repelled if k is not a common friend of i. However, this average effect hides a more nuanced pattern: when disaggregating by tie strength, we observe that the negative effect is driven by situations where the i-j tie is weak but the j-k tie is strong — in other words, when the migrant has a tenuous connection to the destination and that tenuous connection has strong connections to other people in the destination.

These results are presented in Figure 8, which summarizes the regression coefficients from Tables A9 and A10. The figure indicates the sign of the regression coefficient (using $+/-$ labels) from a regression of i's migration decision on the number of different types of i-j links, where type is determined by the strength of the i-j link (strong ties shown with thick lines, weak ties shown with thin lines) and the existence and strength of the j-k link. The four figures on the left, based on Tabls A9, indicate that migrants are generally drawn to places where their contacts have many ties, but that they are deterred when their weak ties have a large number of strong ties. Similarly, the set of triangles on the right show all possible configurations of a supported i-j tie (based on Table A10), and indicate that supported links are positively correlated with migration in all cases except when the i-j tie is weak and the j-k tie is strong.

This heterogeneity is consistent with the notion, proposed by Dunbar (1998) and others, that people might have a capacity constraint in the number of friendships they can effectively support, which in turn might induce a degree of rivalry for the attention of a friend. In our context, migrants may be drawn to places where they receive their friends' undivided attention.[24]

[24]Dunbar originally proposed that humans could maintain roughly 150 stable relationships, since "the limit imposed by neocortical processing capacity is simply on the number of individuals with whom a stable inter-personal relationship can be maintained." In the migration context, Beaman (2012) and Dagnelie et al. (2019) find evidence that migrants may compete with each other for economic opportunities. See also Wahba and Zenou (2005), who empirically test the tradeoff between information and rivalry in an Egyptian labor market survey. They show that up to a certain (network) size, the network information effect dominates the competition (rivalry) effect so that network is always beneficial for finding a job. However, above a certain size, the second effect dominates the first one so that agents have less chance of finding job when network size increases.

Beyond location-specific subnetworks

The regression results presented above calculate network extensiveness and interconnectedness based on location-specific subnetworks at home and in the destination. It is possible, however, that the social capital from network connections may cross geographic boundaries. For instance, a potential migrant i in home district h might receive information about a destination district d from a person k (who lives in d) via a common friend j that lives at home h or in a district other than d. We therefore show how results change when we relax restrictions on the location of the intermediate contact j.

Results in Table A11 suggest that the main results — and in particular the negative role of extensiveness — do not depend on restrictions on the location of intermediate connections. For reference, Column 1 of Table A11 replicates the prior result from Column 3 of Table 2. Column 2 of Table A11 then shows results when i's direct contact j lives in the home district d; we observe that the coefficient associated with network extensiveness (i.e., unique friends of friends) remains negative, and is in fact much larger in magnitude than in column 1 – the increase is likely due to the fact that when i's home network includes an additional contact j to intermediate the connection to k, this also directly increases i's propensity to remain at home.[25] Column 3 allows for both types of network support (intermediated by destination friends and intermediated by home friends) to jointly influence the migration decision; both coefficients remain negative.

4.4 Robustness and identification (revisited)

Section 3.3 introduced our identification and estimation strategy. For our estimates to be causal requires the identifying assumption that $E[\epsilon_{idt}|\pi_{dt}, \mu_i, \eta_k] = 0$. In other words, we assume that the variation in higher-order network structure we observe is exogenous, conditional on the identity of the individual making the migration decision, the origin-destination-month choice being made, and the number of direct contacts the individual has in that destination in that month. Here, we discuss and test the limitations of that assumption, focusing on the two main threats to identification highlighted in Section 3.3: simultaneity and omitted variable bias.

[25] Likewise, the negative coefficient on destination support (intermediated by common friends at home) in Column 2 is likely do to the fact that the additional contact at home required to increase destination support has a dampening effect on the migrant's propensity to leave home.

Evidence against simultaneity: Temporal lags and 'shift-share' analysis

Our identification relies, in part, on the assumption that migrants do not strategically shape the higher-order structure of their social networks after making the decision to migrate. To support this assumption, Figures A4 and A5 indicate that, even among eventual migrants, there are no sharp changes in average higher order network structure in the months leading up to migration. Here, we provide additional analysis related to this identifying assumption.

First, we increase the lag between measurement of network structure and migration. Our main specifications (e.g., Table 2) test how migration decisions in month t (e.g., August 2008) relate to social network structure in month $t-2$ (e.g., June 2008). Table A12 shows that results are qualitatively unchanged when migration in t is regressed on network structure in $t-6$ instead. Migrants may plan migrations more than 6 months in advance, but the similarity of the results using 2-month vs. 6-month lagged networks suggests that strategic network formation is not driving our results.

Second, we test a "shift-share" specification that relates migration decisions to changes in an individual's higher-order destination network structure, holding lower-order network structure fixed. Specifically, we define an early period t_0 (e.g., 12 months prior to migration) and a late period t_1 (e.g., 2 months prior to migration), and measure the change in the higher-order network structure of each individual i between t_0 and t_1. In these specifications, we "freeze" the set of i's direct contacts at t_0, in an effort to reduce the endogenous decisions that i makes about their direct contacts (over whom they presumably have more direct control). The identifying variation in the shift-share specification comes from changes in the contacts of i's "frozen" contacts.

As intuition, Figures A8 and A9 reconstruct Figures A4 and A5, but hold fixed each migrant's contacts at $t-12$ to show how higher-order structure of that fixed network changes over time in the months leading up to, and immediately following, the eventual migration. In Figure A8, we observe no sudden or gradual changes in the average number of friends of i's friends from $t-12$. In Figure A9, we observe a gradual decrease in network support, but no sudden changes in the months immediately prior to migration.[26] The lack of sudden changes suggests that migrants are not systematically altering the high-order structure of their social networks in anticipation of migration.

[26] The gradual decrease in support is likely due to the fact that any specific edge in the network (including the j-k edge that provides support to the migrant i's edge with k) has a positive probability of disappearing over time. For each j-k edge that disappears, support will decrease unless it is replaced by a different j-m edge involving an m who is also connected to i. By contrast, we do not observe a gradual decrease in i's friends of friends. This is likely because j can replace a lost contact with any m — including those not connected to i – and maintain the same number of friends of friends.

Table A13 presents regression results of the shift-share approach, in a format similar to Table 2, but now holding fixed i's contacts from t_0={6,12} months prior to migration. In both cases, we measure changes in higher-order network structure based on the connections of the intersection of i's contacts at t_0 and t_1, i.e., the set of contacts who were connected to i in both t_0 and t_1. All specifications set the late period t_1 at 2 months prior to migration. The results of the shift-share analysis are broadly consistent with the main results in Table 2. Increases in friends of friends that occur within i's frozen-in-time contacts in the destination are insignificantly or negatively correlated with migration. Increases in support in the destination are positively associated with migration. We also note that the total predictive power of changes in network structure is limited (i.e., the partial R^2 values in Table A13 are all less than 0.02). If migrants were systematically shaping their higher-order networks in anticipation of migration (and in advance of the $t-2$ lag used in our main specification), it is likely that such behavior would better predict migration.

Evidence against omitted variable bias: Increasingly restrictive fixed effects

Our preferred conditional logit specification Specification (3) includes fixed effects for each individual (μ_i), each destination-month combination (π_{hdt}), and each destination degree centrality (η_k). While these account for the most likely sources of omitted variable bias, there are scenarios in which this assumption could be violated (as in the carpenter/farmer example in Section 3.3). We therefore run a series of robustness checks that further isolate the identifying variation behind the regression results presented above.

In particular, we temporarily switch to a linear regression specification, which makes it possible to include a very restrictive set of fixed effect that we could not estimate with a conditional logit specification. In the linear specification, we define migration M_{ihdt} as a binary variable equal to 1 if the individual chooses to move from h to d at t and 0 otherwise, and estimate:

$$M_{ihdt} = \beta' Z_{ihd(t-s)} + \pi_{hdt} + \mu_i + \sum_k \eta_k \mathbb{1}(D_{id(t-s)} = k) + \epsilon_{ihdt} \tag{6}$$

where π_{hdt} are (home district * destination district * month) fixed effects, μ_i are individual fixed effects, and $D_{id(t-s)}$ is i's degree centrality.[27]

Column 1 of Table A14 presents the results of estimating (6), using fixed effects similar

[27] When estimating (6), we include one observation for each potential destination d of each individual i in each month t. We define a 'potential destination' as any non-home district $d \neq h$. Thus, the regression includes at most 26 observations for each individual i in each month t.

to those in our preferred conditional logit specification (Table 2): π_{hdt}, μ_i, and η_k. The linear model results are qualitatively similar: while additional network support increases the likelihood of migration to the destination, additional friends of friends do not increase migration likelihood. Column 2 in Table A14 then includes fixed effects for each *individual-month* pair, so that the identifying variation comes *within individuals in a given month* but across potential destination districts.[28] Column 3 instead includes separate fixed effects for each *individual-destination* pair, so that the β coefficients are identified solely by variation within individual-destination over time.[29] Column 4 includes fixed effects for each *individual-Degree*, exploiting variation between all destinations where a single individual has the exact same number of contacts. Column 5, which includes over 600 million fixed effects, isolates variation within individual-home-destination observations over time. In all instances, the coefficients of interest are quite stable, and in particular, the average effect of additional friends of friends is either negative or insignificant (or both).

Additional tests of robustness

We perform several additional tests to check whether the main results are sensitive to different measurement strategies used to process the mobile phone data. Since these results show a very similar picture and are highly repetitive, we omit them for brevity:

- **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 the '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 between these individuals. As a robustness check, we take a more restrictive definition of social network and only include reciprocated edges, i.e., cases where i calls j at least once and j calls i at least once.

[28]Such variation would occur if, for example, in a given month, a single migrant were choosing between two destination districts, had the same number of contacts in each district, and then decided to migrate to the district where his contacts were more interconnected — and if that additional interconnectedness exceeded the extent to which all networks in that destination were more interconnected.

[29]This could reflect a scenario where an individual had been considering a move to a particular destination for several months, but only decided to migrate after his friends in the destination became friends with each other (the G_2 vs. G_1 comparison of Figure 1) — and where that tightening of his social network exceeds the average tightening of networks in that destination (as might occur around the holidays, for instance).

- **Treatment of outliers (removing low- and high-degree individuals)**: Our main specifications remove from our sample all individuals (and calls made by individuals) with more than 200 unique contacts in a single month (this represents the 95th percentile). This is intended to remove spammers, calling centers, "public" phones, and large businesses. In robustness tests, we also remove individuals from our regressions with fewer than 2 contacts, to address concerns that individuals with just one or two friends could bias linear regression estimates, and that network support is sometimes considered undefined for individuals with fewer than two contacts.

Summary

The fact that social networks are not randomly assigned makes it difficult to firmly establish the causal effect of networks on migration. In our setting, we exploit the rich data at our disposal to develop an identification and estimation strategy that offers, in our view, a plausible method to study the influence of higher order network structure on migration. Specifically, we make the identifying assumption that $E[\epsilon_{idt}|\pi_{dt}, \mu_i, \eta_k] = 0$, and use the large quantity of data to estimate these fixed effects $(\pi_{dt}, \mu_i, \eta_k)$. This allows us to focus on how *higher order* network structure, conditional on lower order structure, relates to lagged migration decisions. The preceding sections provide evidence in support of this identifying assumption, and against many common alternative explanations for our results. Still, it remains an assumption, and we acknowledge that our identification is not bulletproof.

If this causal interpretation does not seem justified, the analysis nonetheless reveals a striking and hitherto undocumented relationship between social networks and migration. In particular, through all the robustness tests we have run, we consistently find that migrants are more likely to go to places where their social networks have certain types of higher order structure. In particular, migrants are more likely to go to places where their contacts are interconnected. The presence of this positive correlation is accentuated by the fact that migrants are *not* more likely to migrate to places where their networks are more extensive, i.e., where their friends have more unknown friends.

5 Conclusion

Social networks play a critical role in economic decision-making. This paper uses an extremely detailed dataset to understand how networks influence the decision to migrate. Our analysis suggests several new stylized facts about the relationship between social networks and migration. We find that migrants are consistently drawn to locations where their social networks are interconnected but that, perhaps most surprising, the average migrant is *not* drawn to places where their networks are extensive, i.e., where their friends have lots of friends. Additional analysis suggests that this unexpected result may be due to the fact that migrants may have limited information about unknown destinations, and that they may feel competition for the attention of their well-connected friends. In additional, the granularity of our data allows us to document rich heterogeneity in how different types of migrants respond to social networks. For instance, we find that unlike the "average migrant," repeat migrants and long-term migrants are drawn to more extensive networks.

More broadly, these results highlight how new sources of digital data can provide nuanced insight into the role of social networks in consequential economic decisions. In contexts ranging from product adoption (Banerjee et al., 2013) and disease transmission (Keeling and Eames, 2005) to the spread of new ideas and innovations (Rogers, 1962, Kitsak et al., 2010), simple models of information diffusion have seen remarkable success. Such models often support the stylized narrative that the primary function of networks is to diffuse information about economic opportunities (cf. Rees, 1966, Ioannides and Datcher Loury, 2004). However, the patterns revealed by our data are hard to reconcile with these models, and suggest that some of the value of social networks comes from higher-order network interconnections. We thus hope that one broader contribution of this paper is to illustrate how, as rich social network data become available to researchers, those data can be used to test and distinguish between different models of network utility. Likewise, we expect that more nuanced empirical insights derived from such data can in turn help inspire advances in the theory of social networks.

Data Availability

Replication files for this paper are available at `https://doi.org/10.5281/zenodo.10020030`. This includes all of the analysis code required to replicate the tables and figures in this paper, as well as sample data files. The original data used in the paper were derived from mobile phone metadata obtained from a mobile phone operator in Rwanda. Due to privacy and confidentiality restrictions, these raw data cannot be shared publicly. Please see the replication files for instructions on how to access the original data.

References

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

Ambrus, A., M. Mobius, and A. Szeidl, "Consumption risk-sharing in social networks," *American Economic Review*, 2015, *104* (1), 149–82.

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou, "Who's who in networks. Wanted: The key player," *Econometrica*, 2006, *74* (5), 1403–1417.

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

_ , _ , _ , and _ , "Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials," *The Review of Economic Studies*, 2019.

Banerjee, Abhijit V. and Andrew F. Newman, "Information, the Dual Economy, and Development," *The Review of Economic Studies*, 1998, *65* (4), 631–653.

_ , Emily Breza, Arun Chandrasekhar, Esther Duflo, and Matthew O. Jackson, "Come play with me: Experimental evidence of information diffusion about rival goods," in "in" 2012.

Barwick, Panle Jia, Yanyan Liu, Eleonora Patacchini, and Qi Wu, "Information, Mobile Communication, and Referral Effects," Working Paper 25873, National Bureau of Economic Research May 2019.

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

Beaman, Lori, Ariel BenYishay, Jeremy Magruder, and Ahmed Mushfiq Mobarak, "Can network theory based targeting increase technology adoption," *Unpublished Manuscript*, 2015.

Bengtsson, Linus, Xin Lu, Anna Thorson, Richard Garfield, and Johan von Schreeb, "Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti," *PLoS Med*, August 2011, *8* (8), e1001083.

Bertoli, S., J. Fernández-Huertas Moraga, and F. Ortega, "Crossing the border: Self-selection, earnings and individual migration decisions," *Journal of Development Economics*, March 2013, *101*, 75–91.

Bertoli, Simone and Ilse Ruyssen, "Networks and migrants' intended destination," *Journal of Economic Geography*, March 2018.

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

_ **and Nathan Eagle**, "Mobile divides: gender, socioeconomic status, and mobile phone use in Rwanda," in "Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development" ACM 2010, p. 6.

_ **and** _ , "Divided We Call: Disparities in Access and Use of Mobile Phones in Rwanda," *Information Technology and International Development*, 2012, *8* (2), 1–16.

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

_ **, Stephen G. Bronars, and Stephen J. Trejo**, "Self-selection and internal migration in the United States," *Journal of Urban Economics*, September 1992, *32* (2), 159–185.

Bramoullé, Yann, Rachel Kranton, and Martin D'amours, "Strategic interaction and networks," *American Economic Review*, 2014, *104* (3), 898–930.

Burt, Ronald S., *Structural holes: The social structure of competition*, Harvard university press, 1992.

Büchel, Konstantin, Maximilian V. Ehrlich, Diego Puga, and Elisabet Viladecans-Marsal, "Calling from the outside: The role of networks in residential mobility," *Journal of Urban Economics*, September 2020, *119*, 103277.

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

_ **and Matthew O. Jackson**, "The Effects of Social Networks on Employment and Inequality," *The American Economic Review*, June 2004, *94* (3), 426–454.

Card, David, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

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

Carrington, William J., Enrica Detragiache, and Tara Vishwanath, "Migration with Endogenous Moving Costs," *The American Economic Review*, 1996, *86* (4), 909–930.

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

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

Chandrasekhar, Arun G., Cynthia Kinnan, and Horacio Larreguy, "Social Networks as Contract Enforcement: Evidence from a Lab Experiment in the Field," *American Economic Journal: Applied Economics*, October 2018, *10* (4), 43–78.

Chi, Guanghua, Fengyang Lin, Guangqing Chi, and Joshua Blumenstock, "A general approach to detecting migration events in digital trace data," *PLOS ONE*, October 2020, *15* (10), e0239408. Publisher: Public Library of Science.

Chuang, Yating and Laura Schechter, "Social Networks in Developing Countries," *Annual Review of Resource Economics*, October 2015, *7*, 451–472.

Coleman, James S, "Social capital in the creation of human capital," *American journal of sociology*, 1988, *94*, S95–S120.

Comola, Margherita and Mariapia Mendola, "The Formation of Migrant Networks," *Scandinavian Journal of Economics*, 2015, *117* (2), 592–618.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin, "ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects," March 2019.

Dagnelie, Olivier, Anna Maria Mayda, and Jean-François Maystadt, "The labor market integration of refugees in the United States: Do entrepreneurs in the network help?," *European Economic Review*, January 2019, *111*, 257–272.

Deshingkar, Priya and Sven Grimm, *Internal Migration and Development: A Global Perspective*, United Nations Publications, 2005.

Dinkelman, Taryn and Martine Mariotti, "The Long Run Effects of Labor Migration on Human Capital Formation in Communities of Origin," Working Paper 22049, National Bureau of Economic Research February 2016.

Dolfin, Sarah and Garance Genicot, "What Do Networks Do? The Role of Networks on Migration and "Coyote" Use," *Review of Development Economics*, May 2010, *14* (2), 343–359.

Dunbar, Robin I. M., "The social brain hypothesis," *Evolutionary Anthropology: Issues, News, and Reviews*, 1998, *6* (5), 178–190.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker, "Referral-based Job Search Networks," *The Review of Economic Studies*, April 2016, *83* (2), 514–546.

Easley, David and Jon Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, 1 edition ed., New York: Cambridge University Press, July 2010.

Eaton, Jonathan, Samuel S Kortum, and Sebastian Sotelo, "International Trade: Linking Micro and Macro," Working Paper 17864, National Bureau of Economic Research February 2012. Series: Working Paper Series.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund, "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment," *The Quarterly Journal of Economics*, February 2003, *118* (1), 329–357.

