DISCUSSION PAPER SERIES

DP13611

MIGRATION AND THE VALUE OF SOCIAL NETWORKS

Joshua Blumenstock, Guanghua Chi and Xu Tan

DEVELOPMENT ECONOMICS AND LABOUR ECONOMICS

MIGRATION AND THE VALUE OF SOCIAL NETWORKS

Joshua Blumenstock, Guanghua Chi and Xu Tan

This Discussion Paper is issued under the auspices of the Centre's research programme in **DEVELOPMENT ECONOMICS AND LABOUR ECONOMICS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

MIGRATION AND THE VALUE OF SOCIAL NETWORKS

Abstract

What is the value of a social network? 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 use a rich 'digital trace' dataset to link the migration decisions of millions of individuals to the topological structure of their social networks. We find that migrants systematically prefer 'interconnected' networks (where friends have common friends) to 'expansive' networks (where friends are well connected). A micro-founded model of network-based social capital helps explain this preference: migrants derive more utility from networks that are structured to facilitate social support than from networks that efficiently transmit information.

Joshua Blumenstock - jblumenstock@gmail.com
University of California, Berkeley and CEPR

Guanghua Chi - guanghua@berkeley.edu
University of California, Berkeley

Xu Tan - tanxu@uw.edu
University of Washington

Migration and the Value of Social Networks[*]

Joshua Blumenstock[†]   Guanghua Chi[‡]   Xu Tan[§]

U.C. Berkeley   U.C. Berkeley   University of Washington

March 19, 2019

Abstract

What is the value of a social network? 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 use a rich 'digital trace' dataset to link the migration decisions of millions of individuals to the topological structure of their social networks. We find that migrants systematically prefer 'interconnected' networks (where friends have common friends) to 'expansive' networks (where friends are well connected). A micro-founded model of network-based social capital helps explain this preference: migrants derive more utility from networks that are structured to facilitate social support than from networks that efficiently transmit information.

JEL classification: 015, R23, D85, Z13, O12, C55

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

[†]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.

The central goal of this paper is to better understand exactly *how* social networks influence an individual's decision to migrate, and through the analysis of migration, to provide more general insight into how social networks provide utility. 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); and second, that networks act as a safety net for migrants by providing material or social support (Carrington, Detragiache and Vishwanath, 1996, Dolfin and Genicot, 2010, Munshi, 2014, Comola and Mendola, 2015). This distinction between the 'information' and 'social support' value of social networks made in migration literature parallels the contrast between *information capital* and *cooperation capital* made in the theoretical network literature (Jackson, 2018). More broadly, network theory suggests that the utility an individual receives from a social network depends, in part, on the topological structure of the network. Information capital, which reflects the network's ability to efficiently transmit information, is associated with *expansive* subnetworks (e.g., stars and trees) where an individual is linked to many others via short network paths.[1] Cooperation capital is usually motivated by repeated game models of network interaction, where *interconnected* networks (e.g., cliques) best support social reinforcement and sanctioning.[2]

However, there is considerable ambiguity about which types of social capital matter most, and even the nature of each type of social capital in isolation. For instance, the prevailing view in the migration literature is that migrants tend to go to places where they have larger networks,[3] but several studies argue that larger networks may actually deter migration, for

---

[1]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]Jackson, Rodriguez-Barraquer and Tan (2012) and Ali and Miller (2016) provide recent examples. See also Ligon and Schechter (2011), Jackson, Rodriguez-Barraquer and Tan (2012), Ambrus, Mobius and Szeidl (2015) and Chandrasekhar, Kinnan and Larreguy (2018).

[3]Classic papers documenting this effect include Rees (1966), Greenwood (1969), Granovetter (1973), Montgomery (1991), and Borjas, Bronars and Trejo (1992). More recent examples include Munshi (2003), Winters, de Janvry and Sadoulet (2001), Dolfin and Genicot (2010), Patel and Vella (2012), Fafchamps and Shilpi (2013), Mahajan and Yang (2017), Giulietti, Wahba and Zenou (2018), Bertoli and Ruyssen (2018).

instance if migrants compete with one another over opportunities and resources (Calvó-Armengol, 2004, Calvó-Armengol and Jackson, 2004, Beaman, 2012). 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).

These ambiguities arise because it has historically been difficult to differentiate between distinct sources of social capital in a single empirical setting. In the migration case, linking social network structure to migration decisions is not feasible with traditional data. As Chuang and Schechter (2015) note, "there is little evidence making use of explicit network data on the impact of networks on the initial migration decision... Collecting migration data is quite difficult, and collecting network data is quite difficult; combining the two is even more so" (p.464).[4] 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.[5] Such proxies provide a reasonable approximation of the size of a migrant's social network, but obscure the higher-order topological network properties that can help disambiguate the mechanism through which social networks provide utility. This higher-order network structure plays a critical role in decisions about employment, education, health, finance, product adoption, and the formation of strategic alliances.[6] 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

---

[4]The difficulty of measuring migration is exacerbated in developing countries, where short-term migration is common (Deshingkar and Grimm, 2005, McKenzie and Sasin, 2007, Carletto, de Brauw and Banerjee, 2012, Lucas, 2015). The challenges of measuring social network structure are discussed in Chuang and Schechter (2015) and Breza et al. (2017).

[5]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, Wang and Wang (2018) take advantage of a resettlement program in China that sent 18 million urban youth to rural areas. Related approaches are used by Card (2001), Hanson and Woodruff (2003) and Dinkelman and Mariotti (2016).

[6]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, Mobius and Szeidl (2015) and Chandrasekhar, Kinnan and Larreguy (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.

Figure 1: Schematic diagrams of the social networks of three migrants
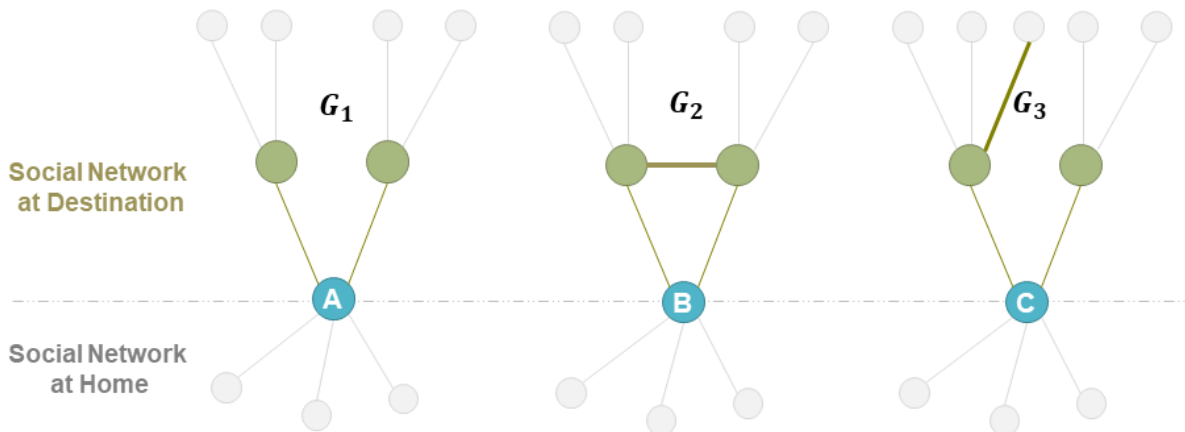
Notes: Each of the blue circles (A, B, C) represents a different individual considering migrating from their home to a new destination. Each individual has exactly three contacts in the home district (grey circles below the dashed line) and two contacts in the destination district (green circles above the dashed line). The social network of these three individuals is denoted by G_1, G_2, and G_3.

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.[7]

We begin with a reduced form analysis that links each individual's migration decision to the structure of his or her social network in the months prior to migration. 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 (identification is discussed in detail below). 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.

The reduced form analysis establishes a new set of stylized facts about the relationship between migration and social networks. First, we 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 are expected, an advantage of our setting is that we can observe the nonparametric relationship between migration and

---

[7]Limitations of these data are discussed in Section 3. Identification and estimation are the focus of Section 4.

network size. We find this relationship to be monotonic and approximately linear with elasticity one, 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). We also find that the probability of leaving home decreases proportional to the size of the home network.