Fafchamps, Marcel and Forhad Shilpi, "Determinants of the Choice of Migration Destination," *Oxford Bulletin of Economics and Statistics*, June 2013, *75* (3), 388–409.

FEWS NET Rwanda, "Rwanda Food Security Outlook: July to December 2014," *Famine Early Warning Systems Network Food Security Update*, 2014.

Friedkin, Noah E., "Horizons of Observability and Limits of Informal Control in Organizations," *Social Forces*, September 1983, *62* (1), 54–77.

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

Giulietti, Corrado, Jackline Wahba, and Yves Zenou, "Strong versus weak ties in migration," *European Economic Review*, May 2018, *104*, 111–137.

Granovetter, M. S., "The strength of weak ties," *American journal of sociology*, 1973, *78* (6), 1360.

Greenwood, Michael J., "An Analysis of the Determinants of Geographic Labor Mobility in the United States," *The Review of Economics and Statistics*, 1969, *51* (2), 189–194.

Guiteras, Raymond, James A. Levinsohn, and Ahmed Mushfiq Mobarak, "Demand Estimation with Strategic Complementarities: Sanitation in Bangladesh," Technical Report 13498, C.E.P.R. Discussion Papers January 2019.

Hanson, Gordon H. and Christopher Woodruff, "Emigration and educational attainment in Mexico," Technical Report, Mimeo., University of California at San Diego 2003.

Hare, Denise, "'Push' versus 'pull' factors in migration outflows and returns: Determinants of migration status and spell duration among China's rural population," *The Journal of Development Studies*, February 1999, *35* (3), 45–72.

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

Ioannides, Yannis M. and Linda Datcher Loury, "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, December 2004, *42* (4), 1056–1093.

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

_ , "A typology of social capital and associated network measures," *Social Choice and Welfare*, March 2020, *54* (2), 311–336.

_ **and Asher Wolinsky**, "A strategic model of social and economic networks," *Journal of economic theory*, 1996, *71* (1), 44–74.

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

_ **and Stephen Nei**, "Networks of military alliances, wars, and international trade," *Proceedings of the National Academy of Sciences*, December 2015, *112* (50), 15277–15284.

Jackson, Matthew O and Yves Zenou, "Games on networks," in "Handbook of game theory with economic applications," Vol. 4, Elsevier, 2015, pp. 95–163.

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

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

Keeling, Matt J. and Ken T. D. Eames, "Networks and epidemic models," *Journal of The Royal Society Interface*, September 2005, *2* (4), 295–307.

Kempe, David, Jon Kleinberg, and Éva Tardos, "Maximizing the Spread of Influence Through a Social Network," in "Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining" KDD '03 ACM New York, NY, USA 2003, pp. 137–146.

Kermack, William Ogilvy and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proc. R. Soc. Lond. A*, August 1927, *115* (772), 700–721.

Kinnan, Cynthia, "Distinguishing barriers to insurance in Thai villages," *Journal of Human Resources*, 2019.

_ **, Shing-Yi Wang, and Yongxiang Wang**, "Access to Migration for Rural Households," *American Economic Journal: Applied Economics*, October 2018, *10* (4), 79–119.

Kitsak, Maksim, Lazaros K. Gallos, Shlomo Havlin, Fredrik Liljeros, Lev Muchnik, H. Eugene Stanley, and Hernán A. Makse, "Identification of influential spreaders in complex networks," *Nature Physics*, November 2010, *6* (11), 888–893.

Kranton, Rachel E, "The formation of cooperative relationships," *The Journal of Law, Economics, and Organization*, 1996, *12* (1), 214–233.

König, Michael D., Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti, "Networks in Conflict: Theory and Evidence From the Great War of Africa," *Econometrica*, 2017, *85* (4), 1093–1132.

Lai, Shengjie, Elisabeth zu Erbach-Schoenberg, Carla Pezzulo, Nick W. Ruktanonchai, Alessandro Sorichetta, Jessica Steele, Tracey Li, Claire A. Dooley, and Andrew J. Tatem, "Exploring the use of mobile phone data for national migration statistics," *Palgrave Communications*, March 2019, *5* (1), 1–10. Number: 1 Publisher: Palgrave.

Lenormand, Maxime, Miguel Picornell, Oliva G. Cantú-Ros, Antònia Tugores, Thomas Louail, Ricardo Herranz, Marc Barthelemy, Enrique Frías-Martínez, and José J. Ramasco, "Cross-Checking Different Sources of Mobility Information," *PLOS ONE*, August 2014, *9* (8), e105184.

Ligon, Ethan A. and Laura Schechter, "Motives for Sharing in Social Networks," *SSRN eLibrary*, February 2011.

Lu, Xin, David J. Wrathall, Pål Roe Sundsøy, Md. Nadiruzzaman, Erik Wetter, Asif Iqbal, Taimur Qureshi, Andrew Tatem, Geoffrey Canright, Kenth Engø-Monsen, and Linus Bengtsson, "Unveiling hidden migration and mobility patterns in climate stressed regions: A longitudinal study of six million anonymous mobile phone users in Bangladesh," *Global Environmental Change*, May 2016, *38*, 1–7.

Lucas, Robert E. B., "Internal migration in developing countries," in "Handbook of Population and Family Economics," Vol. 1, Elsevier, January 1997, pp. 721–798.

_ , "Chapter 26 - African Migration," in Barry R. Chiswick and Paul W. Miller, ed., *Handbook of the Economics of International Migration*, Vol. 1 of *Handbook of the Economics of International Migration*, North-Holland, 2015, pp. 1445–1596.

Mahajan, Parag and Dean Yang, "Taken by Storm: Hurricanes, Migrant Networks, and U.S. Immigration," Working Paper 23756, National Bureau of Economic Research August 2017.

McKenzie, David and Hillel Rapoport, "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks," *Review of Economics and Statistics*, June 2010, *92* (4), 811–821.

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

Miller, David and Xu Tan, "Seeking Relationship Support: Strategic network formation and robust cooperation," *working paper*, 2018.

Monderer, Dov and Lloyd S Shapley, "Potential games," *Games and economic behavior*, 1996, *14* (1), 124–143.

Montgomery, James D., "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *The American Economic Review*, 1991, *81* (5), 1408–1418.

Morten, Melanie, "Temporary Migration and Endogenous Risk Sharing in Village India," *Journal of Political Economy*, February 2019, *127* (1), 1–46.

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

_ , “Community Networks and the Process of Development,” *The Journal of Economic Perspectives*, October 2014, *28* (4), 49–76.

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

National Institute of Statistics of Rwanda, “The Evolution of Poverty in Rwanda from 2000 to 2011: Results from the Household Surveys (EICV),” Technical Report, Kigali, Rwanda February 2012.

_ , “Migration and Spatial Mobility,” Technical Report, Kigali, Rwanda January 2014.

Patel, Krishna and Francis Vella, “Immigrant Networks and Their Implications for Occupational Choice and Wages,” *The Review of Economics and Statistics*, October 2012, *95* (4), 1249–1277.

Pindolia, Deepa K., Andres J. Garcia, Zhuojie Huang, Timothy Fik, David L. Smith, and Andrew J. Tatem, “Quantifying cross-border movements and migrations for guiding the strategic planning of malaria control and elimination,” *Malaria Journal*, May 2014, *13* (1), 169.

Rees, Albert, “Information Networks in Labor Markets,” *The American Economic Review*, 1966, *56* (1/2), 559–566.

Rogers, Everett M., *Diffusion of Innovations*, Simon and Schuster, 1962.

Simmel, Georg, *The sociology of georg simmel*, Vol. 92892, Simon and Schuster, 1950.

Sotelo, Sebastian, “Practical aspects of implementing the multinomial pml estimator,” Technical Report, Mimeo 2019.

Todaro, Michael P., “Internal Migration in Developing Countries: A Survey,” in “Population and Economic Change in Developing Countries,” University of Chicago Press, 1980.

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

Ugander, Johan, Lars Backstrom, Cameron Marlow, and Jon Kleinberg, “Structural diversity in social contagion,” *Proceedings of the National Academy of Sciences*, 2012, p. 201116502.

United Nations Population Division, “Cross-national Comparisons of Internal Migration: An Update on Global Patterns and Trends,” *Technical Paper No. 2013/1*, 2013.

Wahba, Jackline and Yves Zenou, “Density, social networks and job search methods: Theory and application to Egypt,” *Journal of Development Economics*, December 2005, *78* (2), 443–473.

Wesolowski, Amy, Caroline O. Buckee, Deepa K. Pindolia, Nathan Eagle, David L. Smith, Andres J. Garcia, and Andrew J. Tatem, "The Use of Census Migration Data to Approximate Human Movement Patterns across Temporal Scales," *PLOS ONE*, January 2013, *8* (1), e52971.

_ , Nathan Eagle, Abdisalan M. Noor, Robert W. Snow, and Caroline O. Buckee, "The impact of biases in mobile phone ownership on estimates of human mobility," *Journal of The Royal Society Interface*, April 2013, *10* (81), 20120986.

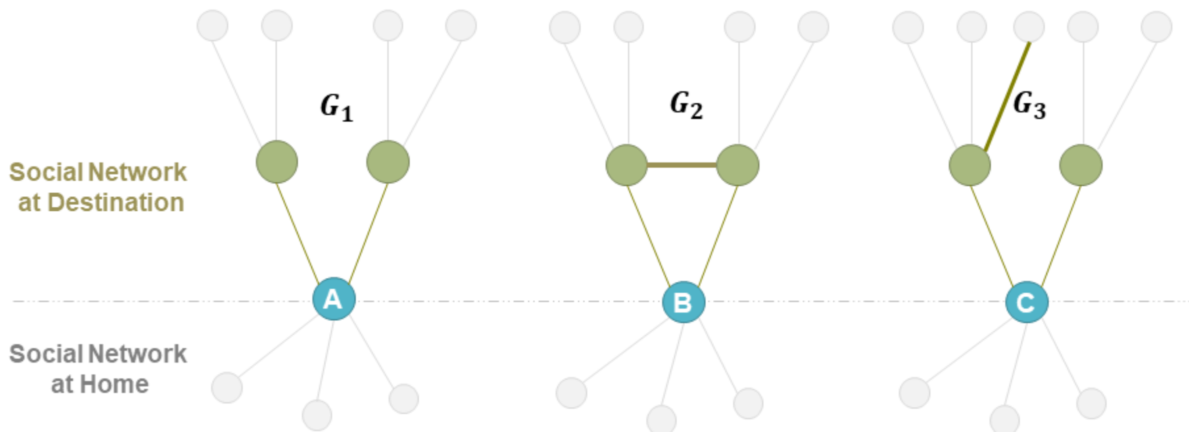
Williams, Nathalie E., Timothy Thomas, Matt Dunbar, Nathan Eagle, and Adrian Dobra, "Measurement of human mobility using cell phone data: developing big data for demographic science," in "Population Association of America Annual Meeting" Citeseer 2013.

Winters, Paul, Alain de Janvry, and Elisabeth Sadoulet, "Family and Community Networks in Mexico-U.S. Migration," *The Journal of Human Resources*, 2001, *36* (1), 159–184.

World Bank Group, *Reshaping Urbanization in Rwanda: Economic and Spatial Trends and Proposals*, Washington, DC: World Bank, 2017.

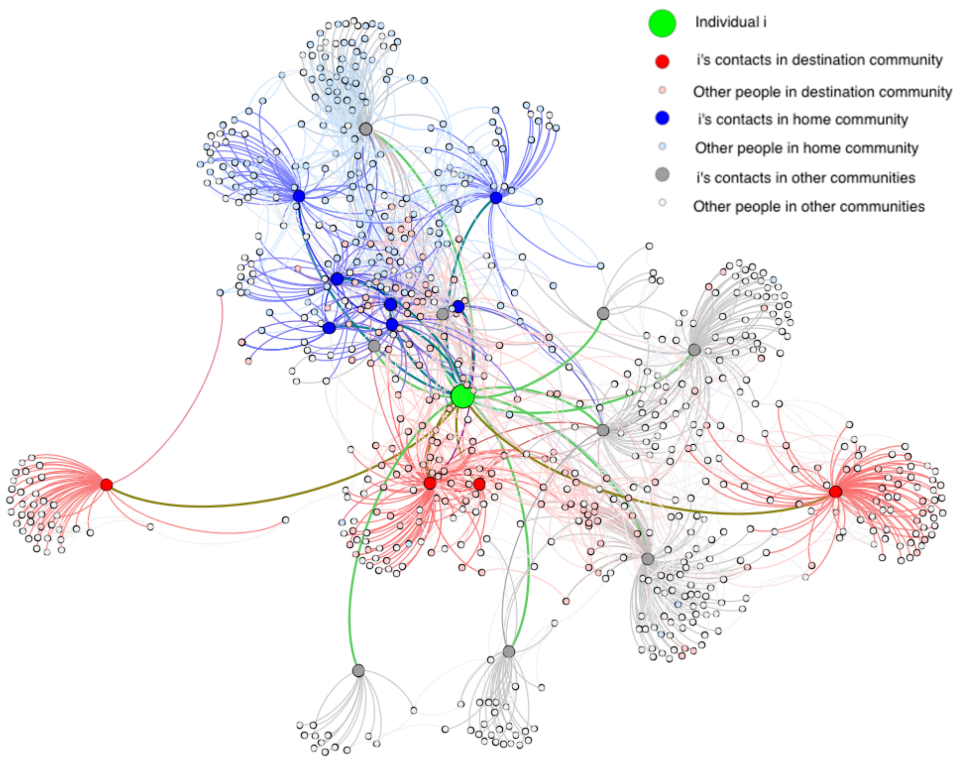
Figures

Figure 1: Schematic diagrams of the social networks of three migrants



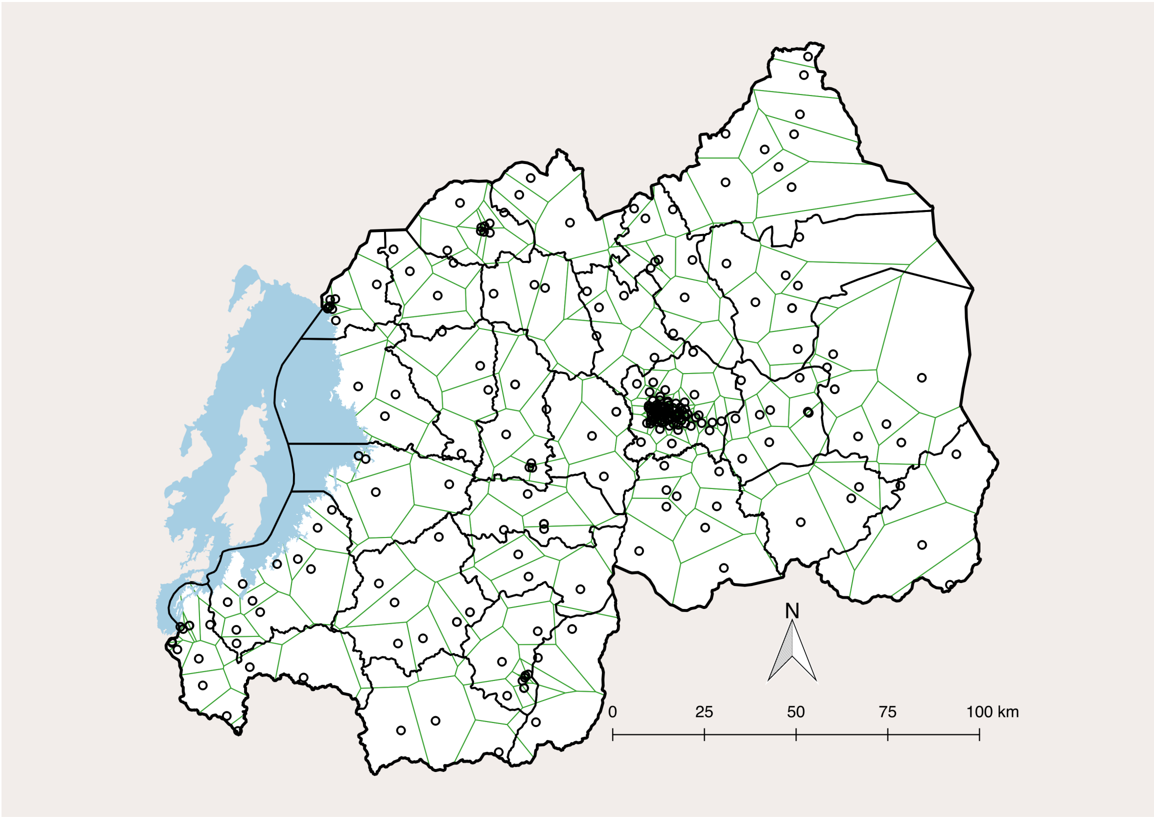
Notes: Each of the blue circles (A, B, C) represents a different individual considering migrating from their home to a new destination. Each individual has exactly three contacts in the home district (grey circles below the dashed line) and two contacts in the destination district (green circles above the dashed line). The social network of these three individuals is denoted by G_1, G_2, and G_3.

Figure 2: The social network of a single migrant



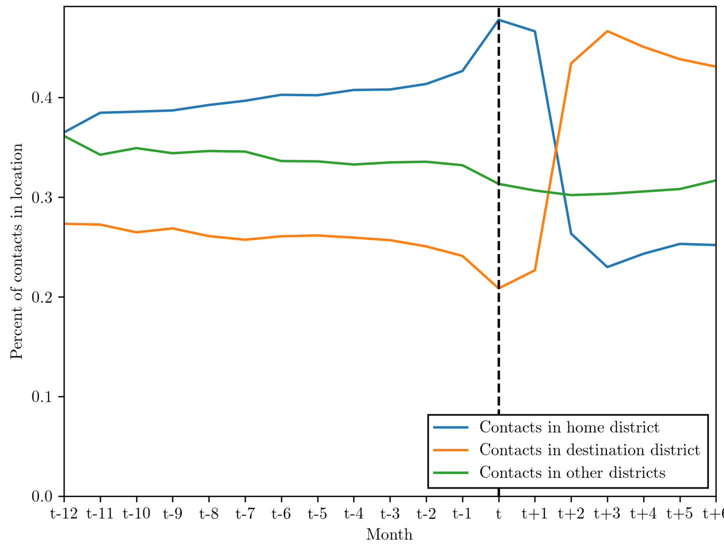
Notes: Diagram shows the social network, as inferred from phone records, of a single migrant i. Nodes represent individuals; edges indicate that two individuals communicated in the month prior to i's migration. Direct contacts of i are shown in blue (for people i's home district), red (for people in i's destination district), and solid grey (for people in other districts). Small hollow 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. All individuals within two hops of i are shown. Nodes are spaced using the force-directed algorithm described in Hu (2005).

Figure 3: Location of all mobile phone towers in Rwanda, circa 2008

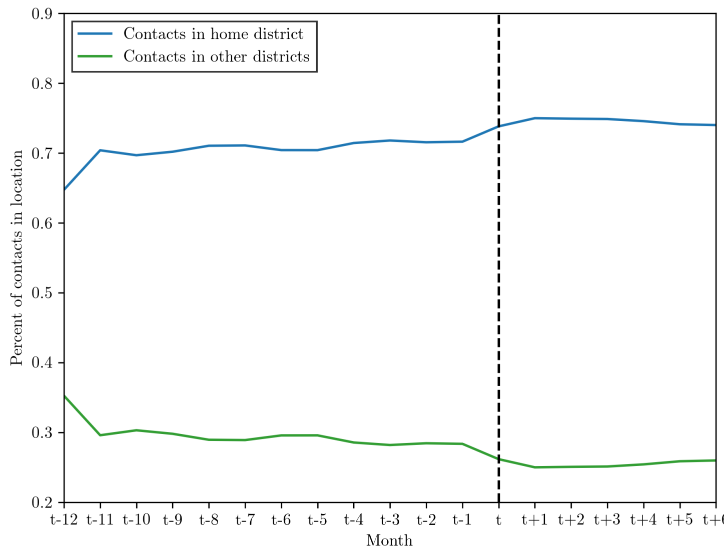


Notes: Black circles indicate cell tower locations. Black lines represent district borders. Green lines show the voronoi polygons roughly divide the country into the coverage region of each tower.

Figure 4: Changes in network structure over time



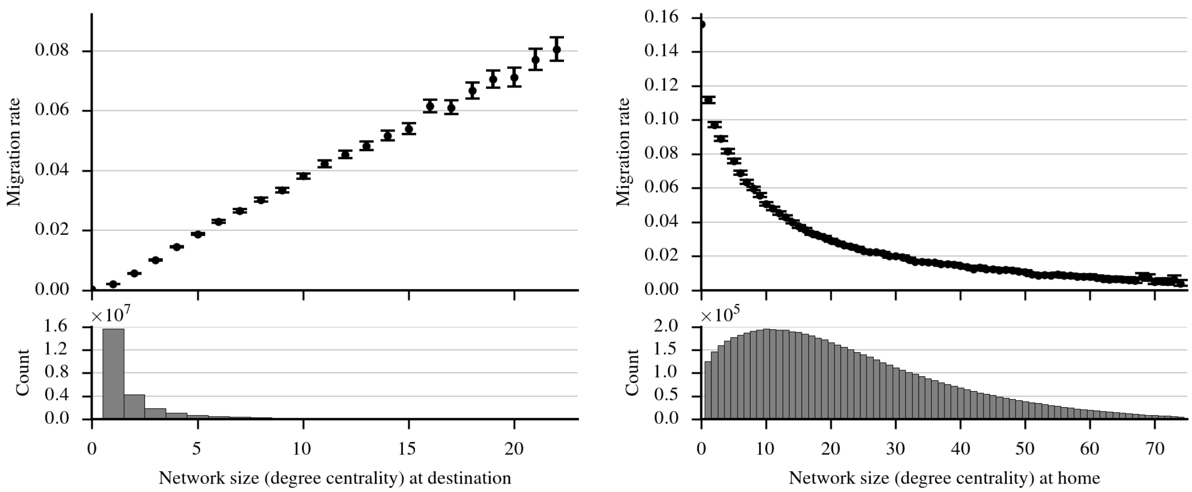
(a) Migrants



(b) Non-Migrants

Notes: Figures show how the network connections of (a) migrants and (b) non-migrants evolves over time. For (a), we draw a random sample of 10,000 migrants, and plot the average percentage of contacts those individuals have in the home, destination, and other districts, in each of the 12 months before and 6 months after migration. The dashed vertical line indicates the date of migration. For (b), we draw a random sample of 10,000 non-migrants by selecting, for each migrant who is sampled to appear in (a) and observed to migrate in month t, a non-migrant from the same reference month t.

Figure 5: Migration and degree centrality (number of unique contacts in network)

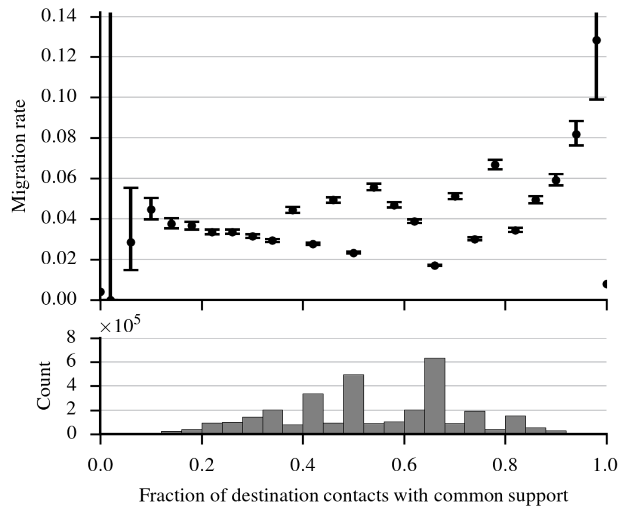


(a) Degree Centrality at Destination

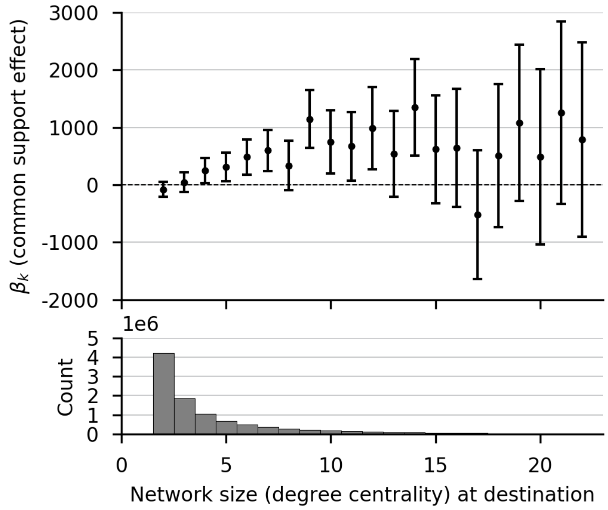
(b) Degree Centrality at Home

Notes: Lower histograms indicate the unconditional degree distribution, i.e., the number of individual-month observations for each degree centrality (i.e., the number of unique contacts) in the (a) destination network and (b) home network. The upper figures show, at each level of degree centrality (in month $t-2$), the average migration rate (in month t). Error bars indicate 95% confidence intervals, using the Wilson Score interval for binomial proportions.

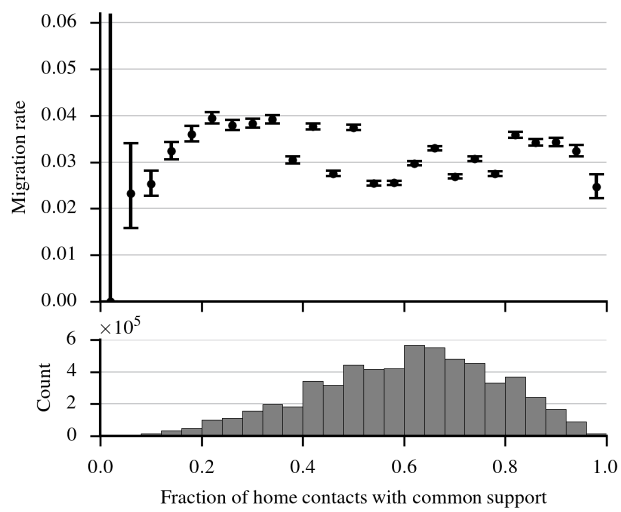
Figure 6: Migration and network "interconnectedness" (friends with common support)



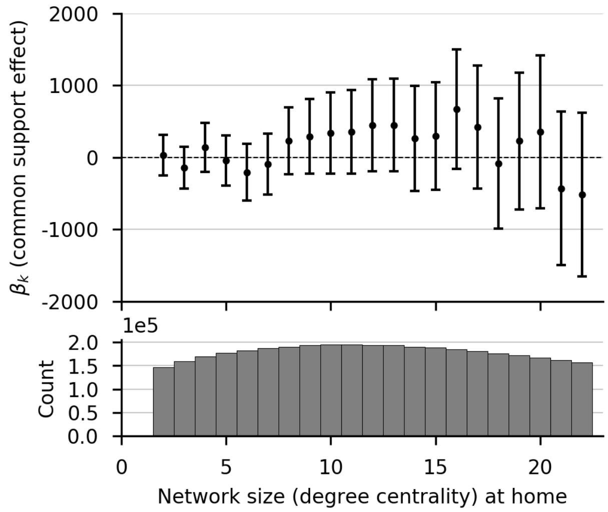
(a) Network support at destination



(b) Network support at destination, by degree



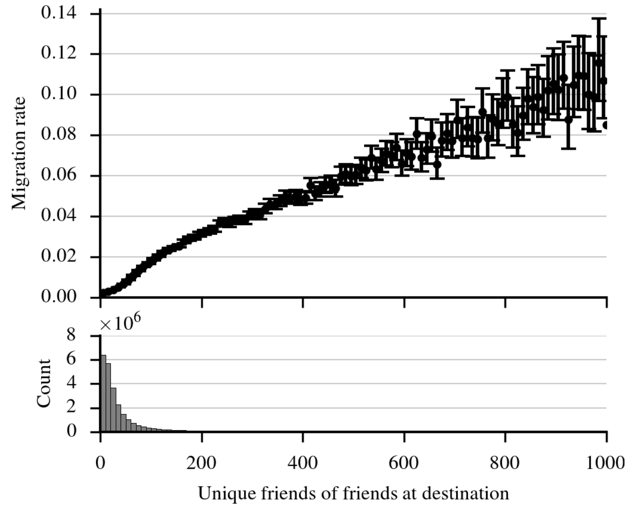
(c) Network support at home



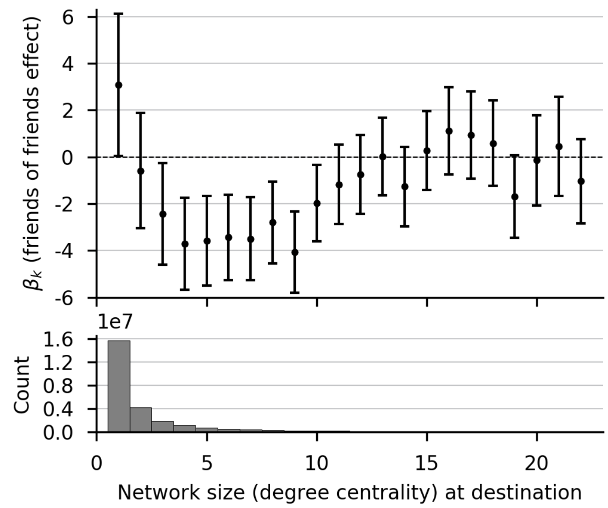
(d) Network support at home, by degree

Notes: Network support indicates the fraction of contacts supported by a common contact (see Section 2.1). In all figures, the lower histogram shows the unconditional distribution of the listed variable. Figures in the left column (a and c) show the average migration rate for different levels of network support. Figures in the right column show the β_k values estimated with model 5, i.e., the correlation between migration and support for individuals with different sized networks (network degree) after conditioning on fixed effects. Top row (Figures a and b) characterizes the destination network; bottom row (Figures c and d) characterizes the home network. Error bars for a and c indicate 95% confidence intervals, using the Wilson Score interval for binomial proportions. Error bars for b and d indicate 95% confidence intervals, two-way clustered by individual and by home-destination-month. Coefficients and standard errors on b and d are multiplied by 1000 to make figures legible.

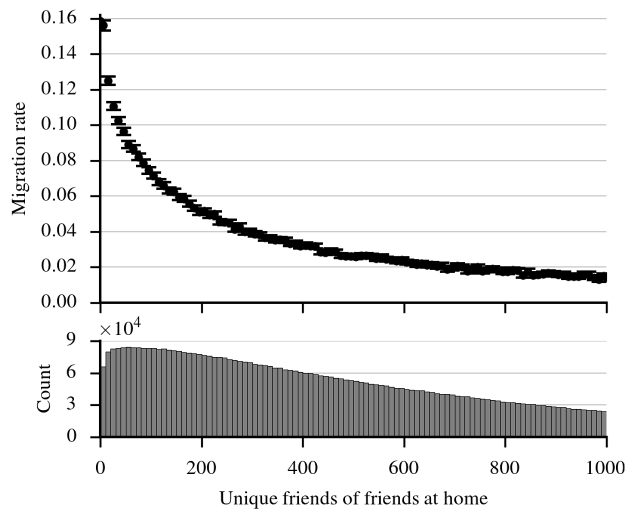
Figure 7: Relationship between migration and "extensiveness" (unique friends of friends)



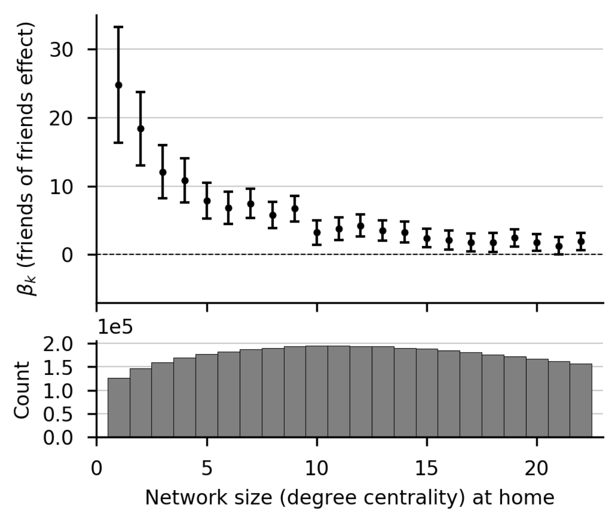
(a) Friends of friends at destination



(b) Friends of friends at destination, by degree



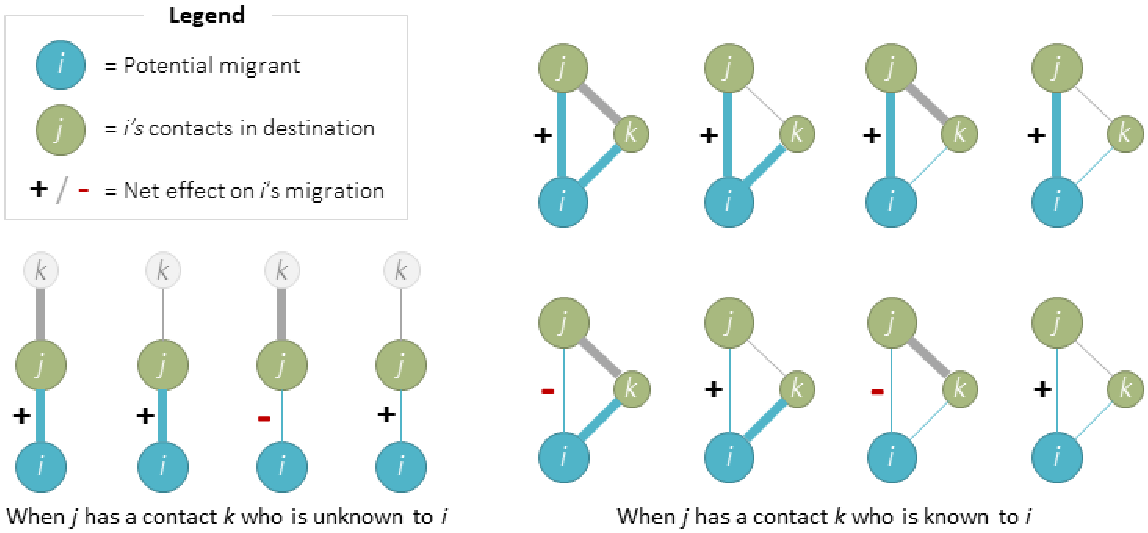
(c) Friends of friends at home



(d) Friends of friends at home, by degree

Notes: Main figures in the left column (a and c) show the average migration rate for people with different numbers of unique friends of friends. Figures in the right column show the β_k values estimated with model 5, i.e., the correlation between migration and unique friends of friends for individuals with different numbers of friends, after conditioning on fixed effects. Top row (Figures a and b) characterizes the destination network; bottom row (Figures c and d) characterizes the home network. Lower histograms show the unconditional distribution of the independent variable. Error bars for a and c indicate 95% confidence intervals, using the Wilson Score interval for binomial proportions. Error bars for b and d indicate 95% confidence intervals, two-way clustered by individual and by home-destination-month. Coefficients and standard errors on b and d are multiplied by 1000 to make figures legible.

Figure 8: The role of (higher order) strong and weak ties in a migrant's network



Notes: Thick edges represent "strong" ties and thin edges represent "weak ties" The +/− signs summarize the effect that j has on i's likelihood of migration, based on the coefficients along the diagonal of Tables A9 and A10.

Tables

Table 1: Summary statistics of mobile phone metadata

	(1) In a single month (January 2008)	(2) Over two years (July 2006 - June 2008)
Number of unique individuals	432,642	793,791
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

Notes: Migration statistics calculated from Rwandan mobile phone data. Column (1) is based on data from a single month; column (2) includes two years of data, potentially counting each individual more than once. A "migration" in this table is defined as occurring when an individual remains in one district for 2 consecutive months and then remains in a different district for the next 2 consecutive months. We denote as urban the three districts in the capital of Kigali; the remaining districts are considered rural. Appendix Table A1 provides additional summary statistics for different types of migration events.

Table 2: Effects of home & destination network structure on migration

	(1)	(2)	(3)
Destination Degree (network size)	260.24***	315.86***	
	(2.48)	(2.53)	
Destination % friends with support	2586.98***	2270.11***	182.84***
	(11.50)	(12.03)	(11.75)
Destination friends of friends	-2.16***	-5.03***	-0.14***
	(0.08)	(0.09)	(0.06)
Home Degree (network size)	41.67***	81.34***	
	(1.15)	(1.27)	
Home % friends with support	845.24***	815.48***	105.62***
	(17.17)	(17.38)	(17.89)
Home friends of friends	2.35***	0.52***	1.38***
	(0.04)	(0.05)	(0.04)
Home district	8229.08***		
	(9.54)		
Observations	184,637,637	184,637,637	184,637,637
Pseudo R^2	0.68	0.68	0.68
Degree fixed effects	No	No	Yes
Destination*Month fixed effects	No	Yes	Yes
Individual*Month fixed effects	Yes	Yes	Yes

Notes: Each observation corresponds to an individual-month-district tuple. Each column indicates a separate regression of a binary variable indicating 1 if an individual i chose to live in district d in month t. Results are estimated using a conditional logit model, using social network characteristics of the location calculated in month $t-2$. "Home District" is a binary variable indicating whether the destination choice in t is i's home in $t-1$. Degree fixed effects, as well as Destination*Month fixed effects, are interacted with the Home District fixed effect. See discussion in Section 3.4. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. Standard errors are clustered by individual. *p<0.1; **p<0.05; ***p<0.01.

Appendices - For Online Publication

A1 Model and structural estimation

A1.1 A model of social capital and migration

We develop a model that links social network structure (in both the home and destination) to subsequent migration decisions. Formally, we say that an individual i receives social capital, or utility for short, $u_i(G)$ from social network G. In deciding whether or not to migrate, the individual weighs the utility of her home network G^h against the utility of the network G^d in the potential destination, and migrates if the difference is greater than an idiosyncratic cost ε_i that can reflect, among other things, wage differentials and i's idiosyncratic costs of migrating.

$$u_i(G^d) > u_i(G^h) + \varepsilon_i. \tag{7}$$

How people derive utility from their social networks — and equivalently, how we parameterize $u_i(G)$ — is not known ex ante. The network theory literature links this network-based utility to the topological structure of the underlying network (i.e., to the configuration of connections between nodes in the network). Jackson (2020) summarizes this work, and provides a taxonomy of *social capital* in networks. We focus on two types of social capital that prior studies have emphasized in the decision to migrate: *information capital* and *cooperation capital*.