Second, we document, to our knowledge for the first time, the role that higher-order network structure plays in migration decisions. As a proxy for the 'interconnectedness' of the network, we measure the extent to which the individual's local subnetwork is clustered, where a large proportion of neighbors have common friends. As a proxy for the network's 'expansiveness', we measure the size of the individual's distance-2 and distance-3 neighborhood. We find that migrants are drawn to locations where their networks are interconnected, but that, on average, they are actually *less* likely to go to places where their networks are expansive — 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 expansive networks is driven by settings where a migrant's direct contacts have a large number of "strong ties" in the destination (where tie strength is defined by the frequency of communication); when a migrant's destination contacts have many weak ties, 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 competitition for attention (as in Dunbar (1998) and 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), long-term migrants, and short-distance migrants — all 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 expansive.

Building on these reduced-form estimates, our final set of results provide structural insight into the more general question of how people derive value from their social networks.

This structure allows us to be more precise about the utility that comes from 'expansive' and 'interconnected' subnetworks, and accounts for more complex network structure than the proxy measures used in the reduced-form analysis. Our model characterizes the migration decision as, ceteris paribus, a tradeoff between the utility an individual receives from the home network and the utility received from a potential destination network, net an idiosyncratic cost of migrating. The focus of the model is on understanding the utility $u_i(G)$ an individual i receives from an arbitrary social network G. We assume that agents derive utility from their networks in two archetypal ways. First, as a source of information capital, where information transmission is modeled as a diffusion process with possible loss of information, as in Banerjee et al. (2013). And second, as a source of cooperation capital, where agents engage in repeated cooperation games with their neighbors, as in Jackson, Rodriguez-Barraquer and Tan (2012) and Ali and Miller (2016).

We estimate this model by maximizing the likelihood of hundreds of thousands of observed migration decisions, and note several results. First, in a departure from benchmark models of diffusion, we find strong support for competition or rivalry in information transmission: a model where information passes from i to j (inversely) proportional to the size of each individual's immediate network fits the data better than standard models where information passes with constant probability. In particular, our results suggest that two people share information with probability roughly inversely proportional to the square root of the (product of the) numbers of their contacts. Our model also allows us to decompose the total utility of an agent's network into two components. Consistent with the reduced-form regressions, we find that when information transmission is constrained to be non-rival, most agents receive very little utility from information capital (provided by structures that efficiently diffuse information) relative to cooperation capital (derived from network structures that facilitate repeated cooperation). However, when rivalry is empirically parameterized, information capital and cooperation capital contribute relatively evenly to the migrant's total utility.

Since our approach to studying migration with mobile phone data is new, we devote considerable attention to causal identification, and perform a large number of tests to check the robustness of our results.[8] Perhaps the most important limitation of our approach is that we lack exogenous variation in the structure of an individual's network, so that the social networks we observe are almoendogenous to migration decisions. We address this concern

---

[8]Our baseline results assume each individual faces an independent migration decision in each month. She can either stay put, or migrate to one of the 26 other districts in the country of Rwanda. We regress the binary migration decision on (lagged) properties of the migrant's social network, using either a discrete choice (multinomial logit) model or a panel fixed effects specification. Our measurement strategy, these specifications, and the robustness tests are described in detail in Sections 3 and 4.

in two principal ways. First, we relate migration decisions in each month to the structure of the social network several months prior in order to minimize the likelihood that the decision to migrate shaped the social network, rather than vice versa.[9] Second, and more important, identification is achieved through an extremely restrictive set of fixed effects that limit the potential for many of the most common sources of endogeneity. 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 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 (the G_2 vs. G_1 comparison of Figure 1) — 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).[10]

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.[11] Our data make it possible to establish a more nuanced set of stylized facts — highlighting, in particular, the value migrants place on interconnected networks,

---

[9]One concern is that migrants might begin to strategically reshape their networks long in advance of migrating. We perform several tests to check for such an effect, but find no evidence of anticipatory changes in network structure — see Section 4 for an extensive discussion.

[10]In addition to the preferred specification, we perform a series of robustness tests to more precisely isolate the source of identifying variation. In particular, we show the results from regressions that include fixed effects for (a) each individual-month, which isolates the variation between a migrant's potential destinations in a single month; (b) each individual-destination, which isolates variation over time in the structure of an individuals network in a single destination; (c) each individual j in the destination, which removes variation that might be driven by specific destination contacts who are singularly capable of facilitating migration. In these and related cases, the main results are qualitatively unchanged.

[11]Prominent examples include Munshi (2003), McKenzie and Rapoport (2010), Dolfin and Genicot (2010), Beaman (2012), Patel and Vella (2012), Bertoli, Fernández-Huertas Moraga and Ortega (2013), Ambrus, Mobius and Szeidl (2015), Morten (2019), Munshi and Rosenzweig (2016), Bertoli and Ruyssen (2018).

and substantial heterogeneity in how different types of migrants value networks differently — that have not been documented in prior work. Second, through the study of migration, we shed light on the more fundamental question of how individuals can derive utility from social networks (cf. Jackson, 2010, Banerjee et al., 2013, 2014). Specifically, we use millions of revealed-preference migration decisions to estimate a model of network utility. This allows us to distinguish between the utility provided by network geometries that facilitate the free flow of information from geometries that facilitate repeated cooperation. While the models we test are highly stylized, we hope it can provide a foundation for future work calibrating structural models of network utility with population-scale social network data.

2   A model of social capital and migration

A central goal of network theory is to understand how the structure of a social network affects the utility that an agent obtains from that network. Our model links social network structure (in both the home and destination) to subsequent migration decisions, to obtain a revealed preference measure of network utility.

Formally, we say that an individual i receives utility $u_i(G)$ 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.

$$u_i(G^d) > u_i(G^h) + \varepsilon_i. \tag{1}$$

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 (2018) 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*.

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. This notion is motivated by a robust theoretical and empirical literature that suggests that the value of a social network stems, at least in part, from its ability to efficiently transmit information (Topa,

2001, Calvó-Armengol and Jackson, 2004, Jackson and Yariv, 2010, Banerjee et al., 2013). Separately, we consider the cooperation capital of a network to be the network's ability to facilitate interactions that benefit from cooperation and community enforcement, such as risk sharing and social insurance (e.g., Jackson, Rodriguez-Barraquer and Tan, 2012, Ligon and Schechter, 2011, Chandrasekhar, Kinnan and Larreguy, 2018).

We therefore make the assumption that, for each agent i, the total utility that the agent receives from a network G can be expressed as a combination of the information capital u_i^I and cooperation capital u_i^C that i receives from G (we omit G when referring to an arbitrary network):

$$u_i = U(u_i^I, u_i^C). \tag{2}$$

Later, in Section 6, we provide micro foundations for both u_i^I and u_i^C. Here, it is important to simply note that information capital and cooperation capital reflect two fundamental functions of social networks — as a channel to spread and receive information, and as an enforcement device to prompt cooperation and favor exchange — that are typically associated with different subnetwork topologies. In particular, efficient information gathering requires an *expansive* subnetwork such that one person is linked to many others via short network paths (Granovetter, 1973); efficient cooperation typically requires an *interconnected* subnetwork where the local community is tightly connected.[12]

When we later make structural assumptions about u_i^I and u_i^C (Section 6.3), we will be able to calculate the utility of arbitrarily complex networks. Before imposing this structure, we present a 'reduced form' analysis that more transparently illustrates how proxy measures of expansiveness and interconnectedness correlate with migration decisions. The data and measurement strategy are described in more detail in the following section. Section 4 then discusses our identification strategy, and the reduced form results are presented in Section 5. The full structural model is developed and estimated in Section 6.

---

[12]In particular, Ali and Miller (2016) model a dynamic game of repeated cooperation and find that a clique network (a complete network) generates more cooperation and higher average utility than any other networks; Jackson, Rodriguez-Barraquer and Tan (2012) model a game of repeated favor exchanges and highlight 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 footnote 2.

3 Data

To study the empirical relationship between networks and migration, we exploit a novel source of data that contains extremely detailed information on the migration histories and evolving social networks of over one million individuals in Rwanda. These data contain the universe of all mobile phone activity that occurred in Rwanda from January 2005 until June 2009. These Call Detail Records (CDR) were obtained from Rwanda's near-monopoly telecommunications company, and contain metadata on every phone call mediated by the mobile phone network. In total, we observe roughly one billion mobile phone calls between roughly one million unique subscribers. For each of these events, we observe a unique identifier for the caller, a unique identifier for the recipient, the date and time of the call, as well as the location of the cellular phone towers through which the call was routed. All personally identifying information is removed from the CDR prior to analysis. 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 data involving phone numbers in the 95th percentile or greater of social network size.[13]

This section describes the methods used to observe the structure of each individual's social network over time (Section 3.1), and to extract each individual's complete migration history (Section 3.2). Section 3.3 discusses external validity and other measurement concerns.