A1.1.1 Information capital: competition and 'extensiveness'

A robust theoretical and empirical literature suggests that the value of a social network stems, at least in part, from its ability to efficiently transmit information (see footnote 8). We build on recent efforts by Banerjee et al. (2013) to model this information capital as an information sharing process with possible loss of information. It is worth noting that Banerjee et al. (2013) study a seeding process in which an agent is injected with one unit of information, and this agent's diffusion centrality measures the impact of his information to the network. We study a receiving process in which each agent is initially endowed with one piece of information. The nature of information could be about job openings, the work environment, compensation, and so on. We seek to measure how much information an agent could receive from the network. Using the same information sharing process as Banerjee et al. (2013), we will show that the (proxy) measure we seek turns out to be diffusion centrality, because the flow of information is symmetric.[30]

In this model, a population of N agents, $N = \{1, \ldots, n\}$, are connected in an undirected network. Let G be the adjacency matrix of the network: $G_{ij} = 1$ if i and j are connected and otherwise $G_{ij} = 0$, including $G_{ii} = 0$. Denote agent i's neighbors as $N_i = \{j : G_{ij} = 1\}$, and agent i's degree as $d_i = |N_i|$, which is the number of his or her neighbors in N_i. Agents meet with their neighbors repeatedly, and when they meet at period t, nodes receiving the information at $t-1$ transmit it to their neighbors with probability $q \in (0, 1)$.

In this benchmark model of information sharing, more extensive networks — where an individual has a large number of short-distance indirect neighbors — provide additional utility. We extend this model by allowing for the possibility that neighbors might compete for the attention of their common neighbor. This is motivated by our earlier observation that more extensive destination networks are not positively correlated with migration, and with the evidence that suggests possible rivalry for attention (see Section 4.3).

We model the source of competition for attention as costly socializing with neighbors, so when an agent has more neighbors, he or she may spend less time with each neighbor. Formally, let cQ^ω be the cost of spending Q amount of time on communicating with neighbors. We assume each agent does not possess additional information about neighbors (such as their degrees), so each agent evenly distributes the total amount of time Q to her d neighbors, that is, she spends $q = Q/d$ amount of time with each neighbor. Her utility from communicating with neighbors is given by $d \cdot v(Q/d)^\beta - cQ^\omega$, in which she receives a value of $v(Q/d)^\beta$ from spending Q/d amount of time with each neighbor, and the total cost of spending time Q is cQ^ω. We assume the cost is convex in time $\omega \geq 1$, the value is concave in time $\beta \leq 1$, and they cannot be linear at the same time $\omega > \beta$. The agent's maximization problem becomes

$$\max_Q \; dv(Q/d)^\beta - cQ^\omega. \tag{8}$$

To maximize her utility, the agent's optimal time per neighbor is

$$Q/d = \frac{1}{d^\lambda} \left(\frac{\beta v}{\omega c} \right)^{\frac{1}{\omega - \beta}}, \text{ where } \lambda = \frac{\omega - 1}{\omega - \beta} \in [0, 1]. \tag{9}$$

[30]Our (proxy) measure counts the sum of expected number of times an agent hears about each piece of information, similar to Banerjee et al. (2013).

Notice that if the cost is linear ($\omega = 1$), then the marginal cost of communicating with one neighbor does not increase when the agent has more other neighbors. Thus, the optimal time per neighbor is independent of her degree: $\lambda = 0$. On the other hand, if the value is linear ($\beta = 1$), time with neighbors are perfect substitutes. Then, the total amount of time Q is independent of her degree, which is then evenly split among neighbors: $\lambda = 1$.

Motivated by this simple exercise, we let the interaction between each pair of linked agents ij depend on their degrees. In particular, let the frequency of their interaction be discounted by $\frac{1}{d_i^\lambda d_j^\lambda}$ due to possible competition for attention. During information sharing, each agent initially has one unit of information. In each period from period 1 up to period T, each agent i shares $\frac{1}{d_i^\lambda d_j^\lambda}q$ fraction of her current information to each neighbor j. Notice that $q < 1$ is the original information sharing discount in Banerjee et al. (2013) that is due to loss of information. Then, agent i's information capital is a sum of all the information that she can receive from the network. The vector of agents' *information capital* is the modified diffusion centrality vector, modified to include possible competition for attention. Then,

$$DC(G; q, \lambda, T) \equiv \sum_{t=1}^{T} (q\tilde{G})^t \cdot \mathbf{1}, \quad \text{and} \quad \forall ij, \; \tilde{G}_{ij} = \frac{1}{d_i^\lambda d_j^\lambda} G_{ij}. \tag{10}$$

When $\lambda = 0$, this is the original diffusion centrality, which assumes that in each period information is shared with probability q and information is useful if heard within T periods. When $\lambda > 0$, there is a tradeoff between the positive discounted utility from indirect neighbors and a negative effect due to competition with them for direct neighbors' attention. We say the *distance* between two agents is 2, if they are not connected but share a common neighbor. To highlight the tradeoff, we compare an agent's information capital with and without a distance-2 neighbor. Let $G \setminus \{k\}$ be the resulting network matrix removing its kth row and kth column.

Proposition 1. *Consider $T = 2$. For any agent i and any of her distance-2 neighbors k, there exists a threshold $\lambda_{ik} \in (0, 1)$ such that when $\lambda < \lambda_{ik}$, agent i's information capital is higher in network G than that in $G \setminus \{k\}$, and when $\lambda > \lambda_{ik}$, the comparison is reverse.*

All proofs are in Appendix A2. This result shows that when λ is small, having more neighbors of neighbors increases one's information capital, whereas when λ is large (i.e., close to one), having more indirect neighbors decreases one's information capital. Thus, λ allows for extensive networks to be either beneficial or harmful.

A1.1.2 Cooperation capital: support and 'interconnectedness'

Social networks also facilitate interactions that benefit from community cooperation and enforcement, such as risk sharing and social insurance. We model this dynamic following the setup of Ali and Miller (2016), which highlights the importance of *supported* relationships, where a link is supported if the two nodes of the link share at least one common neighbor (see also Jackson et al. (2012) and Miller and Tan (2018)).

As before, a population of N players are connected in an undirected network G, with $ij \in G$ and $ji \in G$ if agent i and j are connected (we abuse the notation of G slightly, which differs from the matrix format in the information model). 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 $\delta > 0$. When they meet, instead of sharing information, agent i and j now choose their effort levels a_{ij}, a_{ji} in $[0, \infty)$ as their contributions to a joint project.[31] 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. We normalize the net value of effort a as $b(a) - c(a) = a$, and assume the cost function c is a smooth function satisfying $c(0) = 0$ and the following assumption.

Assumption **1.** *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" $c(a)/a$ is strictly increasing.*

Strict convexity with the limit condition guarantees that in equilibrium effort is bounded. Increasing relative cost means a player requires proportionally stronger incentives to exert higher effort. All players share a common discount rate $r > 0$, and the game proceeds over continuous time $t \in [0, \infty)$.

As has been documented in several different real-world contexts, we assume agents have only local knowledge of the network. Specifically, we assume each agent only observes her local neighborhood, including her neighbors, and the links among these neighbors (in additional to her own links). To be precise, it is common knowledge that agent i observes each $j \in g_i \equiv \{i\} \cup N_i$, and all links in $G_i \equiv \{jk : j, k \in g_i\}$. In addition, we consider local monitoring, such that each agent learns about her neighbors' deviation (shirking behavior), and this information travels instantly.[32]

[31] The variable-stakes formulation is adopted from Ghosh and Ray (1996) and Kranton (1996).

[32] The local monitoring is stronger than the private monitoring in Ali and Miller (2016). It allows us to characterize the optimal equilibrium for any network under only local knowledge of the network, the counterpart of which is unknown with private monitoring (to the best of our knowledge), with the exception that Ali and Miller (2016) find the optimal equilibrium when the network is a triangle.

To begin, we seek to minimize contagion of deviation to the rest of the society off the equilibrium path, which follows from Jackson et al. (2012).

Definition 1. *A strategy profile is* ***robust*** *if an agent's deviation only affects partnerships involving herself and between her neighbors.*

Our first result shows that high levels of cooperation can be sustained in a robust manner, with agents needing only local information about the network and other agents' behavior.

Proposition 2. *For any network G, there exists a robust equilibrium of repeated cooperation that maximizes each agent's utility subject to agents' local knowledge of the network.*

Intuitively, each partnership ij uses the maximal level of effort subject to their shared common knowledge of the network. This maximal level of effort depends on the level of efforts i and j can sustain with each of their common neighbors k, which in turn depends on the level of efforts $\{i, j, k\}$ can sustain with their common neighbors l, and so on. Thus, this problem can be solved inductively, starting from the effort level of the largest clique(s) within $g_{ij} = g_i \cap g_j$, which always exists because the population is finite.

However, the optimal equilibrium in Proposition 2 could demand a high cognitive ability and a lot of computational capacity to solve, because one needs to solve (interdependent) effort levels for all subsets of neighbors in her local network. To address this concern, we instead focus on a simple equilibrium strategy profile that maintains the desired properties and sustains high levels of cooperation from the network enforcement.

To do so, we introduce two benchmark cooperation levels. The first one is *bilateral cooperation*, the maximal cooperation attainable between two partners without the aid of community enforcement.

Bilateral cooperation Consider a strategy profile in which, on the path of play, each agent of the partners exerts effort level a if each has done so in the past; otherwise, each exerts zero effort. The equilibrium path incentive constraints are:

$$b(a) \leq a + \int_0^\infty e^{-rt}\delta a dt. \tag{11}$$

The bilateral cooperation level a^B is the effort level that binds the incentive constraint. Since the grim trigger punishment is a minmax punishment and each partner's effort relaxes the other partner's incentive constraint, these are the maximum efforts that can be supported by any stationary equilibrium that does not involve community enforcement.

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

$$b(a) \leq a + 2\int_0^\infty e^{-rt}\delta a dt. \tag{12}$$

The incentive constraint is binding at effort level a^T. Notice that the future value of cooperation is higher in a triangle because there are two ongoing partnerships for each agent, so it can sustain higher level of efforts $a^T > a^B$ and everyone gets a strictly higher utility.

We characterize a particularly simple equilibrium strategy profile that further highlights the value of supported links. Recall that a link ij is *supported* if there exists k such that $ik \in G$ and $jk \in G$; i.e., if i and j have at least one common friend.

COROLLARY **1.** *There exists a robust equilibrium in which any pair of connected agents cooperate on a^T if the link is supported, and on a^B otherwise.*

As the triangular level of effort can be sustained by three fully-connected agents, this strategy profile is robust. For example, consider a triangle ijk plus a link jk'. Even if k' has shirked on j, which reduces the value j gets from the partnership jk', it does not damage j's incentive to cooperate in the triangle ijk because it can sustain a^T by itself.

A1.2 Complete model

We now return to the migration decision. In equation (13),

$$u_i = U(u_i^I, u_i^C), \tag{13}$$

we assume that i's utility from a network contains information capital and cooperation capital (u_i^I and u_i^C); here, we further assume that the utility can be expressed as a linear combination of these two capitals. This stylized formulation is not meant to imply that u^I and u^C are orthogonal or that other aspects of the network do not weigh in the decision to migrate. Rather, this linear combination is intended to provide a simple benchmark that contrasts two archetypical properties of network structure, which we can also estimate with our data. Appendix A2.1 develops a more general model of network utility, based on a network game approach, which allows for more complex interactions among agents (for instance that an individual's utility may be affected by her position in the global network as well as her local network structure).[33] Appendix A2.2.2 shows that similar results obtain when we consider a log-linear (Cobb-Douglas) utility function.

As outlined in Section A1.1.1, we say that agent i's information capital is proportional to their modified diffusion centrality $DC_i(q, \lambda, T)$, which is the i-th element of the vector in (10). We derive i's cooperation capital from Corollary 1 in Section A1.1.2, which implies that supported links are more valuable than unsupported links:

$$u_i^C = u_1 d_i^{NS} + u_2 d_i^S, \tag{14}$$

where d_i^{NS} is the number of i's unsupported links, d_i^S is the number of i's supported links, u_1 is the utility of cooperating on an unsupported link, and u_2 is the utility of cooperation on a supported link.

The overall utility is thus

$$u_i = u_0 DC_i(q, \lambda, T) + u_1 d_i^{NS} + u_2 d_i^S. \tag{15}$$

We will use this model to contrast the value of information capital against the value of cooperation capital, so we replace the parameters (u_0, u_1, u_2) by (π^I, π^C, α) and rewrite the overall utility:

$$u_i = \pi^I DC_i(q, \lambda, T) + \pi^C \left(d_i + \alpha d_i^S\right). \tag{16}$$

Substituting (16) into the original migration decision (7), we have

$$\begin{aligned} &\pi^{I,d} DC_i(G^d; q, \lambda, T) + \pi^{C,d} \left(d_i(G^d) + \alpha^d d_i^S(G^d)\right) \\ &\qquad > \pi^{I,h} DC_i(G^h; q, \lambda, T) + \pi^{C,h} \left(d_i(G^h) + \alpha^h d_i^S(G^h)\right) + \varepsilon_i. \end{aligned} \tag{17}$$

Notice that we allow agents to have different weights $(\pi^{I,d}, \pi^{C,d}, \pi^{I,h}, \pi^{C,h})$ for the home and destination networks, because it is possible that the relative value of information and cooperation is different in a home network than in a destination network. For the same reason, we allow α to differ between home and destination networks. However, we assume

[33] The network game approach follows in the tradition of Ballester et al. (2006), who use a network game to identify the key player, and König et al. (2017), who study strategic alliances and conflict. This approach is formally attractive, but since each agent's utility depends on their position and the entire network structure, it could not be realistically computed on our data. (As a point of comparison, calibration of the far simpler model (13) takes several days to complete, even after being parallelized across a compute cluster with 96 cores). See also Guiteras et al. (2019) for a related structural approach to dealing with network inter-dependencies.

(q, λ, T) are the same for home and destination networks, because they capture properties of the network that are common across agents and over which the agent has no direct control.[34]

A1.3 Model parameterization

We use the migration decisions made by several hundred thousand migrants over a 4.5-year period to estimate the parameters of model (17). The estimation proceeds in two steps. First, we draw a balanced sample of migrants and non-migrants by selecting, for every migrant who moves from h to d in month t, a non-migrant who lived in h in month t, had ≥ 1 contacts in d, but remained in h after t. This provides a total sample of roughly 270,000 migrants and non-migrants.

Second, we use simulation to identify the set of parameters that maximize the likelihood of generating the migration decisions observed in the data. The structural parameters of primary interest are λ, which we interpret as a measure of the competition or rivalry in information transmission; (α^h, α^d), the added value of a supported link, above and beyond the value of an unsupported link at home and in the destination; and the scaling coefficients $(\pi^{I,d}, \pi^{C,d}, \pi^{I,h}, \pi^{C,h})$, which together indicate the relative importance of information capital and cooperation capital at home and in the destination. We normalize $\pi^{C,h} = 1$, and follow Banerjee et al. (2013) by setting q equal to the inverse of the first eigenvalue of the adjacency matrix, $\mu_1(G)$, and $T = 3$.[35] Since a very large number of combinations of possible parameters exist, we use an iterative grid-search maximization strategy where we initially specify a large set of values for each parameters, then focus and expand the search around local maxima.[36]

To provide more intuition for the model estimation process, Figure A10 shows the estimation plots for the parameters of the model. To produce these figures, we take all possible combinations of 6 parameters, resulting in roughly 50,000 different parameter vectors. We then simulate the migration decisions of the 270,000 migrants and non-migrants using

[34] In particular, the 'destination' network of a given migrant is actually the 'home' network for all of the contacts that live in that destination.

[35] When we treat q as a free parameter and estimate it via MLE, the likelihood-maximizing value of q is very close to $1/\mu_1(G)$. Banerjee et al. (2013) show that this approach to measuring diffusion centrality closely approximates a structural property of "communication centrality." However, we cannot directly estimate this latter property on our empirical network, which contains hundreds of thousands of nodes and tens of millions of edges.

[36] Specifically, for each possible set of parameters $< \lambda, \alpha^d, \alpha^h, \pi^{I,d}, \pi^{C,d}, \pi^{I,h} >$, we calculate the utility of the home and destination network for each migrant, and the change in utility after migration. If the change in utility of migration is positive, we predict that individual would migrate. We choose the set of parameters that minimizes the number of incorrect predictions.

model (17), and calculate the percentage of correct classifications. The figures show the the marginal distributions over a single parameter of the accuracy for the top percentile of parameter vectors. In most cases, the likelihood function is concave around the global maximum.

The structural model is largely being identified by the same variation that drives the reduced-form results. For instance, 97.5% of the variation in the total simulated utility of the destination network can be explained by the three main measures of network structure used in Section 4.[37] This shows that when the rivalry parameter λ is optimally chosen for the structural model, the average effect of one's second-neighborhood becomes negative.

A1.4 Parameterization results

Estimation of the model yields several results. First, we find an optimal value of the rivalry coefficient at $\lambda = 0.5$, as shown in the first panel Figure A10. This suggests a significant departure from the benchmark information diffusion model of Banerjee et al. (2013): having friends who have many friends can actually reduce the utility that the agent receives from the network. The parameterized value of 0.5 implies that the probability of people sharing information with a neighbor is roughly inversely proportional to the (square root of the) size of their social networks. For instance, revisiting individuals A and C from Figure 1 (and assuming a two-period transmission model), with the parameterized $\lambda = 0.5$, we expect that A would receive 1.17 times the information capital as C. By contrast, the benchmark model with $\lambda = 0$ would imply that A would receive slightly less (0.99 times) information capital than C.

Second, using the information diffusion measure with the optimally parameterized rivalry coefficient, we find that the total utility from u_i^I (loosely, the 'information capital') and the total utility from u_i^C (loosely, the 'cooperation capital') contribute relatively evenly to the agent's total utility from the network. This can be seen most clearly in Figure A11, which shows the distribution of predicted utility from u_i^I and u_i^C for each of the individuals used to estimate the simulation. The bulk of this distribution lies around the 45-degree line, which is where $u_i^I = u_i^C$. This result is perhaps surprising given the reduced-form results presented in Section 4, which suggest that friends of friends in the destination have an insignificant (or negative) effect on the migration decision. However, a critical difference between the reduced

[37] Specifically, we regress the total *simulated* utility in the destination network, using the parameterized structural model, on three 'reduced-form' properties of the individual's social network: the destination degree centrality, the number of unique destination friends of friends, and the destination network support (see Section 2.1 for definitions). In this linear regression (no fixed effects), $R^2 = 0.975$.

form and structural results is that the structural results allow for rivalry in information transmission. To further confirm that it is the rivalry parameter drives this difference, we re-estimate a version of model (17) where the rivalry coefficient is fixed at $\lambda = 0$. In other words, we use the original diffusion centrality (without λ) to measure the information capital and redo the whole simulation to identify the likelihood-maximizing set of parameters. We find that information capital (as the original diffusion centrality) contributes very little to total network utility, such that the predicted utility from information is substantially less than from cooperation $u_i^I << u_i^C$.

Third, and consistent with previous results, we find that supported links are valued more than unsupported links. This can be observed in the calibration plots for α^D and α^H in Figure A10. In particular, $\alpha^d = 5$ implies that one supported link in the destination is six times as valuable as an unsupported link in the destination, and similarly, $\alpha^h = 1$ implies that one supported link at home is twice as valuable as an unsupported link at home.

Taken together, the structural estimates provide a micro-founded validation of the reduced-form results described earlier. This is an important step, since the reduced form results are based on statistical properties of networks that are correlated in complex ways, which cannot be easily accounted for in a regression specification. The model parameterization also provides independent support for the presence of some degree of rivalry in information diffusion — a possibility that was suggested by the heterogeneity discussed in Section 4.3, but only directly tested through structural estimation.

As a final step, Appendix A2.2 examines the robustness of the parameterization results. In particular, we allow for the migration decision to include an average migration cost τ, which acts as a linear threshold that is constant across people, in addition to the idiosyncratic error that varies with each individual:

$$u_i(G^d) > u_i(G^h) + \tau + \varepsilon_i. \tag{18}$$

Separately, instead of the linear form of (17), we consider a Cobb-Douglas utility function which implies a log-linear combination of information capital and cooperation capital. Equation (17) becomes

$$\begin{aligned} &\pi^{I,d} \log DC_i(G^d; q, \lambda, T) + \pi^{C,d} \log \left(d_i(G^d) + \alpha^d d_i^S(G^d)\right) \\ &\quad > \pi^{I,h} \log DC_i(G^h; q, \lambda, T) + \pi^{C,h} \log \left(d_i(G^h) + \alpha^h d_i^S(G^h)\right) + \varepsilon_i. \end{aligned} \tag{19}$$

Results in Appendix A2.2 show that the key qualitative results persist under these alternative specifications of model (17).

A1.5 Proofs

Proof of Proposition 1: Consider any agent i and any of her distance-2 neighbors k, and let $G' = G \setminus \{k\}$. To show the existence of such threshold λ_{ik}, it is sufficient to show the following three parts are true. First, when $\lambda = 0$, agent i's diffusion centrality is higher in network G than that in network G'. This is straight forward, because when there is no competition among neighbors, distance-2 neighbors always increase the diffusion centrality which is a sum of information one gets from her neighbors and distance-2 neighbors. Second, when $\lambda = 1$, agent i's diffusion centrality is lower in network G than that in network $G \setminus \{k\}$. Third, the difference in diffusion centrality for any given q (recall $T = 2$)

$$DC_i(G; \lambda, q) - DC_i(G'; \lambda, q)$$

decreases in λ.

For the second part, let $\lambda = 1$ and let agent j be one of i's neighbors who are connected to agent k. Let d_j be agent j's degree in network G, which is at least two since he or she is connected to both i and k. The information capital agent i gets from agent j in network G is then (recall $\lambda = 1$)

$$DC_{ij}(G; q) = q\frac{1}{d_i d_j} + q^2 \sum_{h \in N_j} \frac{1}{d_i d_j^2 d_h}.$$

The first term is the direct information i gets from j, and the second term is the indirect information i gets from j's neighbors. On the other hand, without agent k, the information capital agent i gets from agent j is

$$DC_{ij}(G'; q) = q\frac{1}{d_i(d_j - 1)} + q^2 \left(\sum_{h \in N_j \setminus g_k} \frac{1}{d_i(d_j - 1)^2 d_h} + \sum_{l \in N_j \cap N_k} \frac{1}{d_i(d_j - 1)^2(d_l - 1)} \right).$$

Without agent k, agent j's degree decreases by one and so does any of j and k's common neighbors l. Also, agent i on longer gets indirect information from k, which is reflected as

$(N_j \setminus g_k) \cup (N_j \cap N_k) = N_j \setminus \{k\}$. We have,

$$\begin{aligned}
&DC_{ij}(G'; q) - DC_{ij}(G; q) \\
&\geq q\left(\frac{1}{d_i(d_j-1)} - \frac{1}{d_i d_j}\right) + q^2\left(\sum_{h\in N_j\setminus\{k\}}\left(\frac{1}{d_i(d_j-1)^2 d_h} - \frac{1}{d_i d_j^2 d_h}\right) - \frac{1}{d_i d_j^2 d_k}\right) \\
&\geq q\left(\frac{1}{d_i(d_j-1)} - \frac{1}{d_i d_j}\right) - q^2\frac{1}{d_i d_j^2 d_k} \\
&= q\frac{1}{d_i(d_j-1)d_j} - q^2\frac{1}{d_i d_j^2 d_k} \\
&> 0.
\end{aligned}$$

This is true for all $j \in N_i \cap N_k$. So the second part is true that when $\lambda = 1$, agent i's diffusion centrality in network G' is higher.

Third, we consider the difference in agent i's diffusion centrality from neighbor j:

$$\begin{aligned}
&DC_{ij}(G'; \lambda, q) - DC_{ij}(G; \lambda, q) \\
&= q\left(\frac{1}{d_i^\lambda(d_j-1)^\lambda} - \frac{1}{d_i^\lambda d_j^\lambda}\right) - q^2\frac{1}{d_i^\lambda d_j^{2\lambda} d_k^\lambda} + q^2\sum_{h\in N_j\setminus g_k}\left(\frac{1}{d_i^\lambda(d_j-1)^{2\lambda}d_h^\lambda} - \frac{1}{d_i^\lambda(d_j)^{2\lambda}d_h^\lambda}\right) \\
&\quad + q^2\sum_{l\in N_j\cap N_k}\left(\frac{1}{d_i^\lambda(d_j-1)^{2\lambda}(d_l-1)^\lambda} - \frac{1}{d_i^\lambda(d_j)^{2\lambda}d_l^\lambda}\right). \qquad (20)
\end{aligned}$$

Clearly, each of the four terms in (20) increases as λ increases. So we prove the third part of the monotonicity of the difference in the two diffusion centrality. ∎

Proof of Proposition 2: We construct the equilibrium as follows. Consider the partnership between i and j; the common knowledge they share about the network includes $g_{ij} = g_i \cap g_j$ and $G_{ij} = G_i \cap G_j$.

First, we identify the maximal effort for each clique with m agents.

$$b(a) \leq a + (m-1)\int_0^\infty e^{-rt}\delta a dt,$$

in which $b(a)$ is the gain from deviation and the right hand side is the payoff of each agent from all m agent cooperating at effort a. The effort $a^{c=m}$ binds this inequality.

Then, we claim there exists a maximal effort for the link ij subject to their shared common knowledge. If $g_{ij} = \{i, j\}$, then this maximal effort is $a^{c=2}$, otherwise it can be found by induction as illustrated below. From now on, we focus on the shared local network

(g_{ij}, G_{ij}). We say a subset of agents is *fully-connected* if every agent in the subset is connected to everyone else in the subset. When the largest clique(s) in (g_{ij}, G_{ij}) has $h+2$ agents, then the induction takes h steps:

- In step 1, find the largest clique(s), for example, $g_{ijk_1\ldots k_h}$. Then assign the effort $a(k_m k_l|ijk_1\ldots k_h) = a^{c=h+2}$ to each link $k_m k_l$ within the clique. That is, it is common knowledge among agents in the clique that each link can sustain effort at least $a^{c=h+2}$.

- In step 2, find all subsets of fully-connected agents containing $h+1$ agents, including i and j (this must always hold for all subsets we discuss, so omitted below). For any of them, say $g_{ijk'_1\ldots k'_{h-1}}$, assign $a(k'_m k'_l|ijk'_1\ldots k'_{h-1})$ to each link $k'_m k'_l$ to bind the inequality:

$$b(a) \le a + \int_0^\infty e^{-rt}\delta \left(ha + \sum_{l\in g_{ijk'_1\ldots k'_{h-1}}\setminus\{i,j,k'_1,\ldots,k'_{h-1}\}} a(il|ijk'_1\ldots k'_{h-1}l) \right) dt.$$

  That is, everyone in the clique uses the effort a and for other links that all of them can observe, the effort level is determined in the previous step (step 1).

- ...

- In step η, find all subsets of fully-connected agents containing $(h+3-\eta)$ agents. For any of them, say $g_{ijk''_1\ldots k''_{h+1-\eta}}$, assign $a(k''_m k''_l|ijk''_1\ldots k''_{h+1-\eta})$ to each link $k''_m k''_l$ to bind the inequality:

$$b(a) \le a + \int_0^\infty e^{-rt}\delta \left((h+2-\eta)a + \sum_{l\in g_{ijk''_1\ldots k''_{h+1-\eta}}\setminus\{i,j,k''_1,\ldots,k''_{h+1-\eta}\}} a(il|ijk''_1\ldots k''_{h+1-\eta}l) \right) dt.$$

- ...

- In step $h+1$, the only subset containing 2 agents and including i and j is the set $\{i,j\}$. The effort between them (a^*_{ij}) must bind the inequality:

$$b(a) \le a + \int_0^\infty e^{-rt}\delta \left(a + \sum_{l\in g_{ij}\setminus\{i,j\}} a(il|ijl) \right) dt.$$

By construction, each effort level is the highest effort that is sustainable given the (higher-order) common knowledge of the network. Thus, a^*_{ij} is the maximal effort sustainable between ij subject to their shared knowledge of the network. In addition, as long as no one in g_{ij} has deviated, i and j can sustain a^*_{ij}. Thus, the strategy is robust. ∎

A2 Extension and discussion of the model

A2.1 A network game approach

In the benchmark model, we assume the total utility each agent gets from the network is a linear combination of information capital and cooperation capital as in equation (13). To allow more complex features of network structures to influence the value an agent gets from the social network, one possibility is to consider a network game approach.

Each agent i chooses an action a_i, which could be socializing with friends, cooperating with them or both. Let $\mathbf{a} = (a_1, \ldots, a_n)$ be the strategy profile. We use the matrix format of a network G, such that $G_{ij} = G_{ji} = 1$ when i and j are connected. Let the matrix G^s be the network of links that are supported in the baseline network G, that is $G^s_{ij} = G^s_{ji} = 1$ if and only if ij is supported in G. Agent i derives the following quadratic utility, which has been commonly-used in network games (Jackson and Zenou 2015):

$$u_i(\mathbf{a}, G) = \pi a_i - \frac{a_i^2}{2} + \phi \sum_{j=1}^{n} G_{ij} a_i a_j + \alpha \sum_{j=1}^{n} G^s_{ij} a_i a_j. \tag{21}$$

The first two terms $\pi a_i - \frac{a_i^2}{2}$ represent a linear benefit and a quadratic cost to agent i from choosing a_i. When $\phi > 0$, the third term $\phi \sum_{j=1}^{n} G_{ij} a_i a_j$ reflects the strategic complementarity between neighbors' actions and one's own action.[38] And the last term $\alpha > 0$ reflects the additional complementarity between supported neighbors.

We add two remarks about the utility function. First, the utility differs from a standard network game setup due to the last term, $\alpha \sum_{j=1}^{n} G^s_{ij} a_i a_j$. This is motivated by the theory results in Section A1.1.2 and the empirical results in Section 4 that an agent may derive additional utility from a supported neighbor. Second, if $\alpha = 0$, then the equilibrium action will be in proportion to the diffusion centrality in Section A1.1.1, $DC(G; q, \lambda, T)$ when $q = \phi$, $\lambda = 0$ and $T \to \infty$. In particular, ϕ can be viewed as the information passing probability

[38] While it is unlikely in our setup, ϕ could be negative in some network games, which then reflects the substitution between neighbors' actions and one's own action.

q. The equilibrium action of agent i depends on the entire network structure, including her indirect neighbors and her supported links, and thus, this network approach allows for these network structures to jointly determine the equilibrium utility an agent gets from the network.

Let $\mu_1(G)$ be the spectral radius of matrix G, $\mathbf{I}$ be the identity matrix, and $\mathbf{1}$ be the column vector of 1.

PROPOSITION **3.** *If* $\mu_1(\phi G + \alpha G^s) < 1$, *the game with payoffs* (21) *has a unique (and interior) Nash equilibrium in pure strategies given by:*

$$\mathbf{a}^* = \pi(\mathbf{I} - \phi G - \alpha G^s)^{-1}\mathbf{1}. \tag{22}$$

Consider the first-order necessary condition for each agent i's action:

$$\frac{\partial u_i(\mathbf{a}, G)}{\partial a_i} = \pi - a_i + \phi \sum_{j=1}^{n} G_{ij} a_j + \alpha \sum_{j=1}^{n} G^s_{ij} a_j = 0.$$

This leads to

$$a_i^* = \pi + \phi \sum_{j=1}^{n} G_{ij} a_j^* + \alpha \sum_{j=1}^{n} G^s_{ij} a_j^*. \tag{23}$$

In the matrix form: $\mathbf{a}^* = \pi\mathbf{1} + \phi G\mathbf{a}^* + \alpha G^s \mathbf{a}^*$, which leads to the solution in (22).

A simple way to prove this solution is indeed the unique (and interior) Nash equilibrium, as noted for example by Bramoullé et al. (2014), is to observe that this game is a potential game (as defined by Monderer and Shapley 1996) with potential function:

$$P(\mathbf{a}, G, \phi) = \sum_{i=1}^{n} u_i(\mathbf{a}, G) - \frac{\phi}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} G_{ij} a_i a_j - \frac{\alpha}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} G^s_{ij} a_i a_j.$$

We omit the details of the analogous proof, which can be found in Bramoullé et al. (2014) and Jackson and Zenou (2015).

In the equilibrium, the utility of agent i is given by

$$\begin{aligned} u_i(\mathbf{a}^*, G) &= \pi a_i^* - \frac{a_i^{*2}}{2} + \phi \sum_{j=1}^{n} G_{ij} a_i^* a_j^* + \alpha \sum_{j=1}^{n} G^s_{ij} a_i^* a_j^* \\ &= a_i^* \left(\pi + \phi \sum_{j=1}^{n} G_{ij} a_j^* + \alpha \sum_{j=1}^{n} G^s_{ij} a_j^* \right) - \frac{a_i^{*2}}{2}. \end{aligned}$$

By equation (23), $u_i(\mathbf{a}^*, G) = (a_i^*)^2/2$, which by equation (22) depends on (π, ϕ, α, G). So in this way, we can estimate how an agent's utility depends on the interaction with neighbors ϕ, the added value of a supported link α, and his or her position in the network G.

More generally, the network game can be enriched to capture the possibilities of competition with indirect neighbors, as we modeled in Section A1.1.1. For example, Ballester et al. (2006) consider a global congestion effect by adding the term $-\lambda a_i \sum_{j=1}^{n} a_j$ to each agent i's utility. Using the corresponding equilibrium utility with this congestion λ, one could also estimate the rivalry or competition with indirect neighbors.

A2.2 Robustness of model calibration

Our benchmark model assumes that an individual will migrate if the total utility of the destination network exceeds the total utility of the home network (equation 7), and assumes that the total utility an agent i receives from an arbitrary network G can be expressed as a linear combination of the information capital and cooperation capital of G (equation 13). This highly stylized formulation is intended to contrast, as transparently as possible, what the literature has emphasized are the two main mechanisms through which social networks provide utility. Here, we explore alternative formulations of models (7) and (13), to test the robustness of the calibration results in Section A1.4.

A2.2.1 Fixed migration costs

We first allow for the migration decision (equation 7) to include a fixed threshold (cost) τ, in addition to the idiosyncratic error ε_i:

$$u_i(G^d) > u_i(G^h) + \tau + \varepsilon_i. \tag{24}$$

Here, τ is meant to capture the possibility that all people might share a common aversion to migrating; accounting for this shared cost might help us identify the main parameters of interest.

When model (24) is calibrated with the data, the main observations in Section A1.4 persist. Full calibration plots for all parameters $< \lambda, \alpha^d, \alpha^h, \tau, \pi^{I,d}, \pi^{C,d}, \pi^{I,h} >$ are shown in Figure A12. Most importantly, the optimal value of the rivalry coefficient remains at $\lambda = 0.5$ (top left). Similar to the results presented in the main text, supported links are more valuable than unsupported links (i.e., α^D and α^H are both greater than 0). In particular,

α^D is exactly 5 as in the main model, and α^h decreases slightly from 1 to 0.5. We also find that the total utility from information capital and cooperation capital contribute relatively the same amount to an agent's total utility from the network.The calibration sensitivity plot for the new parameter, τ, is shown in the middle-right panel of Figure A12. This calibration is more noisy, with the optimal calibrated threshold at $\tau = -5$. This is perhaps surprising, since a literal interpretation of τ is as an average migration cost, which should be positive. However, the vast majority of agents in our simulation have considerably larger home networks than destination networks (see the bottom panels of Figure 5); it is likely that the negative τ is offsetting the fact that in our balanced sample home utility generally exceeds destination utility.

A2.2.2 Cobb-Douglas utility

Next, we consider a Cobb-Douglas network utility function, which can be rewritten as the total utility being a log-linear combination of information capital and cooperation capital. Specifically, equation (17) becomes

$$\begin{aligned}\pi^{I,d} \log DC_i(G^d; q, \lambda, T) + \pi^{C,d} \log \left(d_i(G^d) + \alpha^d d_i^S(G^d)\right) \\ > \pi^{I,h} \log DC_i(G^h; q, \lambda, T) + \pi^{C,h} \log \left(d_i(G^h) + \alpha^h d_i^S(G^h)\right) + \varepsilon_i. \end{aligned} \quad (25)$$

We note that the linear utility function and the Cobb-Douglas utility function describe fundamentally different ways that agents value the network. A key difference is that the information capital and cooperation capital are substitutable in the linear utility function, but they are complementary in the Cobb-Douglas utility function. To receive high utility based on the Cobb-Douglas form, an agent needs both a high information capital and a high cooperation capital, while only one is needed based on the linear form.

In results available upon request, we find that the main observations in section A1.4 persist. The log-linear model correctly predicts 68.6% of the migration decisions, which is close to, though slightly below, the accuracy of the linear model, which is 69.5%. As before, the optimal value of the rivalry coefficient remains at $\lambda = 0.5$. Similarly, supported links are more valuable than unsupported links, although the particular values differ from the main model: $\alpha^d = 0.5$ and $\alpha^h = 10$. Cooperation capital contributes roughly twice as much to total utility as information capital, which differs from the equal contribution in the main specification. This shows that the fact that both information capital and cooperation capital contribute significantly to the total social capital is a robust result, but the relative weights of the two may depend on their interactions (substitutes or complementary). We also note that when λ is optimally parameterized, information capital contributes significantly more to total utility than when we remove the possibility for rivalry by setting $\lambda = 0$. In other words, regardless of the specific utility functions, the information capital if in the form of the original diffusion centrality does not contribute to the social capital (relative to the cooperation capital), which further supports the finding of rivalry in competing for neighbors' attention.

A3 Measuring migration with mobile phone data

We use the mobile phone logs to reconstruct the migration history of each individual in three steps.

First, we extract the timestamp and cell phone tower identifier corresponding to every phone call and text message made by each individual in the 4.5-year period. This creates a set of tuples {`subscriber_ID, timestamp, tower_ID`} for each subscriber. The tower identifier allows us to approximately resolve the location of the subscriber, to an area of roughly 100 square meters in urban areas and several square kilometers in rural areas. The physical locations of these towers are shown in Figure 3. We do not observe the location of subscribers in the time between phone calls and text messages.