## 3.1 Measuring social network structure from mobile phone data

The mobile phone data allow us to observe all mobile phone calls placed over a 4.5-year period in Rwanda. These pairwise interactions make it possible to reconstruct a detailed measure of the social network structure of each individual in the dataset. To provide some 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). The large number of friends of friends are also depicted, to provide a sense for the richness of the data.[14]

In the analysis that follows, we relate the network structure of each individual to their

---

[13]Specifically, we calculate the total degree centrality (i.e., the number of unique contacts) for each phone number in the dataset, for each month. Phone numbers in the 95th percentile of this distribution have roughly 200 unique contacts in a single month. We then remove all incoming and outgoing calls from the dataset that involve those numbers in that month.

[14]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.

Figure 2: The social network of a single migrant
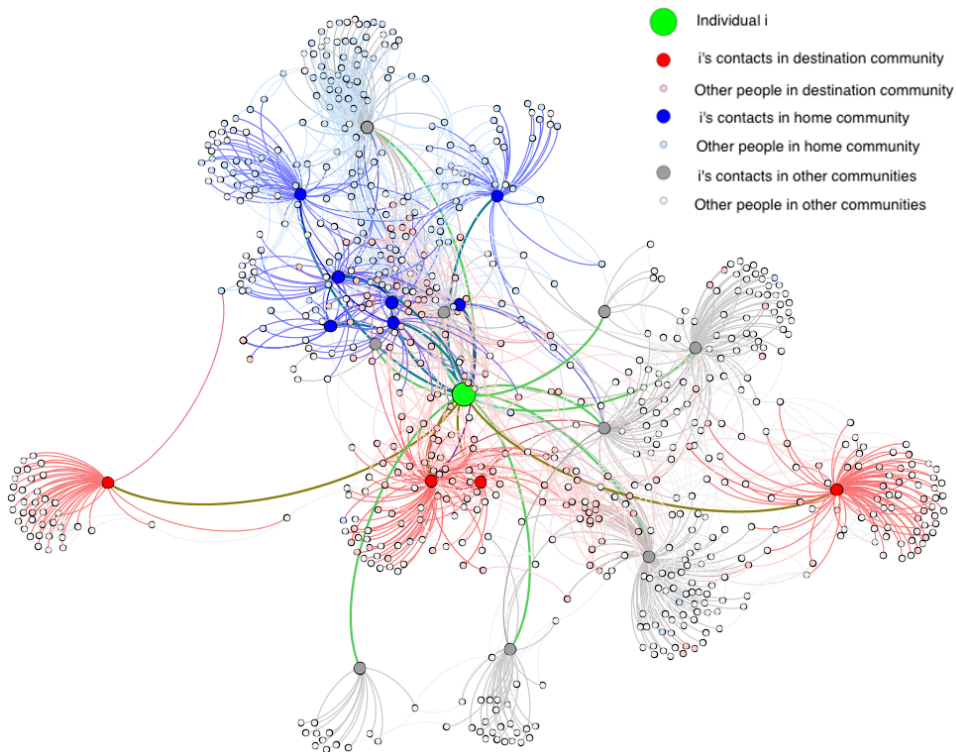
Notes: Diagram shows the social network, as inferred from phone records, of a single migrant i. Nodes represent individuals; edges indicate that two individuals communicated in the month prior to i's migration. Direct contacts of i are shown in blue (for people i's home district), red (for people in 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).

subsequent migration decisions. Following the discussion in Section 2, we focus on a few statistical properties of networks that prior work suggests are important sources of social capital for migrants. The first is *degree centrality*, which simply counts the number of unique individuals with whom each person communicates. This metric most closely reflects the large literature linking migration decisions to the size of an individual's network at the destination (see footnote 3 for classic references). We can separately account for the *strength* of a social tie, which we measure as the number of (undirected) calls between two individuals. In certain analyses we will compare strong and weak ties, where we consider "strong" ties to be those ties in the 90th percentile of the tie strength distribution (equivalent to 5 or more calls per month).[15]

We also show results that correlate migration decisions with crude proxies for the infor-

---

[15]By comparison, Granovetter (1973) defined a weak tie as a tie that was active just once per year.

mation capital u^I and communication capital u^C of a network. Section 6 provides firmer theoretical foundations and a structural approach to measuring u^I and u^C, but in the reduced form analysis that follows, we focus on simple topological properties that roughly differentiate between expansive and interconnected networks.

Information capital. Jackson and Wolinsky (1996) provide one parsimonious measure of information capital as decay centrality, where 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. One concern of the decay centrality is that it assumes an individual receives q^2 from a friend of friend, regardless of how many paths of length two are there between these two individuals. More recently, Banerjee et al. (2013) introduce a notion of diffusion centrality, which accounts for the fact that multiple paths could increase the chance that information makes it from one agent to other. In these two centrality measures, information capital increases with more friends, friends of friends, friends of friends of friends, and so on. Given the empirically estimated discount q is usually far below 1, the closer friends are most important, and thus we construct a simple proxy for information capital using the size of an individual's second-degree neighborhood (or unique *friends of friends*) and third-degree neighborhood (unique friends of friends of friends).

Cooperation capital. Social networks also facilitate interactions that benefit from cooperation and community enforcement, such as risk sharing and social insurance. A consistent set of results has shown that such enforcement is strong when the local network are tightly connected (See footnote 2). In particular, Ali and Miller (2016) model a dynamic game of repeated cooperation and find that a clique network (a complete network) generates more cooperation and higher average utility than any other networks; Jackson, Rodriguez-Barraquer and Tan (2012) model a game of repeated favor exchanges and highlight the importance of *supported* relationships, where a link is supported if the two nodes of the link share at least one common neighbor. Thus, we use two closely related measures of network interconnectedness to proxy for cooperation capital: *network support*, the probability that a friend has one or more common friends; and *network clustering*, the probability that two friends are connected to each other.

Figure 3: Location of all mobile phone towers in Rwanda, circa 2008
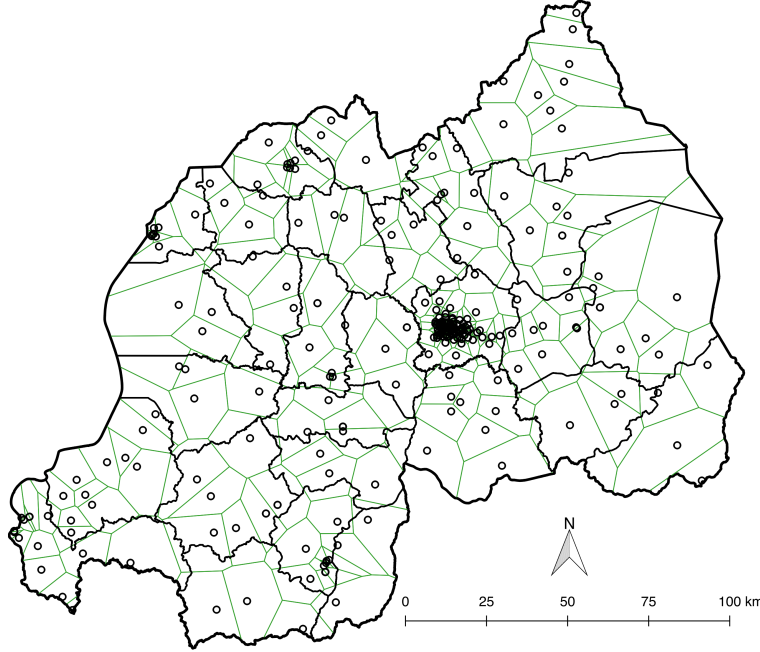
Notes: Black circles indicate cell tower locations. Black lines represent district borders. Green lines show the voronoi polygons roughly divide the country into the coverage region of each tower.

3.2 Measuring migration with mobile phone data

Every time a person uses a mobile phone in Rwanda, the phone company records the time of the event, and the approximate location of the subscriber at the time of the event. We use these 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.[16] The full details of

---

[16]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.

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 is a panel of individual-month districts.[17] 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).[18] Complete details are given in Algorithm 2.

Using these methods, we are able to characterize very granular patterns of internal migration in Rwanda. Summary statistics are presented in Table 1. The first column shows total rates of migration in a single month of the data, using $k = 2$, which defines a migration as an instance where an individual stays in one district for at least 2 months, moves to a new district, and remains in that new district for at least 2 months. The aggregate migration rate in January 2008 is 4.9%; 53.4% of migrants travel from one rural district to another,

---

[17]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.

[18]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 $t + k$.

23.2% travel from rural to urban districts and 23.4% travel from urban to rural districts.[19]

Table 1: Summary statistics of mobile phone metadata

	(1) In a single month (Jan 2008)	(2) Over two years (Jul 2006 - Jun 2008)
Number of unique individuals	432,642	793,791
Number of person-months	432,642	8,121,369
Number of CDR transactions	50,738,365	868,709,410
Number of migrations	21,182	263,208
Number of rural-to-rural migrations	11,316	130,009
Number of rural-to-urban migrations	4,908	66,935
Number of urban-to-rural migrations	4,958	66,264

Notes: Migration statistics calculated from Rwandan mobile phone data. Column (1) based on data from a single month; column (2) includes two years of data, potentially counting each individual more than once. "Migrations" occur when an individual remains in one district for 2 consecutive months and then remains in a different districts for the next 2 consecutive months. We denote as urban the three districts in the capital of Kigali; the remaining districts are considered rural.

To validate these methods, Figure A1 compares the distribution of migration destinations computed from the phone data (red bars) to the distribution of destinations calculated from the 2012 Rwandan census (blue bars), as reported by National Institute of Statistics of Rwanda (2014, p.29). The distributions are not identical, which is expected since the population of phone owners is a non-random sample of Rwandans, but the broad patterns are remarkably consistent across the two approaches to measurement.

While it is reassuring that the aggregate migration rates computed on our data match those reported in traditional surveys, the real advantage of our data is that they can provide a much more granular perspective on internal migration than can be achieved with traditional methods. For instance, the columns of Table A1 disaggregate migration events into several sub-types that are prominent in the literature on internal migration in developing countries (cf. Todaro, 1980, Lucas, 1997, 2015). We observe a striking number of repeat and circular migrants, with a majority of migrants traveling long distances. The data also make it possible to disaggregate migration rates by length of stay. The rows of Table A1 show how 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

---

[19]In Table 1, we classify the three districts that comprise the capital of Kigali as urban, and the remaining 27 districts as rural.

of migration. In later analysis, we show that certain results depend on this definition. But unless otherwise noted, our results define migration as a minimum stay length of $k = 2$, as this most closely matches official statistics on internal migration provided by the Rwandan government.[20]

3.3   Data limitations

While mobile phone data provide uniquely granular insight into the migration decisions and social networks of a large population, there are several important limitations. First, mobile subscribers are not representative of the larger population; in particular, they are wealthier, older, better educated, and are more likely to be male (Blumenstock and Eagle, 2012). While this certainly limits the external validity of our analysis, as we have noted above (and show with Figure A1 and Table A1), the patterns of migration inferred from phone data are broadly consistent with existing data on internal migration in Rwanda. While we do not have survey data that make it possible to directly assess whether phone owners are representative of migrants more generally, we do find that the two populations have similar demographic characteristics.[21]

Second, 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 numbers with very large networks, but this does eliminate potential concerns. When multiple people share the same phone number (which Blumenstock and Eagle (2012) show was not uncommon during this period), we may overestimate the size of an individual's network. Related, it's possible that a single individual might use multiple phone numbers, which would have the opposite effect (in practice, we believe this was less common, since a monopoly operator existed). In principle, our data make it possible to uniquely identify devices and SIM cards, in addition to phone numbers. However, compared to these alternatives, we believe the phone number (which is portable across devices and SIM cards) bears the closest correspondence to the individual subscriber.

Third, 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 the same socio-demographic categories described above; it systematically understates certain

---

[20]According to the 2012 census: 9% of Rwandans are live in a place other than the place they lived in 5 years prior. According to the 2009 Comprehensive Food Security and Vulnerability Analysis, 12% of Rwanda households have a member who migrated in 3 months prior to survey (Feb-Mar 2009).

[21]In particular, separate survey data indicates that the demographic distribution of migrants and non-migrants (i.e., Figures 11 and 12 in National Institute of Statistics of Rwanda (2014)) are quite similar to the demographic distribution of phone owners and non-owners (i.e., Table 2 in Blumenstock and Eagle (2012)).

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 counting edges where communication is reciprocated (see Section 5.4). 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. We find it difficult to imagine how the core results we document below could be a byproduct of non-random selection of true social ties into the sample of ties we observe, but this remains a fundamental limitation of using digital trace data to study social networks.

Finally, the phone data are anonymous and cannot be matched to information about basic economic or demographic information on the individual using each phone. This raises immediate concerns that the network measures we use are simply a proxy for other unobserved confounding variables. However, as we discuss at length in the next section, we use an extremely restrictive set of fixed effects that limits the potential for many of the most worrisome sources of omitted variable bias. However, fixed effects cannot eliminate this potential bias, so in the section below, we carefully articulate the identifying assumption required to interpret our estimates as causal, and provide several robustness tests to explore possible alternative explanations for our results.

4   Identification and estimation

The focus of this paper is on understanding how social networks provide utility that influences 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 try to understand how, holding all such factors fixed, certain variations in social network structure systematically correlate 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, how they compare to someone with network G_3, and so forth. In practice, 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 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 regression estimates.

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 superficially address this concern in two ways. First, we analyze the lagged, rather than contemporaneous, decisions of migrants. Specifically, we relate the migration decision M_{it} made by individual i in month t to the structure of i's social network s months prior. As a concrete example, when $t =$ May 2008 and $s = 2$, we relate the May 2008 migration decision to the structure of the individual's social network in March 2008.[22] 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 (see Figure 1). This is because it seems easier for a migrant to directly control the number of contacts she has in the destination and at home than it is for her to alter the higher-order structure of her social network.

These two 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 4 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. 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

---

[22]Our main specifications use $s = 2$, but in robustness tests we also check $s = 3$ and $s = 1$.

Figure 4: Geographic network structure before and after migration – migrants only
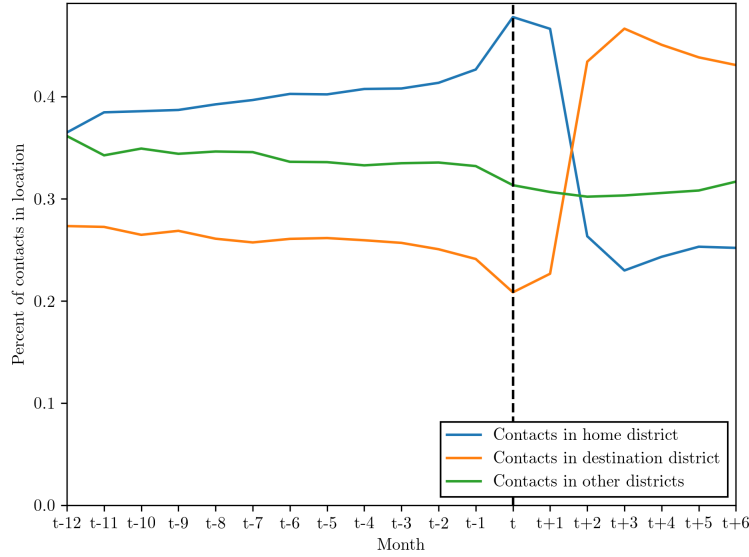
Notes: Figure shows, for a random sample of 10,000 migrants, the average percentage of the migrant's social network in the home and destination districts, in each of the 12 months before and 6 months after migration. Dashed vertical line indicates the date of migration.

in the months before migration; Figure A2b shows analagous results for total call volume. As a sort of 'placebo' test, Figure A3 shows the corresponding figure for non-migrants, where no changes are observed in the "migration" month, as expected (since no migration takes place for this sample).

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 migrants have a relatively constant number of unique friends of friends over time (with no noticeable shift in the months prior to migration). Figure A5 shows similar results for the level of common support in the network.

Omitted Variables

The second 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. 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 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).[23] 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.