Next, we assign each subscriber to a "home" district in each month that she makes one or more transactions. Our goal is to identify the location at which the individual spends the majority of her time, and specifically, the majority of her evening hours.[39] The full details of this assignment procedure are given in Algorithm 1. To summarize, we first assign all towers to a geographic district, of which there are 30 (we treat the three small districts that comprise the capital of Kigali as a single district). Then, for each individual, we compute the most frequently visited district in every hour of the entire dataset (i.e., there will be a maximum of 4.5 years * 365 days * 24 observations for each individual, though in practice most individuals appear in only a fraction of possible hours). We then aggregate these hourly observations, identifying the district where each individual spends the majority of hours of each night (between 6pm and 7am). Finally, we aggregate these daily observations by identifying the district in which the individual spent the majority of nights in each month. The end result

[39] 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.

is a panel of individual-month districts.[40] After this step, we have an unbalanced panel indicating the home location 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. As in Blumenstock (2012), we say that a migration occurs in month $t+1$ 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 $t+1$ is different from d; 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).[41] Complete details are given in Algorithm 2.

The total migration flows that we calculate from the phone data using $k = 2$ are broadly consistent with those captured in the 2012 Rwandan census (Figure A1). In Appendix Figure A13, we observe similar patterns when we define a migration event in the phone data using $k = 6$ (Figure A13a) and $k = 12$ (Figure A13b). As noted in Section 2.2, the measures do not align perfectly; inconsistencies likely arise from the non-representativity of phone owners and from differences in how the two instruments define migration. Since we lose a considerable amount of our data by increasing k (which requires that the individual be observed continuously for k months in one district and then for an additional k months in the same or a different district), our main analysis defines migration using $k = 2$.[42]

[40] 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.

[41] Individuals are treated as missing in month t if they are not assigned a home location in 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 $t+k$.

[42] For instance, as shown in Table A1, we lose roughly half of our sample by moving from $k = 2$ to $k = 6$.

A4 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;
→ Calculate *overall daily modal districts*;
→ Calculate *overall monthly modal districts*;
Step 2 calculate the *hourly modal districts*;

```
if tie districts exist 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 exist 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 exist then
    if overall monthly modal districts can resolve then
        return the district with larger occurance number;
    end
end
```

Algorithm 1: Home location assignment

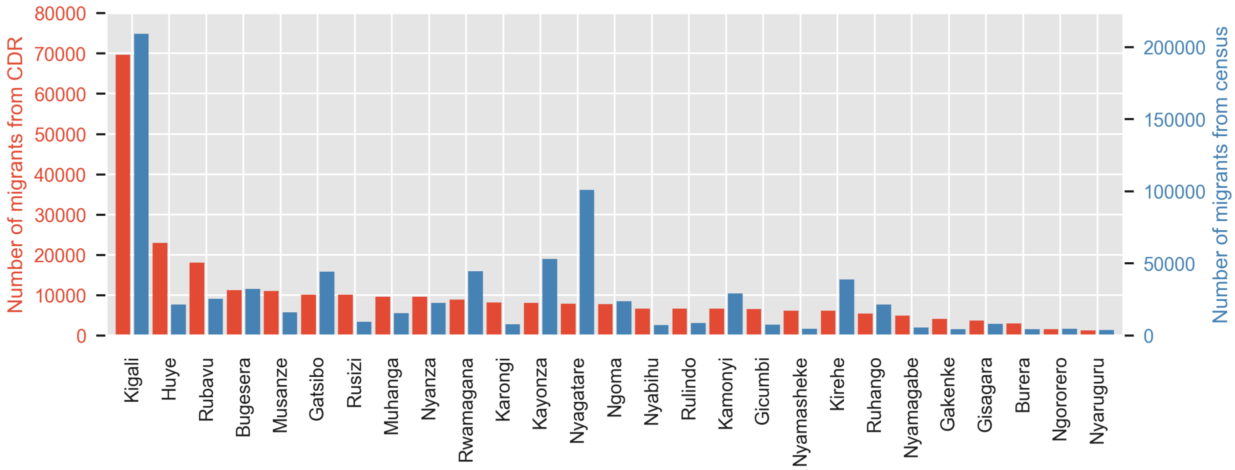
Data: Monthly modal district for four consecutive months: D_1, D_2, D_3, D_4
Result: Migration type

```
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
```

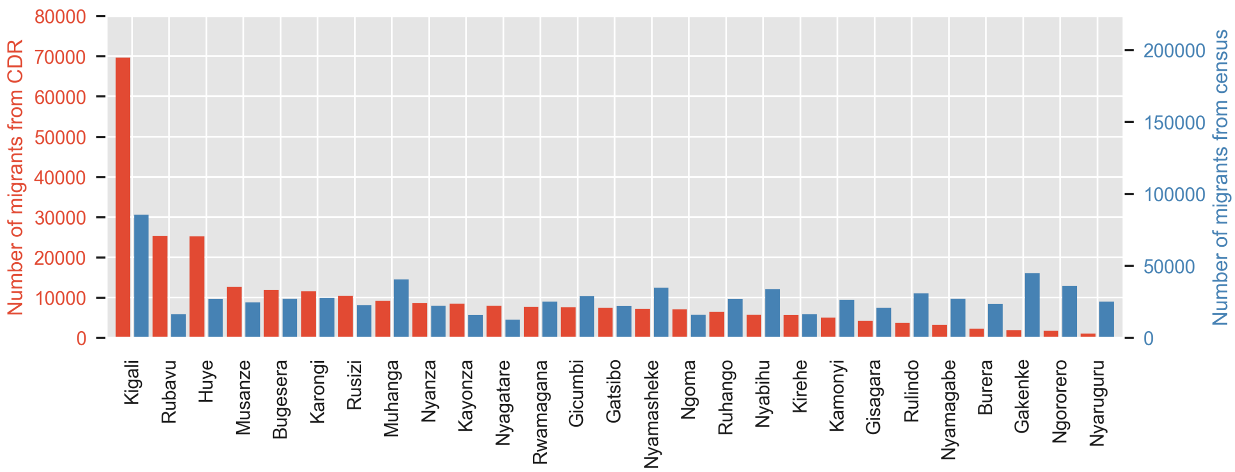
Algorithm 2: Classifying individuals by migrant type for k=2

A5 Appendix Figures and Tables

Figure A1: Validation of Migration Data



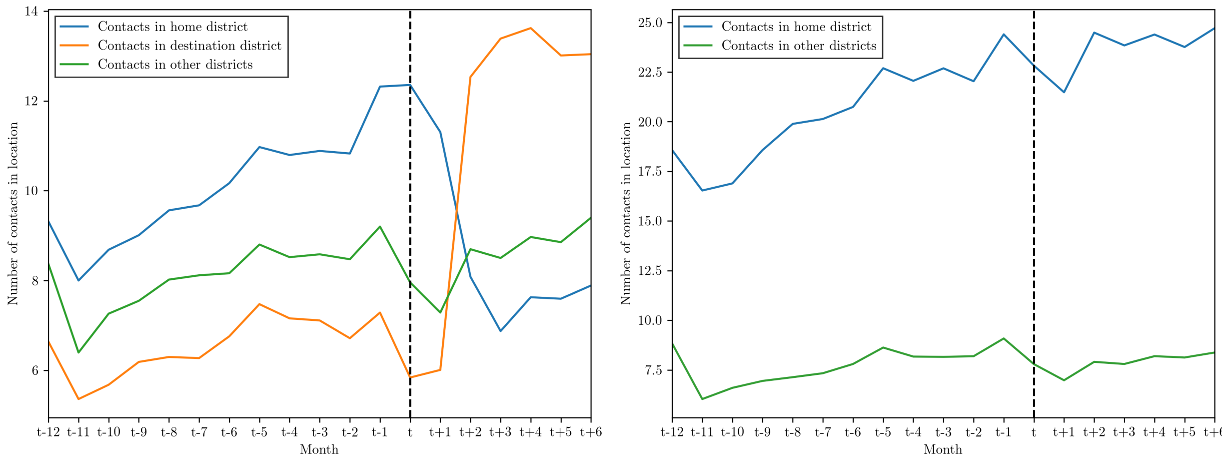
(a) Destination



(b) Origin

Notes: Figure shows the count of Rwandan migrants (a) to each district and (b) from each district, according to two independent sources of data. Red bars indicate the number of migrants inferred from the mobile phone data, using the methods described in Section 2.2. Blue bars indicate the number of recent internal migrants reported in the 2012 Rwandan census data (National Institute of Statistics of Rwanda, 2014). Districts ordered by the number of migrants inferred from phone data.

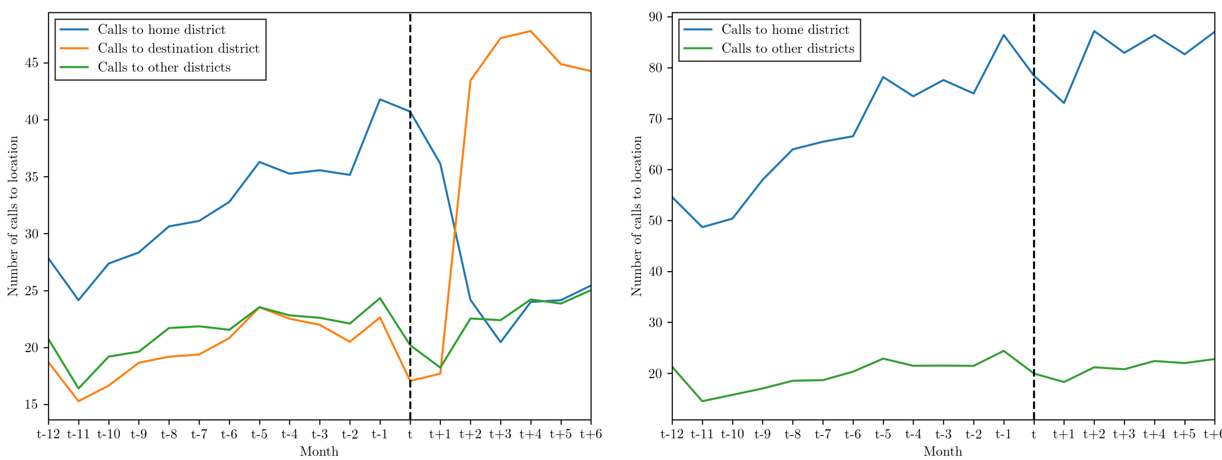
Figure A2: Changes in number of contacts over time



(a) Migrants (# Contacts)

(b) Non-Migrants (# Contacts)

Figure A3: Changes in number of calls over time



(a) Migrants (# Calls)

(b) Non-Migrants (# Calls)

Notes: Figures show how the average number of contacts (top figures) and average number of calls (bottom figures) change over time, averaged across migrants (left figures) and non-migrants (right figures). The top figure differs from Figure 4 by plotting the *number* rather than the *percent* of contacts by location. The dashed vertical line indicates the date of migration. For the right figures, we draw a random sample of 10,000 non-migrants by selecting, for each migrant who is sampled to appear in (a) and observed to migrate in month t, a non-migrant from the same reference month t.

Figure A4: Number of friends of friends, before and after migration (migrants)

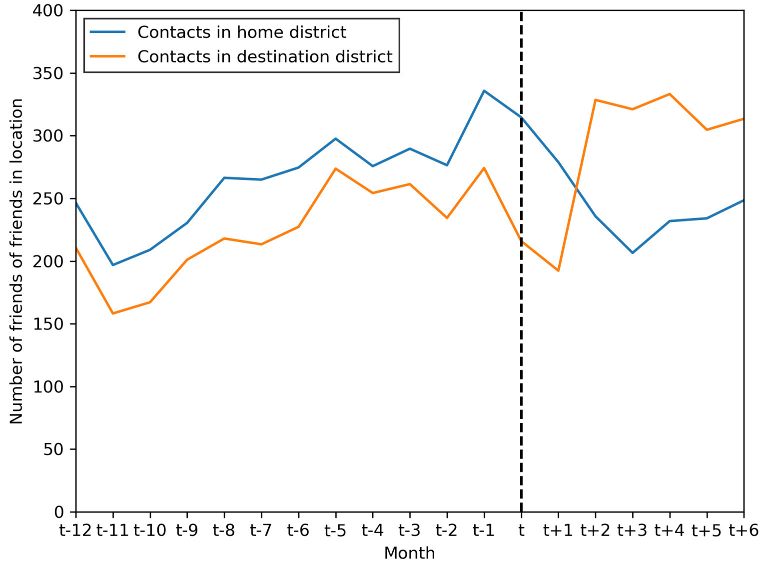
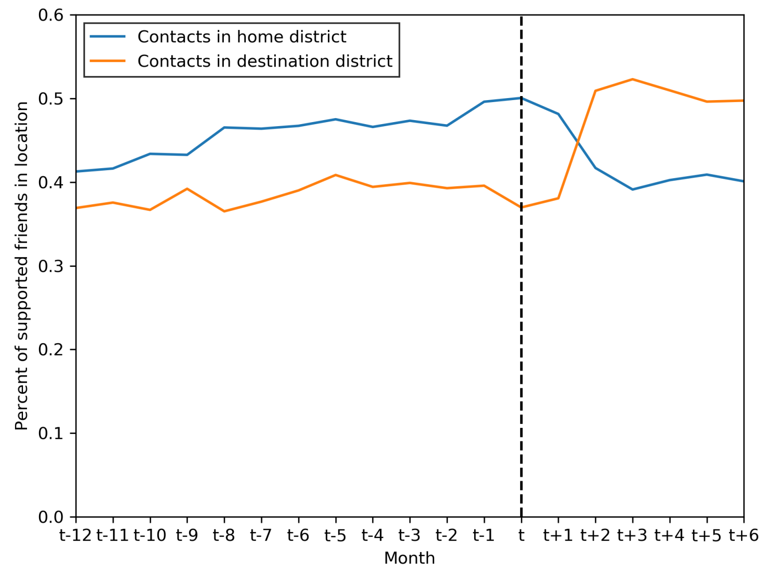
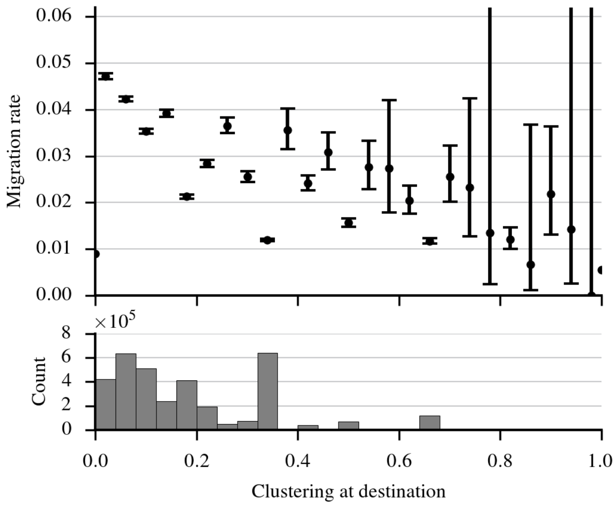


Figure A5: Percent of friends with common support, before and after migration (migrants)

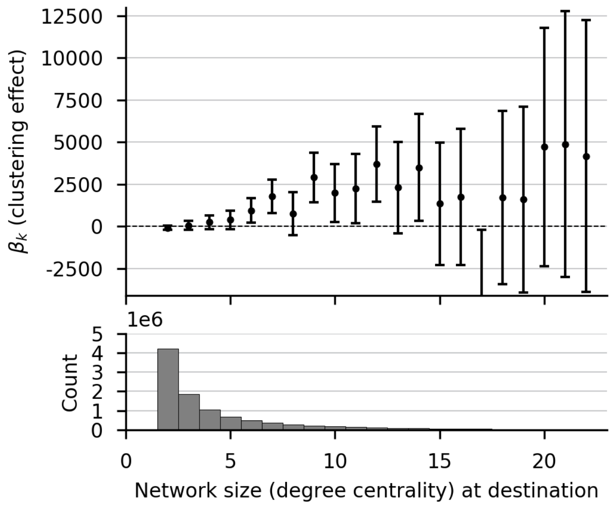


Notes: Top figure shows the average number of friends of friends migrants have in their home district and their destination district, in each of the 12 months before and 6 months after migration. Bottom figure shows the average percent of the migrants friends who have a common friend.

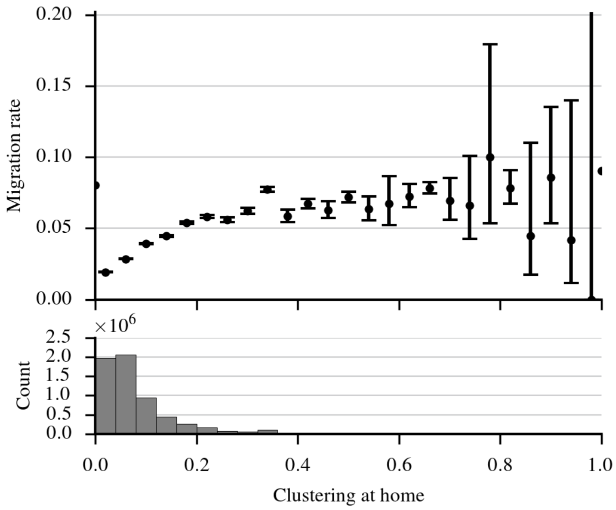
Figure A6: Relationship between migration rate and clustering



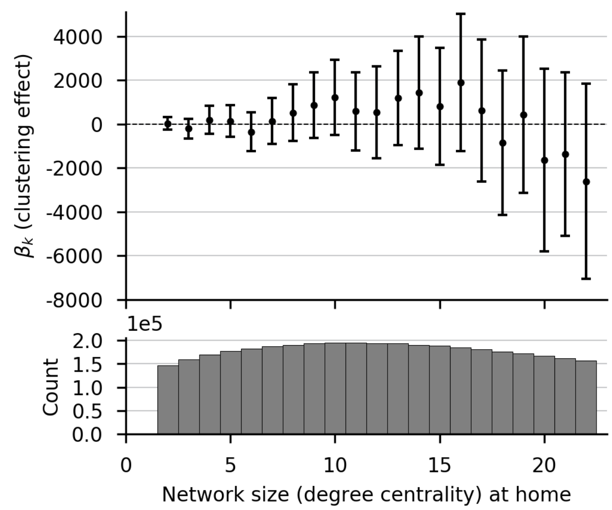
(a) Clustering at Destination



(b) Clustering at Destination, by Degree



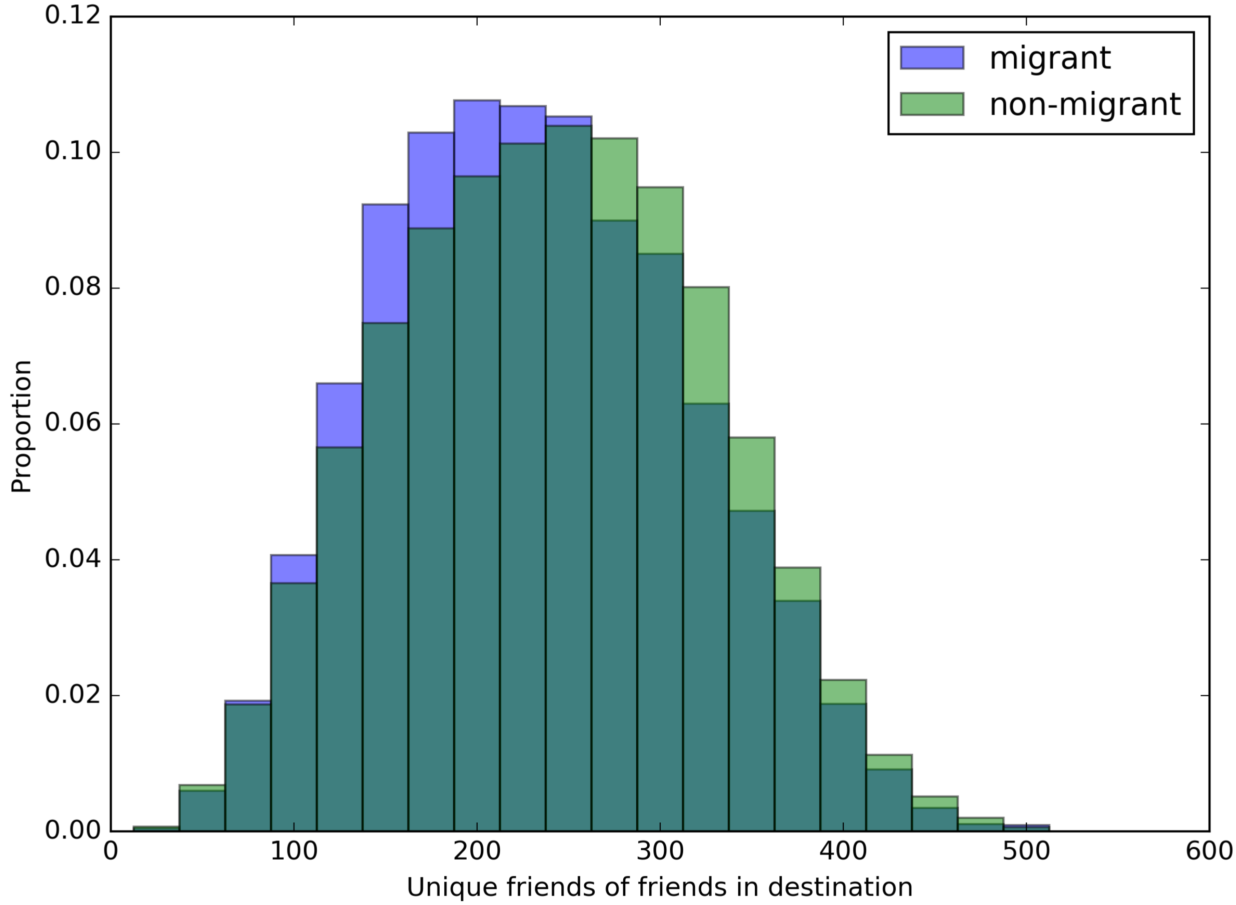
(c) Clustering at Home



(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 for a and c indicate 95% confidence intervals, using the Wilson Score interval for binomial proportions. Error bars for b and d indicate 95% confidence intervals, two-way clustered by individual and by home-destination-month. Coefficients and standard errors on b and d are multiplied by 1000 to make figures legible.

Figure A7: Migrants have fewer friends of friends than non-migrants



Notes: The figure focuses on all individuals who have exactly 10 unique contacts in a potential destination, and shows the distribution of the number of unique "friends of friends" in that destination. Counterintuitively, migrants have fewer unique friends of friends than non-migrants.

Figure A8: Number of friends of friends, before and after migration, shift-share approach

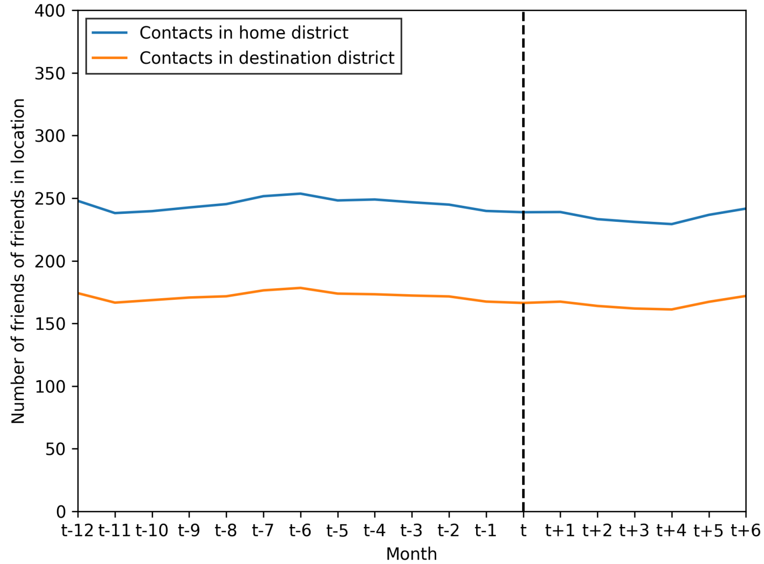
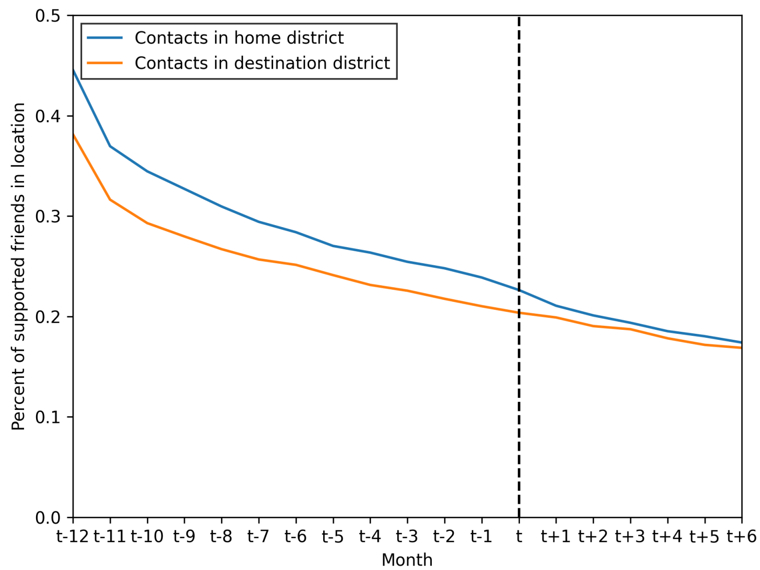
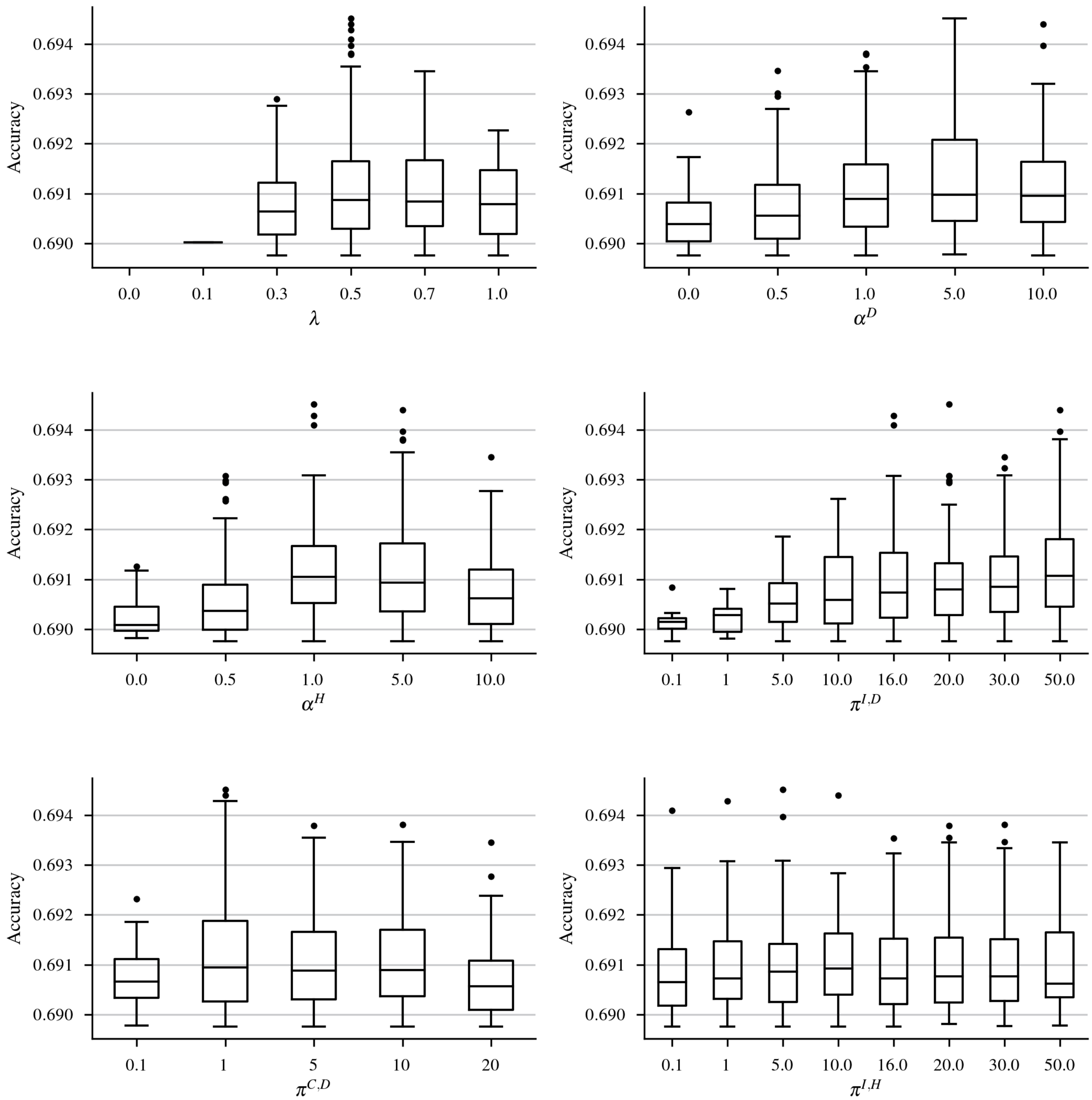


Figure A9: Percent of friends with common support, before and after migration, shift-share approach



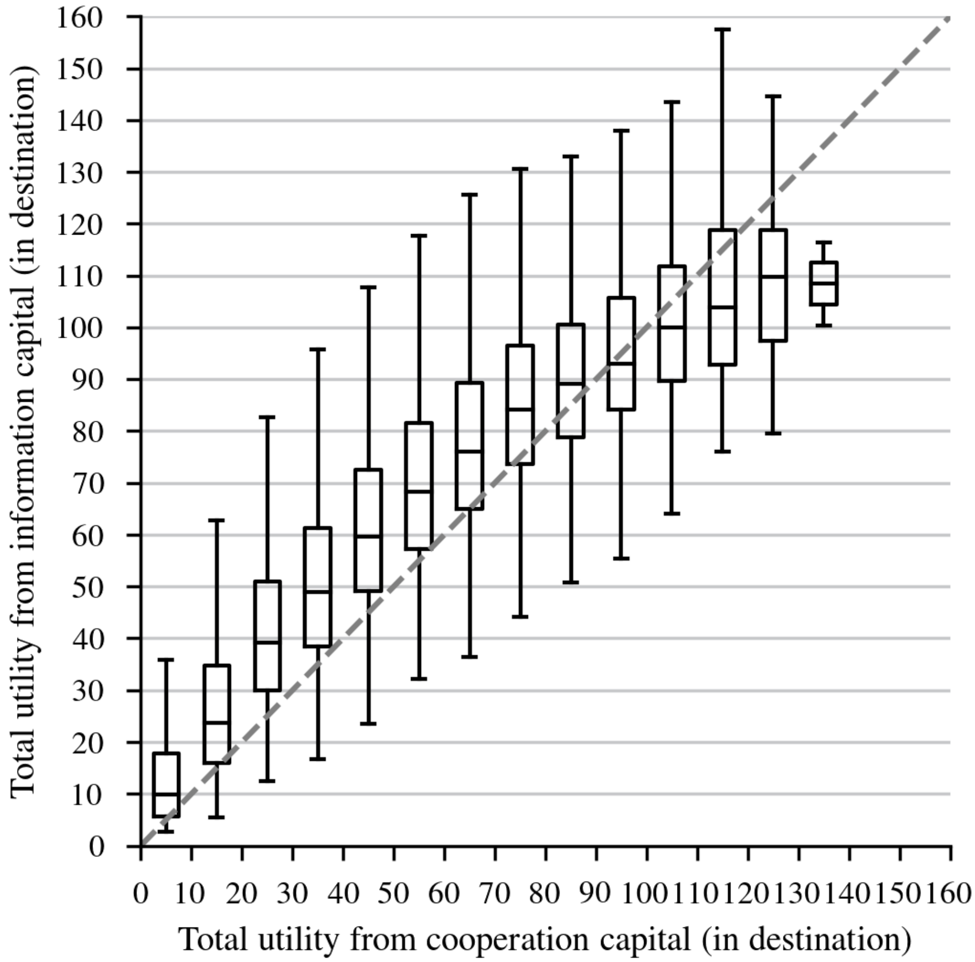
Notes: Top figure shows total number of friends of friends migrants have in their home district and their destination district, in each of the 12 months before and 6 months after migration. Bottom figure shows the percent of the migrants friends who have a common friend. Different from Figures A4 and A5, these two figures hold fixed the migrant's contacts at $t-12$, so that changes in friends of friends (top panel) and changes in common support (bottom panel) come only from changes in higher-order network structure, not from changes in the migrant's immediate contacts.

Figure A10: Calibration results: marginal plots



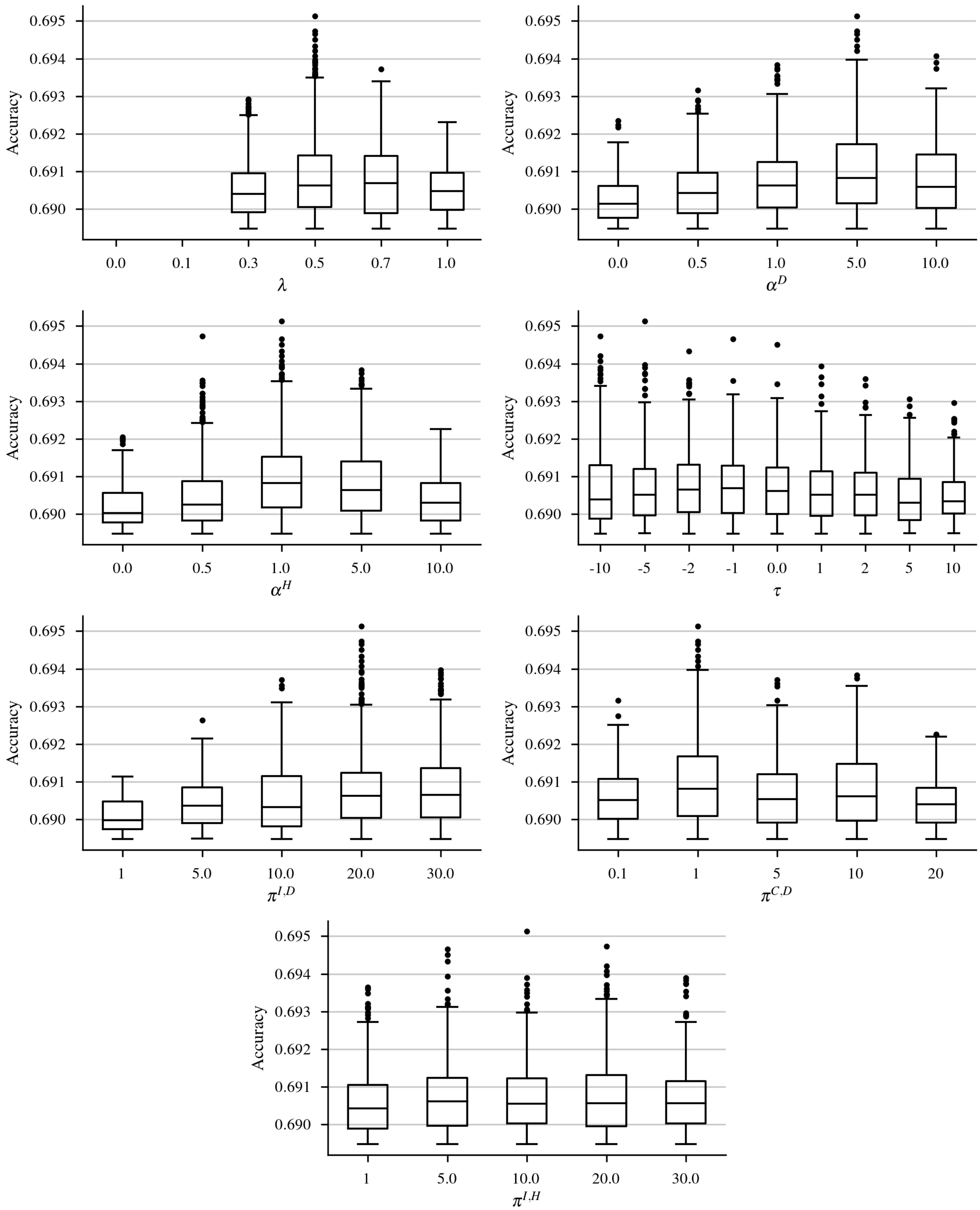
Notes: Figures show the marginal effect of varying λ, α_d, α^h and $(\pi^{I,d}, \pi^{C,d}, \pi^{I,h})$ when calibrating Model 17. Each of roughly 50,000 different parameter combinations is tested; the top percentile (by accuracy)of simulations are used to construct the box and whisker plots shown.

Figure A11: Calibration results: 'information' and 'cooperation' utility



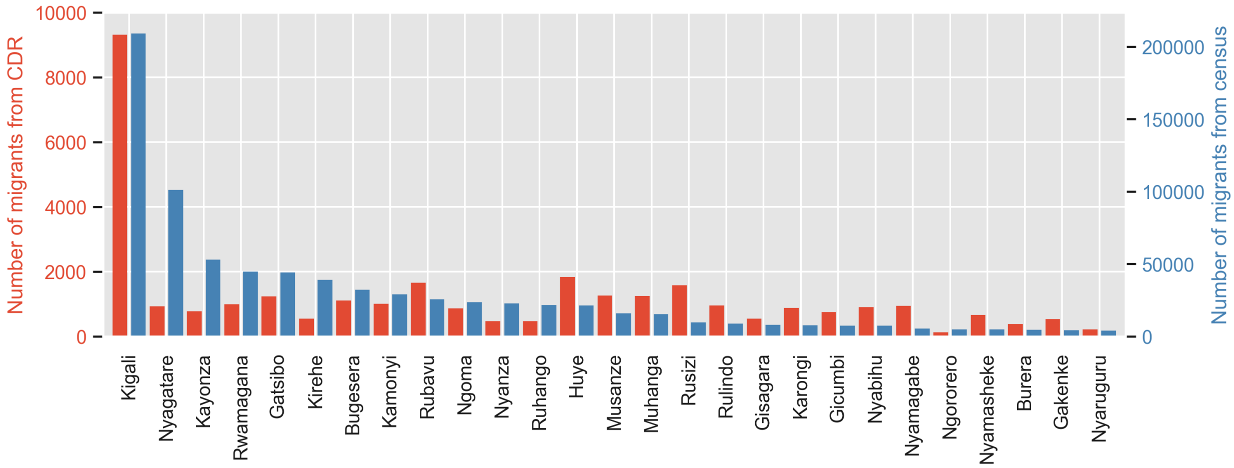
Notes: Figure shows the distribution of predicted utility from 'information' capital and 'cooperation' capital (i.e., equation 13) for 270,000 migrants and non-migrants.