Identification

To summarize, the identifying variation we exploit in our main specification is within-individual over time and over potential destinations, net any factors that are shared by all people considering the same origin-destination trip in the same month, and net any effects that are common to all people with exactly the same number of friends in the destination. We would observe such variation over time 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 (the G_2 vs. G_1 comparison of Figure 1) — 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). An example of identifying variation within individual over potential destinations would 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 — and if that additional interconnectedness exceeded the

---

[23]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.

extent to which all networks in that destination were more interconnected in that particular month.

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, the main fixed effects would not absorb this variation. This might occur if carpenters' networks in a particular district are more interconnected (relative to carpenter networks other districts) than 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 the network. To address such concerns, Section 5.2 shows that our main estimates are stable under a series of even more restrictive specifications that include fixed effects for the *individual*-destination (this isolates variation within individual-destination over time and would address the carpenter/farmer concern, if we assume that those trends are temporally stable), for the *individual*-month (which isolates variation across potential destinations for a single individual in a single month), and a few other scenarios.

Estimation

Formally, for a migrant i considering moving from home district h to destination district d in month t, we wish to estimate the effect of (s-lagged) network structure $Z_{ihd(t-s)}$ on the migration decision M_{ihdt}, where M_{ihdt} is a binary variable equal to 1 if the migrant chooses to move from h to d at t and 0 otherwise. We estimate this in two ways, using either a linear model or a discrete choice (multinomial logit) model.

In the linear model:

$$M_{ihdt} = \beta Z_{ihd(t-s)} + \pi_{hdt} + \mu_i + \nu_D + \epsilon_{ihdt} \tag{3}$$

where π_{hdt} are the (home district * destination district * month) fixed effects; and μ_i are the individual fixed effects. We also condition on i's degree centrality in the destination D using a set of fixed effects ν_D that non-parametrically control for effects that are invariant across all people with the same number of contacts in the destination. The coefficient of interest is β, which indicates the average effect of network property $Z_{ihd(t-s)}$ on the probability of migration. Standard errors are two-way clustered by individual and by home-destination-month.

Specification (3) has several attractive properties: it makes it possible to condition on

a rich set of fixed effects, and can be estimated relatively quickly even on a very large dataset. The difficulty with estimating equation (3) arises in how an observation is defined in the regression. In particular, for non-migrants, it is not clear what should be considered the destination network. We address this by defining an observation at the level of the individual-month-*potential destination*. Thus, in each month, each individual comprises 26 observations, one for each of the 26 potential districts to which that individual could migrate in that month.[24]

Our second approach uses a discrete choice (multinomial logit) model of the migration decision, to address the fact that the 26 observations for each individual in each month are not i.i.d. The multinomial logit is becoming increasingly common in the migration literature (Davies, Greenwood and Li, 2001, Dahl and Sorenson, 2010), and has the advantage of providing a sound microeconomic foundation of utility maximization with a random utility model (Mcfadden, 1974, Revelt and Train, 1998). It treats each monthly decision as a single decision with 27 alternatives (one corresponding to staying at home, and 26 migration options).[25] While more natural in this regard, the multinomial logit has several limitations: it is not possible (or at least, quite difficult) to include the same restrictive set of fixed effects as we include in the linear regression, thus increasing the scope for omitted variable bias; it is similarly ill-suited to estimating the impact of individual-specific characteristics (in our case, the attributes of the individual's home network); and the IIA assumption is problematic. Finally, the computational requirements of the multinomial logit are several orders of magnitude greater than that of the corresponding regressions.[26] In practice, the results from the multinomail logit are always qualitatively the same as those from linear regression, so our main analysis is based on specification (3), with multinomial logit results provided as robustness tests in the appendix (see Table A5).

---

[24] An individual is only considered in months where she can be classified as a migrant or a non-migrant in that month. When an individual is classified as "other" (See Section 3), she is excluded for that month.

[25] Another possibility is to model the decision to migrate with a nested logit model, where the individual makes two independent decision: the first is whether or not to migrate and the second is, given the decision to move, the choice of destination (McFadden, 1984, Knapp, White and Clark, 2001). We believe this approach is less appropriate to our context, as the decision to migrate is closely related to the possible destination choices — Davies, Greenwood and Li (2001) provides a more complete discussion of this point.

[26] Whereas equation 3 can be estimated, even with millions of fixed effects and two-way clustered standard errors, in several minutes on our high-performance computing cluster, the panel logit takes several hours, even with minimal fixed effects. This computational constraint is particularly problematic when estimating our effects non-parameterically, as discussed below.

Non-parametric estimation

Equation (3) and the corresponding multinomial logit indicate the average effect of network characteristic Z on the decision to migrate. We are also interested in disaggregating these effects non-parametrically, to understand how such effects differ for migrants with destination networks of different sizes. We thus present a series of figures that show the coefficients from estimating the model:

$$M_{ihdt} = \sum_{k=1}^{D^{max}} \beta_k Z_{ihd(t-s)} \cdot \mathbb{1}(D = k) + \pi_{hdt} + \mu_i + \nu_D + \epsilon_{ihdt} \tag{4}$$

The vector of β_k coefficients from the above model indicates, for migrants with a fixed number of contacts k, the relationship between the migration decision and the higher order network characteristic $Z_{ihd(t-s)}$.

5   Results

Table 2 summarizes the main results from estimating model (3). We find that on average, each additional contact in the destination is associated with a 0.37% increase in the likelihood of migration (Panel A, column 1), and each contact at home is associated with a 0.04% decrease in that likelihood (Panel B, column 1). Columns 2-4 indicate the average effect of changes in high-order structure, after controlling for the immediate contacts of the individual (i.e., the "degree centrality" fixed effects). In column 4, for instance, the second row in Panels A and B indicates that migrants are more likely to go to places where their destination networks are more interconnected, and less likely to leave interconnected home networks. The third row indicates that, perhaps surprisingly, people are not more likely to migrate to destinations where their contacts have a large number of contacts, but they are less likely to leave such places.

Where the first column of Table 2 separately estimates the "pull" and "push" forces of networks on migration (cf. Hare, 1999), the first two columns of Table A2 jointly estimate both effects, to allow for a more direct comparison. Comparing the first two coefficients in the first and second rows, we note that in determining migration outcomes, the marginal effect of an additional contact in the destination is roughly 6.5 to 7.5 times as important as an additional contact at home.

In the subsections below, we discuss these "reduced form" results in greater detail, re-

Table 2: Migration and social network structure - base specification

|	(1)	(2)	(3)	(4)
Panel A: Destination network characteristics				
Degree (network size)	0.0036547***			
	(0.0000102)			
% Friends with common support		0.0014813***		0.0014808***
		(0.0001146)		(0.0001146)
Unique friends of friends			−0.0000005	−0.0000002
			(0.0000009)	(0.0000009)
Observations	9,889,981	9,889,981	9,889,981	9,889,981
R^2	0.1851423	0.1853017	0.1852869	0.1853017
Panel B: Home network characteristics				
Degree (network size)	−0.0003985***			
	(0.0000049)			
% Friends with common support		−0.0003467		−0.0005710**
		(0.0002422)		(0.0002424)
Unique friends of friends			−0.0000089***	−0.0000089***
			(0.0000004)	(0.0000004)
Observations	9,889,981	9,889,981	9,889,981	9,889,981
R^2	0.1743203	0.1750909	0.1751320	0.1751325
Degree fixed effects	No	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes
Home*Destination*Month F.E.	Yes	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. Standard errors are two-way clustered by individual and by home-destination-month. *p<0.1; **p<0.05; ***p<0.01.

estimate each average effect non-parametrically, and discuss heterogeneity in the migration response by migrant and location type. The analysis reveals considerable nuance in the relationship between networks and migration, helps explaining the "surprising" result in Table 2, and establishes a set of stylized facts that form the basis for structural model of social capital.

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

Our first result validates a central thesis of prior research on networks and migration, which is that individuals are more likely to migrate to places where they have more connections. The unconditional relationship between degree centrality at destination (i.e., the number of unique contacts of the individual) is shown in Figure 5a. A point on this figure can be

24

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

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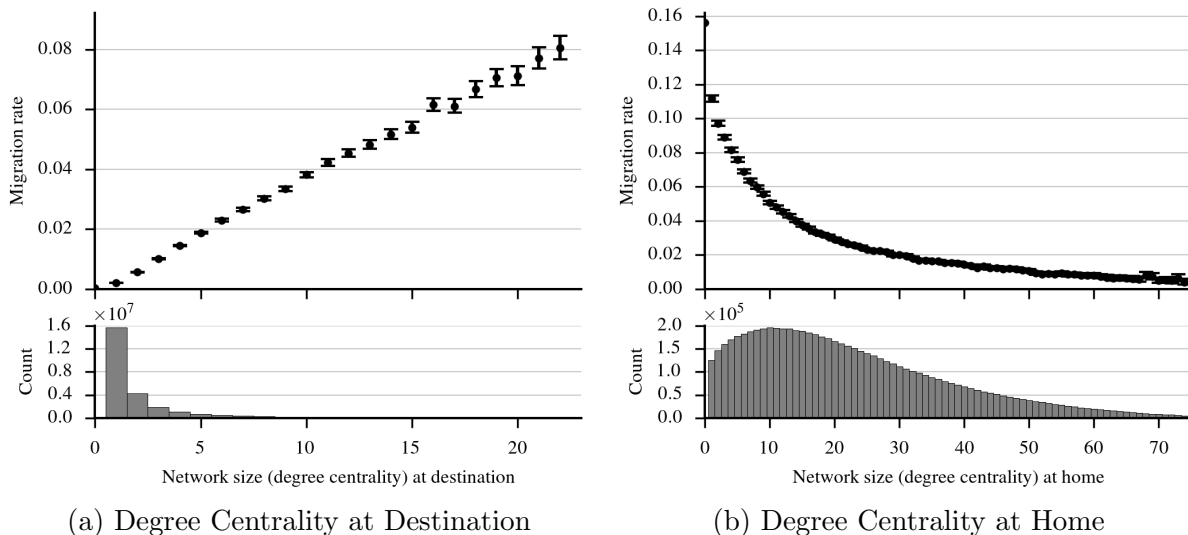
Notes: In both (a) and (b), the lower histogram shows the unconditional degree distribution, i.e., for each individual in each month, the total number of contacts in the (a) destination network and (b) home network. The upper figure shows, at each level of degree centrality, the average migration rate.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

interpreted as the average migration rate (y-axis) across individuals with a fixed number of contacts in the destination (x-axis). For instance, roughly 4% of individuals who have 10 contacts in a potential district d' in month $t-2$ are observed to migrate to d' in month t. The bottom panel of the figure shows the distribution of destination degree centrality, aggregated over individuals, months (24 total), and potential destinations (26 per individual).

This figure also provides intuition for our identification strategy and preferred empirical specification. The average migration rates depicted Figure 5a are likely confounded by a variety of omitted variables. For instance, people in rural districts typically know more people in the urban capital of Kigali than in other districts, and rates of migration to Kigali are higher than to other districts. Thus, Figure A6 re-estimates the migration rates of Figure 5a, conditioning on a series of increasingly restrictive fixed effects. In the first panel, Figure A6a reports the ν_k coefficients and standard errors from estimating:

$$M_{ihdt} = \sum_{d=1}^{D^{max}} \nu_d \mathbb{1}(D = d) + \epsilon_{ihdt} \tag{5}$$

Mechanically, these coefficients are identical to unconditional correlations shown in Figure 5a, albeit shifted down because of the omitted global intercept. In subsequent panels, Figure A6b includes destination district fixed effects (which most immediately addresses the

Kigali concern described above). Figure A6c replaces destination fixed effects with more stringed destination-origin-month fixed effects. Finally, Figure A6d adds individual fixed effects, resulting in an estimating equation similar to equation 4:

$$M_{ihdt} = \sum_{d=1}^{D^{max}} \nu_d \mathbb{1}(D = d) + \pi_{hdt} + \mu_i + \epsilon_{ihdt} \tag{6}$$

In all figures, the qualitative relationship is remarkably unchanged. Individuals with more contacts in a destination community are more likely to migrate to that community. We also see that this relationship is positive, monotonic, and approximately linear with elasticity one. In other words, individuals with k times as many contacts in a destination district are k times more likely to migrate to that district.

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.

5.2 Higher-order network structure

We next examine how the high-order structure of the individual's network — i.e., the connections of the individual's contacts — relate to subsequent migration decisions. We focus on the proxies for network interconnectedness and expansiveness described in Section 3.1.

Network 'interconnectedness'

Figure 6 documents the relationship between migration decisions and the interconnectedness of the individual's social networks, making the generalized comparison between G_1 and G_2 in Figure 1. As described in Section 3.1 and originally proposed in Jackson, Rodriguez-Barraquer and Tan (2012), we measure this interconnectedness as network "support," or the fraction of i's contacts who have one or more friends in common with i. In later robustness tests, we show that related measures of network interconnectedness, tightness, and clustering, produce qualitatively similar results.[27]

Both at home and in the destination, the unconditional relationship between migration and interconnectedness is ambiguous. Figures 6a and Figure 6c show how migration

---

[27]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).

Figure 6: Migration and network "tightness" (friends with common support)
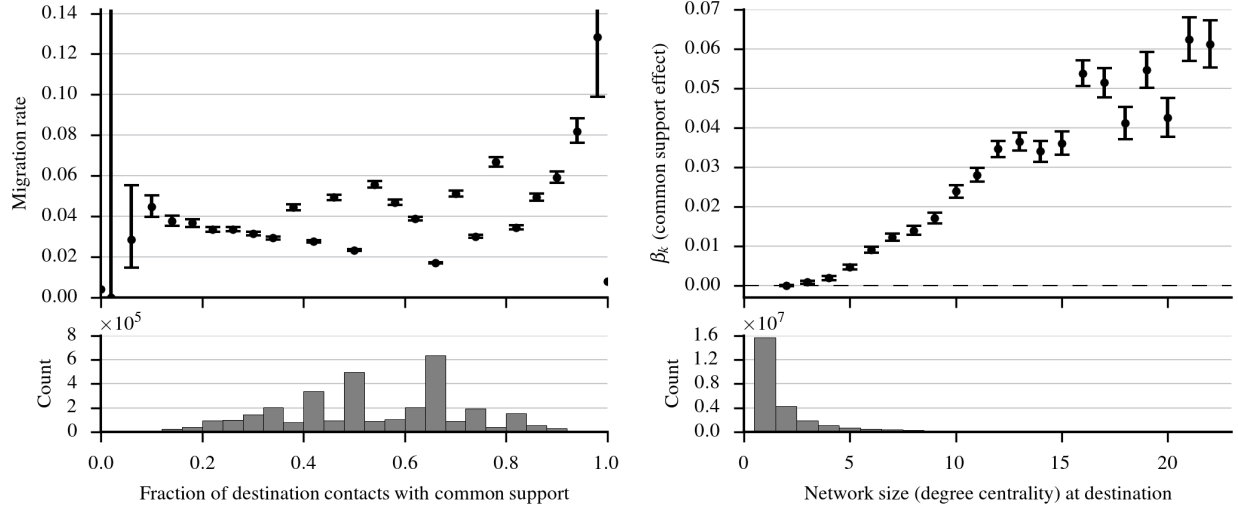
(a) Network support at destination

(b) Network support at destination, by degree
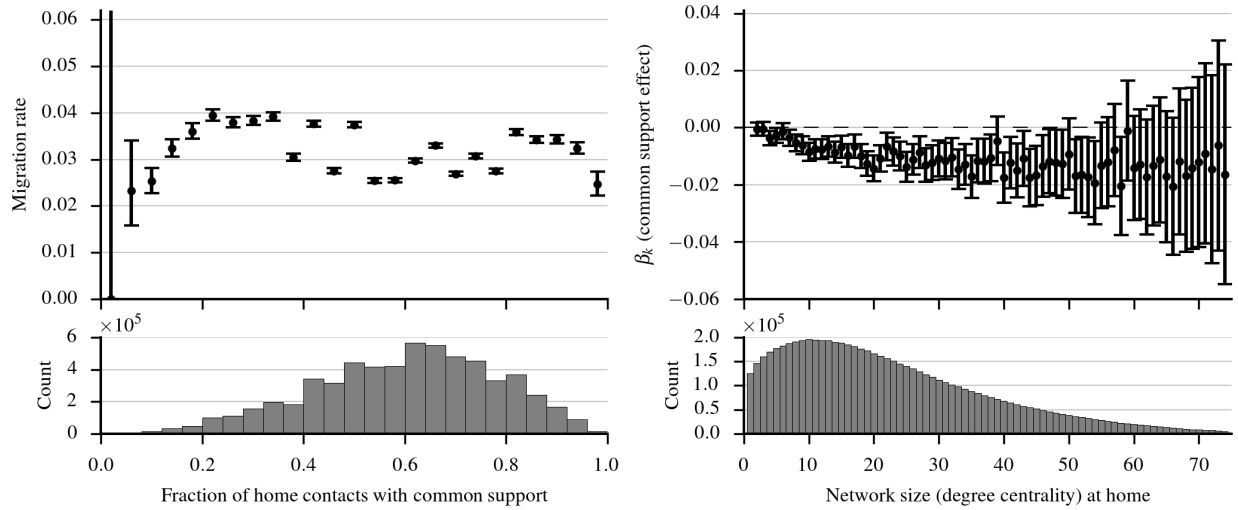
(c) Network support at home