Figure A12: Calibration results (with τ): marginal plots

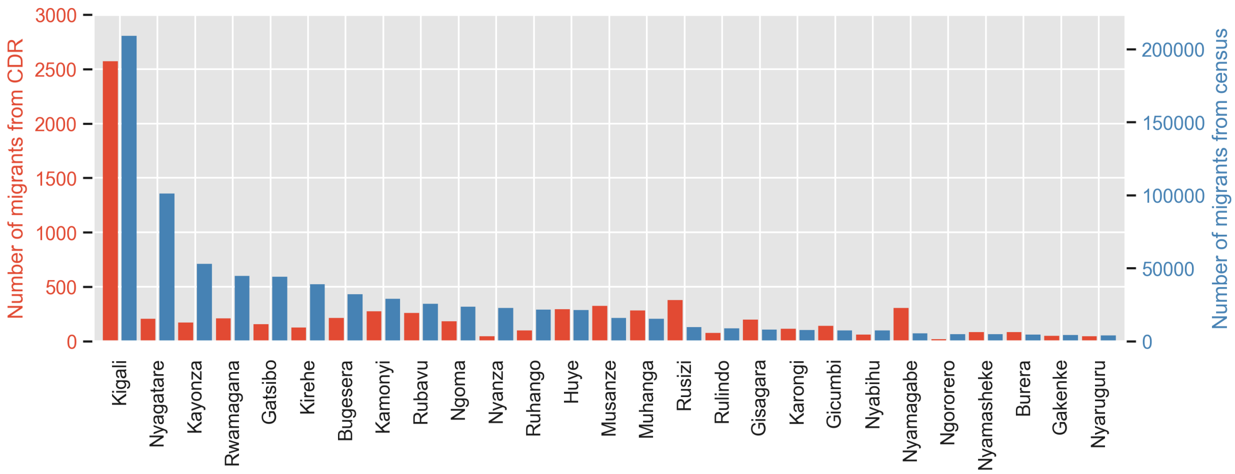


Notes: Figures show the marginal effect of varying λ, α, τ, and π when calibrating Model (24). Each of roughly 50,000 different parameter combinations is tested; the top percentile of simulations are used to generate this marginal plot.

Figure A13: Validation of Migration Data - Varying Definition of Migration



(a) 6-month definition



(b) 12-month definition

Notes: Figure shows the count of Rwandan migrants to each district, according to two independent sources of data. Red bars indicate the number of migrants inferred from the mobile phone data, using (a) a 6-month definition of migration and (b) a 12-month definition of migration (using the methods described in Section 2.2). Blue bars indicate the number of recent internal migrants reported in the 2012 Rwandan census data. Districts ordered by the number of migrants in the census.

Table A1: Detailed migration statistics derived from phone data, for different definitions of 'migration'

Definition of Migrant (k)	Total Individuals (N)	% Ever Migrate	% Repeat migrants (to same district)	% Repeat migrants (to any district)	% Long-distance migrants (non-adjacent districts)	% Circular Migrants
1	849,809	34.565	11.171	21.923	23.181	18.457
2	793,791	21.634	1.933	8.244	13.828	5.934
3	675,773	13.960	0.405	2.893	9.216	2.007
6	470,490	5.294	0.000	0.192	3.547	0.128

Notes: Table counts number of unique individuals meeting different definitions of a "migration", over the two-year period from July 2006 - June 2008. Each row of the table defines a migration by a different k, such that an individual is considered a migrant if she spends k consecutive months in a district d and then k consecutive months in a different district $d' \neq d$ – see Section 2.2 for details. The migration rate is defined as the ratio of the number of migrants over the number of migrants + non-migrants (where a non-migrant is observed to remain in the same district for $2k$ consecutive months). Repeat migrants are individuals who have migrated one or more times prior to a migration observed in month t. Long-distance migrants are migrants who travel between non-adjacent districts. Circular migrants are migrants who have migrated from d to h prior to being observed to migrated from h to d.

Table A2: Migration and destination network structure - Migrants only

	(1)	(2)	(3)
Degree (network size) in destination	638.67***	649.81***	
	(4.21)	(4.08)	
% Friends with common support in destination	1454.74***	1611.79***	371.23***
	(16.73)	(17.12)	(15.18)
Unique friends of friends in destination	-8.66***	-11.50***	-2.12***
	(0.13)	(0.13)	(0.10)
Observations (person-months)	6,766,396	6,766,396	6,766,396
Pseudo R^2	0.04	0.11	0.16
Degree fixed effects	No	No	Yes
Destination*Month fixed effects	No	Yes	Yes
Individual*Month fixed effects	Yes	Yes	Yes

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. Results are estimated using a conditional logit model, using social network characteristics of the destination network calculated in month $t-2$. This specification includes observations only for migrants in the month of migration (i.e., all possible destinations are considered in month t for individuals who migrate in t; individuals who do not migrate in month t are excluded from the regression). See discussion in Section A1.3. Standard errors are clustered by individual. *p<0.1; **p<0.05; ***p<0.01.

Table A3: Heterogeneity by Migration Frequency (Repeat and First-time)

Migration Frequency	(1) Any	(2) Repeat	(3) First-Time
% Destination support	182.84***	116.72*	173.61***
	(11.75)	(70.94)	(12.10)
Destination friends of friends	−0.14***	1.37***	−0.28***
	(0.06)	(0.24)	(0.07)
% Home support	105.62***	-53.53	109.84***
	(17.89)	(82.16)	(18.18)
Home friends of friends	1.38***	0.59***	1.52***
	(0.04)	(0.17)	(0.04)
Observations	184,637,637	7,045,450	184,430,618
R^2	0.68	0.09	0.69
Degree fixed effects	Yes	Yes	Yes
Home*Destination*Month fixed effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes

Notes: All specifications include degree fixed effects, (home * destination * month) fixed effects, and individual fixed effects. Repeat migrants are individuals who have migrated one or more times from h to d prior to a $h-d$ migration observed in month t. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

Table A4: Heterogeneity by Migration Duration (Long-term vs. Short-term)

Migration Distance	(1) Any	(2) Long Stay (> 12 months)	(3) Short Stay (< 6 months)
% Destination support	182.84***	287.19***	164.81***
	(11.75)	(23.58)	(15.39)
Destination friends of friends	−0.14***	0.86***	−0.62***
	(0.06)	(0.11)	(0.09)
% Home support	105.62***	71.39***	142.73***
	(17.89)	(32.34)	(24.12)
Home friends of friends	1.38***	1.65***	1.35***
	(0.04)	(0.08)	(0.05)
Observations	184,637,637	179,534,421	154,216,355
R^2	0.68	0.74	0.71
Degree fixed effects	Yes	Yes	Yes
Home*Destination*Month fixed effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes

Notes: All specifications include degree fixed effects, (home * destination * month) fixed effects, and individual fixed effects. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

Table A5: Heterogeneity by Distance (Adjacent districts vs. Non-adjacent districts)

	(1)	(2)	(3)
Migration Distance	Any	Short Distance (adjacent districts)	Long-Distance (non-adjacent districts)
% Destination support	182.84***	279.08***	278.75***
	(11.75)	(23.74)	(16.56)
Destination friends of friends	−0.14***	−0.145	−1.13***
	(0.06)	(0.12)	(0.10)
% Home support	105.62***	239.37***	−43.21*
	(17.89)	(28.19)	(22.93)
Home friends of friends	1.38***	2.17***	0.80***
	(0.04)	(0.06)	(0.05)
Observations	184,637,637	37,259,713	154,216,355
R^2	0.68	0.56	0.70
Degree fixed effects	Yes	Yes	Yes
Home*Destination*Month F.E.	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes

Notes: All specifications include degree fixed effects, (home * destination * month) fixed effects, and individual fixed effects. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

Table A6: The role of recent migrants and co-migrants

	(1)	(2)	(3)
Destination Degree (network size)	315.86***	295.23***	291.00***
	(2.53)	(2.62)	(2.59)
Destination % friends with support	2270.11***	2281.45***	2278.57***
	(12.03)	(11.99)	(11.98)
Destination friends of friends	−5.03***	−4.61***	−4.52***
	(0.085)	(0.09)	(0.09)
Home Degree (network size)	81.34***	−471.41***	−386.20***
	(1.27)	(3.01)	(2.76)
Home % friends with support	815.48***	657.52***	640.16***
	(17.38)	(17.30)	(17.27)
Home friends of friends	0.52***	−2.04***	−2.19***
	(0.05)	(0.05)	(0.05)
Contacts who recently migrated		665.29***	590.97***
		(3.44)	(3.22)
Observations	184,637,637	184,637,637	184,637,637
R^2	0.68	0.68	0.68
Degree fixed effects	No	No	No
Home*Destination*Month fixed effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes
Definition of "Recent"	NA	Ever	Last month

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. Column (1) replicates the original result from Table 2; column (2) controls for the number of migrants that i knows, who ever migrated from h to d prior to t; column (3) controls for the number of recent migrants that i knows, who migrated from h to d in the month prior to t. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

Table A7: Migration and networks, controlling for prior visits to the destination

	(1)	(2)	(3)
Destination Degree (network size)	315.86***	218.29***	218.74***
	(2.53)	(2.10)	(2.13)
Destination % friends with support	2270.11***	1266.60***	1409.82***
	(12.03)	(12.58)	(13.20)
Destination friends of friends	−5.03***	−3.50***	−3.90***
	(0.085)	(0.07)	(0.07)
Home Degree (network size)	81.34***	82.35***	79.80***
	(1.27)	(1.24)	(1.23)
Home % friends with support	815.48***	866.92***	772.09***
	(17.38)	(17.38)	(17.30)
Home friends of friends	0.52***	0.73***	0.60***
	(0.05)	(0.04)	(0.04)
Prior visit to destination		2717.40***	
		(8.18)	
Prior visit to destination (evening hours)			2654.68***
			(8.75)
Observations	184,637,637	184,637,637	184,637,637
R^2	0.68	0.68	0.68
Degree fixed effects	No	No	No
Home*Destination*Month fixed effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. Column (1) replicates the original result from Table 2; column (2) controls for whether i ever used their phone in d in the month prior to t; column (3) controls for whether i ever used their phone in d during the evening hours (i.e., between 6pm and 7am, as a proxy for staying overnight) in the month prior to t. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

Table A8: The role of strong ties and weak ties

	(1)	(2)	(3)	(4)
Destination "Weak tie"	227.25***	224.05***	257.36***	213.94***
	(4.90)	(4.90)	(10.42)	(9.29)
Destination "Strong tie"	252.83***	340.04***	247.91***	322.66***
	(1.37)	(2.59)	(1.12)	(2.56)
% Destination "Support"		2294.91***		2278.33***
		(12.00)		(12.02)
Destination friends of friends		-5.17^{***}		-5.07^{***}
		(0.08)		(0.08)
Home "Weak tie"	97.22***	75.66***	99.77***	77.16***
	(0.77)	(1.40)	(0.61)	(1.31)
Home "Strong tie"	140.01***	100.72***	197.47***	141.80***
	(2.33)	(2.52)	(4.92)	(4.88)
% Home "Support"		798.61***		796.67***
		(17.46)		(17.45)
Home friends of friends		0.55***		0.53***
		(0.05)		(0.05)
Observations	184,637,637	184,637,637	184,637,637	184,637,637
R^2	0.68	0.68	0.68	0.68
Degree fixed effects	No	No	No	No
Home*Destination*Month FE's	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes
Definition of "Strong"	90th Percentile	90th Percentile	95th Percentile	95th Percentile

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. This table disaggregates contacts at home and destination by the strength of the relationship, where strength is defined in terms of the number of phone calls observed between the two parties. Columns 1 and 2 consider strong ties to be relationships with 5 or more phone calls (the 90th percentile of tie strength); columns 3 and 4use a threshold of 12 calls (the 95th percentile of tie strength). Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

Table A9: Disaggregating the friend of friend effect by the strength of the 2nd-degree tie

	(1)	(2)	(3)	(4)	(5)	(6)
Destination friends of friends (all)	6.92***					
	(0.04)					
Destination friends of friends (strong-strong)		132.27***				8.21***
		(0.81)				(2.17)
Destination friends of friends (strong-weak)			30.40***			6.14***
			(0.18)			(0.47)
Destination friends of friends (weak-strong)				-0.49		-1.26
				(0.36)		(0.75)
Destination friends of friends (weak-weak)					9.19***	5.34***
					(0.05)	(0.14)
Home friends of friends (all)	3.58***					
	(0.02)					
Home friends of friends (strong-strong)		67.82***				11.11***
		(0.50)				(1.10)
Home friends of friends (strong-weak)			14.72***			2.44***
			(0.10)			(0.24)
Home friends of friends (weak-strong)				26.42***		19.62***
				(0.15)		(0.42)
Home friends of friends (weak-weak)					4.92***	0.53***
					(0.03)	(0.08)
Observations	184,637,637	184,637,637	184,637,637	184,637,637	184,637,637	184,637,637
R^2	0.68	0.67	0.67	0.68	0.68	0.68

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. We show the destination "friend of friend" coefficient separately for geometries of different tie strength. "Strong-strong" (column 2) indicates the effect of friends of friends when the potential migrant i is connected to j via a strong tie, and j is connected to k via a strong tie. "Strong-weak" (column 3) indicates the effect when i and j have a strong tie and j and k have a weak tie. Columns 4 and 5 follow this nomenclature. Strong ties are defined as relationships with 5 or more phone calls (the 90th percentile of tie strength) in a given month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01

Table A10: Disaggregating the network support effect by the strength of supported ties

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Destination support (all)	136.47***									
	(11.91)									
Destination support (sss)		184.22***								196.73***
		(49.55)								(49.90)
Destination support (sws)			367.25***							370.71***
			(46.01)							(46.15)
Destination support (ssw)				25.94**						−5.12
				(9.93)						(10.96)
Destination support (sww)					106.63***					104.71***
					(6.87)					(7.42)
Destination support (wss)						−137.20**				−160.66***
						(37.98)				(40.48)
Destination support (wws)							219.32***			51.98
							(32.76)			(35.30)
Destination support (wsw)								−389.10***		−399.92***
								(44.02)		(44.20)
Destination support (www)									219.86***	195.44***
									(33.16)	(33.20)
Destination strong tie	60.62***	63.03***	59.21***	65.15***	65.39***	66.38***	62.16***	64.87***	65.66***	57.62***
	(3.53)	(3.55)	(3.57)	(3.50)	(3.51)	(3.52)	(3.54)	(3.50)	(3.50)	(3.69)
Observations	184,637,637	184,637,637	184,637,637	184,637,637	184,637,637	184,637,637	184,637,637	184,637,637	184,637,637	184,637,637
R^2	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. We show the Destination network "support" coefficient separately for geometries of different tie strengths. "SSS' (column 2) indicates the effect of network support for triangles where the potential migrant i is connected to j via a strong tie, j is connected to k via a strong tie, and k and i are connected by a strong tie. "SWS" (column 3) indicates the effect when i and j have a strong tie, j and k have a weak tie, and k and i have a strong tie. Columns 4-8 follow a similar nomenclature. Strong ties are defined as relationships with 5 or more phone calls (the 90th percentile of tie strength) in a given month. All specifications also include the corresponding home support characteristics; the coefficients associated with home networks are omitted from the table to enable display on a single page. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01

Table A11: Beyond location-specific subnetworks

	(1)	(2)	(3)
Panel A: Migrant connected to friend of friend in destination via friend in destination:			
Destination % friends with common support	182.84*		253.43***
	(0.1381)		(13.48)
Destination unique friends of friends	−0.14***		−0.08
	(0.0035)		(0.07)
Home % friends with common support	105.62***		352.83***
	(17.89)		(27.10)
Home unique friends of friends	1.38***		2.79***
	(0.04)		(0.10)
Panel B: Migrant connected to friend of friend in destination via friend at home:			
Destination % friends with common support		−85.74***	−268.27***
		(18.42)	(21.05)
Destination unique friends of friends		−1.53***	−1.25***
		(0.08)	(0.08)
Home % friends with common support		87.49***	
		(17.92)	
Home unique friends of friends		1.53***	
		(0.04)	
Observations	184,637,637	184,637,637	184,637,637
R^2	0.68	0.68	0.68
Destination degree fixed effects	Yes	Yes	Yes
Home degree fixed effects	Yes	Yes	Yes
Home*Destination*Month fixed effects	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. Panels vary the location of the intermediate connection j between the migrant i and the friend-of-friend k in the destination. Column (1) reproduces column 3 of Table 2, Panel A. Column 2 calculates 'friends of friends' and 'friends with support' when j is at home. Column 3 shows the joint regression with all four regressors. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make tables more readable. *p<0.1; **p<0.05; ***p<0.01.

Table A12: Jointly estimated effects (6 month network lag)

	(1)	(2)	(3)
Destination Degree (network size)	274.53***	302.00***	
	(2.48)	(2.36)	
Destination % friends with support	2594.77***	2281.92***	306.28***
	(13.52)	(13.77)	(14.09)
Destination friends of friends	-6.47***	-5.03***	-0.43***
	(0.08)	(0.09)	(0.08)
Home Degree (network size)	14.39***	61.92***	
	(1.52)	(1.67)	
Home % friends with support	905.79***	939.91***	290.46***
	(19.46)	(19.70)	(20.53)
Home friends of friends	4.72***	2.16***	2.65***
	(0.06)	(0.07)	(0.06)
Home district	8433.02***		
	(9.48)		
Observations	145,546,740	145,546,740	145,546,740
Pseudo R^2	0.68	0.68	0.68
Degree fixed effects	No	No	Yes
Home*Destination*Month fixed effects	No	Yes	Yes
Individual fixed effects	No	Yes	Yes

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. Social network characteristics calculated in month $t-6$. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

Table A13: "Shift share" regression results

	(1)	(2)
Destination change in % friends with common support	46.44*	9.72***
	(26.50)	(37.35)
Destination change in friends of friends	−2.08***	−0.41
	(0.32)	(0.38)
Home change in % friends with common support	38.91*	−28.50
	(22.75)	(31.21)
Home change in friends of friends	2.42***	4.05***
	(0.27)	(0.35)
Observations	144,287,946	101,758,194
R^2	0.68	0.69
Partial R^2 (from network changes)	0.02	0.02
"Early period" t_0	6 months	12 months
"Late period" t_1	2 months	2 months
Degree fixed effects	Yes	Yes
Home*Destination*Month fixed effects	Yes	Yes
Individual fixed effects	Yes	Yes

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. Social network characteristics are calculated based on changes in higher-order network structure between an early period t_0 and a later period t_1, holding fixed the set of contacts who are connected to i in both t_0 and t_1. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

Table A14: Robustness to alternative fixed effect specifications

	(1)	(2)	(3)	(4)	(5)
Destination friends of friends	-0.0002	0.0011	-0.0064^{**}	-0.0077^{***}	-0.0028
	(0.0035)	(0.0040)	(0.0030)	(0.0025)	(0.0044)
% Destination friends with support	1.4808^{***}	1.3719^{***}	0.3458^{***}	0.6663^{***}	0.1123
	(0.1435)	(0.2125)	(0.1339)	(0.1241)	(0.1365)
Observations	9,889,981	9,889,981	9,889,981	9,889,981	9,889,981
R^2	0.1853	0.5081	0.5952	0.6681	0.6333
Fixed effects	$D, h*d*t, i$	$D, h*d*t, i*t$	$D, h*d*t, i*d$	$D, h*d*t, i*D$	$D, h*d*i, t$

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. All specifications control non-parametrically for the number of unique contacts D that i has in district d. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.