(d) Network support at home, by degree

Notes: Network support indicates the fraction of contacts supported by a common contact (see Section 3.1). In all figures, the lower histogram shows the unconditional distribution of the independent 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 4, 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 indicate 95% confidence intervals, clustered by individual.

varies with network support in the destination and at home, respectively. However, this unconditional relationship is potentially confounded by a large number of omitted variables, including the fact that network support is generally decreasing in degree, since the larger an individual's network, the harder it is to maintain a constant level of support.

Holding degree fixed, a clear pattern emerges: people are systematically drawn to places where their networks are more interconnected. This pattern is evident in Figure 6b, which plots the β_k coefficients estimated from model (4) on the destination social network, all of which are positive. Figure 6d show that, holding degree fixed, people are significantly less likely to leave home if their home contacts are more interconnected. Appendix Figure A7 replicates this analysis using the network clustering, instead of network suppport, as a measure of interconnectedness. Results are qualitatively unchanged.

The fact that people are more likely to go to places where their networks are interconnected may not be surprising, but in other settings, the opposite result has been documented. For instance, Ugander et al. (2012) show that people are more likely to sign up for Facebook when their pre-existing Facebook friend network is less interconnected.

Network 'expansiveness'

The relationship between migration and network expansiveness is more surprising and subtle. Here, we focus on the number of unique friends of friends a person has in a given region, i.e., the generalized comparison between G_1 and G_3 in Figure 1. Without controlling for the size of an individual's network, there is a strong positive relationship between migration and expansiveness in the destination (Figure 7a), and a strong negative relationship with expansiveness in the origin (Figure 7c). The shape of these curves resemble the relationship between migration rate and degree shown earlier in Figure 5: the average migration rate increases roughly linearly with the number of friends of friends in the destination, and decreases monotonically but with diminishing returns to friends of friends at home.

Of course, the number of friends of friends a person has is largely determined by the number of friends that person has. Thus, Figures 7b and 7d show how the number of friends of friends relates to migration, holding fixed the number of friends (as well as the other fixed effects in model (4)). For the home network, Figure 7d indicates the expected pattern: the fact that all of the coefficients are negative suggests that given a fixed number of friends at home, people are less likely to leave when those friends have more friends.

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 $1-3$ destination contacts, but negative for people with > 4 destination contacts. Averaged over all migrants, this effect is negative and insigificant (row 3 of Tables 2 and A2). This result is difficult to reconcile with most standard models of information diffusion, such as

28

Figure 7: Relationship between migration and "expansiveness" (unique friends of friends)
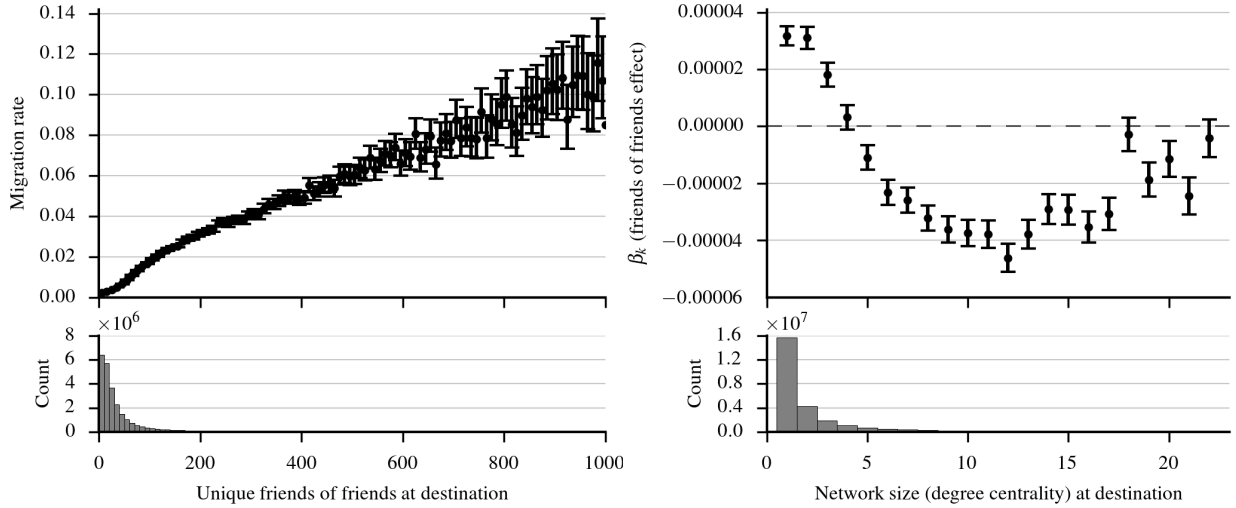
(a) Friends of friends at destination

(b) Friends of friends at destination, by degree
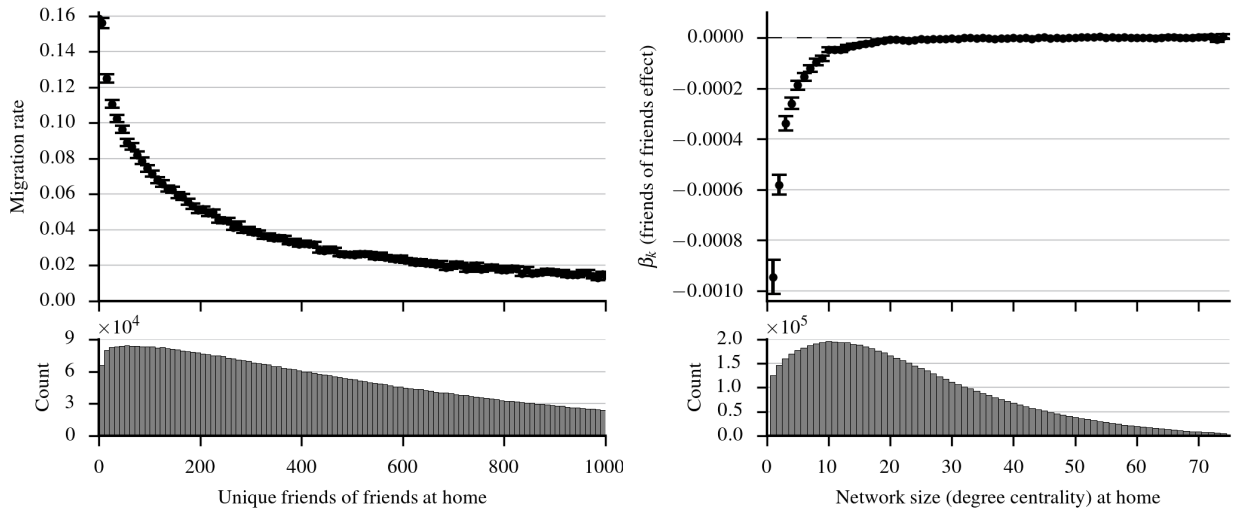
(c) Friends of friends at home

(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 4, 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 indicate 95% confidence intervals, clustered by individual.

those proposed in Banerjee et al. (2013) and Kempe, Kleinberg and Tardos (2003). Indeed, much of the literature on migration and social networks seems to imply that, all else equal, individuals would be more likely to migrate if they have friends with many friends, as such networks would provide more natural conduits for information about job opportunities and

the like.[28]

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 5.4. 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 A9, 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 of the population, it is apparent that, on average, non-migrants have more friends of friends in the destination networks than migrants.

5.3   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. In particular, Tables A6-A10 disaggregate the results from Table A2 along several dimensions that are salient in the migration literature: whether the migrant has previously migrated to the destination (Table A6); whether the migration is between adjacent districts or over longer distances (Table A7); whether the migrant stays in the destination for a long period of time (Table A8); and whether the migration is to an urban or rural destination (Tables A9 and A10).

Heterogeneity and unawareness of the broader network

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

This heterogeneity suggests one possible explanation for the unexpected 'friend of friend' result of Figure 7b. Namely, many of the migrants who are positively influenced by expansive

---

[28]A very similar pattern appears in Figure A8 when we look at the *home* friends of the friend in the destination. In other words, if migrant i in home district h has a friend j in destination district d, we find that people are less likely to migrate to places where j has more friends located in h. (Where Figure 7bb analyzes the relationship between migraiton and the number of j's friends in d, Figure A8 analyzes the number of j's friends in h).

networks are the migrants who seem likely to be more familiar with the structure of their destination networks. Such an interpretation is consistent with the possibility that the average migrant may simply be unaware of the extent to which their friends are connected to other unknown individuals (which would predict a null average effect),[29] but that these "in the know" migrants do value having more friends of friends.

Strong ties, weak ties, and recent migrants

A different explanation for the 'friend of friend' result is suggested by a closer analysis of the role of strong and weak ties in migration. Here, and consistent with recent work by Giulietti, Wahba and Zenou (2018), we find that both strong and weak ties matter in migration: the effect of a strong destination tie is roughly 1.5 times that of a weak destination tie; at home, the effect of a strong tie is roughly twice as large as the effect of a weak tie. These results are shown in Table A11, which defines a strong tie as one that supports five or more communication events in the reference month (the 90th percentile of communication frequency) — see Section 3.1 for details and justification. Recent migrants have a similar effect: people are more likely to go to places where they know recent migrants (defined as a contact who previously made the origin-destination migration that the individual is considering).[30] However, neither strong ties nor recent migrants dominate the migration decision: when controlling for either factor, the main effects reported in Table 2 are qualitatively unchanged.

More interesting is the role that *higher order* tie strength plays in modulating the migration decision. In particular, the results in Section 5.2 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 — or 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 A12 and A13. 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,

---

[29]For other instances where people appear to have incomplete information about the friends of their friends, see Friedkin (1983), Casciaro (1998), and Chandrasekhar, Breza and Tahbaz-Salehi (2016).

[30]Coefficient estimates in Table A14 indicate that knowing a recent migrant in the destination increases the likelihood of migration by roughly 3.5X the amount as knowing anyone else in the destination. The effect is slightly larger for recent migrants who arrived in the destination very recently (last month) than for recent migrants who arrived at any point prior.

Figure 8: The role of (higher order) strong and weak ties in a migrant's network
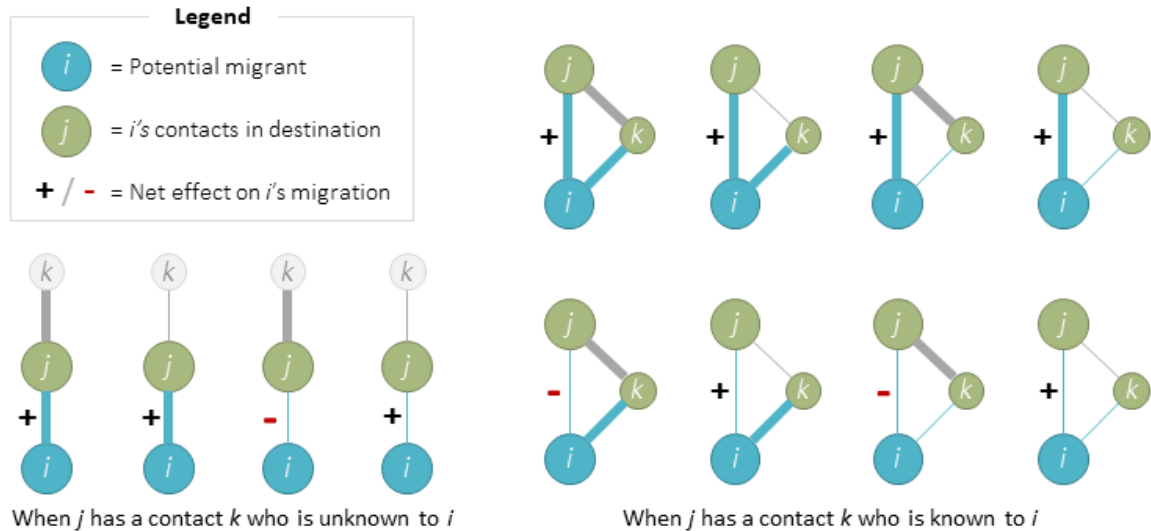
Notes: Thick edges represent "strong" ties and thin edges represent "weak ties" The $+/-$ signs summarize the effect that j has on i's likelihood of migration, based on coefficients in Tables A12 and A13.

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 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, which show all possible configurations of a supported i-j tie, 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,[31] 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. However, these results — and particularly the results concerning the "friend of friend" effect — are more speculative than conclusive. We take these ambiguities as motivation to develop a more coherent model of how migrants derive utility from networks, which we turn to in Section 6.

---

[31]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."

5.4   Robustness and identification (revisited)

Section 4 describes the identifying assumptions behind our regressions. In particular, when estimating models (3) and (4), we assume $E[\epsilon_{ihdt}|\pi_{hdt}, \mu_i, \nu_D] = 0$. 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. While we believe these fixed effects address the most concerning sources of bias, it is of course possible to concoct a scenario in which this assumption would be violated (as in the carpenter/farmer example in Section 4).

We therefore run a series of robustness checks that further isolate the identifying variation behind the regression results presented above. In particular, Appendix Table A4 re-estimates the main effect shown in column 4 of Table 2 under a variety of increasingly restrictive fixed effect specifications. Column 1 replicates the prior result, including fixed effects for π_{hdt}, μ_i, and ν_D. Column 2 in Table A4 then includes fixed effects for each *individual-month* pair, so that the identifying variation comes within individual in a given month but across potential destination districts.[32] Column 3, by contrast, includes separate fixed effects for each *individual-destination* pair, so that the β coefficients are identified solely by variation within individual-destination over time.[33] 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).

In addition to these variations on the core regression specification, we also re-estimate our results using a discrete choice (multinomial logit) model. As noted earlier, this is a more natural specification as it treats each monthly decision as a single decision with 27 alternatives (one corresponding to staying at home, and 26 migration options). Results are shown in Table A5, and are broadly consistent with the main regression results presented

---

[32]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.

[33]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).

earlier.

Finally, 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 from the paper but can provide them to interested readers upon request:

- **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 or text message between these individuals. As a robustness check, we take a more restrictive definition of social network and only include edges if i initiates a call or sends a text message to j and j initiates a call or sends a text message to i.

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

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

6   Structural estimation

The reduced form results presented in Section 5 highlight *how* social networks influence migration decisions, but offer limited insight into *why* some network structures matter more than others. Since the phone data contain no identifying or socio-demographic information about the individual subscribers, we have limited ability to infer whether, for instance, interconnected networks are influential because they tend to consist of family members, co-ethnics, or some other tightly knit community. The regression specifications are also

limited by the fact that different measures of higher-order network structure are highly inter-dependent, so it is difficult to isolate the effect of marginal changes to the network.

For these reasons, we return to the stylized model of Section 2, which describes how different subnetwork topologies provide utility to migrants, and use the revealed preference decisions in our data — to migrate or not to migrate — to parameterize a model of network-based social capital and migration. Recall that we say that an individual i receives utility $u_i(G)$ from a social network G. As emphasized in the literature, we assume that $u_i(G)$ is primarily comprised of information capital and cooperation capital. The next two subsections provide micro foundations for these two types of social capital.

6.1 Information capital: competition and 'expansiveness'

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 1). 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; that agent's diffusion centrality measures the impact of this information to the network. We study a receiving process in which each agent is endowed with one unit of information, and 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 measure we seek turns out to be the diffusion centrality, because the flow of information is symmetric.

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, they share information with each other with probability $q \in (0, 1)$.

In this benchmark model of information sharing, more expansive 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 expansive destination networks are not positively correlated with migration, and with the evidence that suggests some rivalry for attention (see Section 5.3).

We model the source of competition for attention as costly socializing, 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 and socializing 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 socializing 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

$$\max_{Q} \; dv(Q/d)^\beta - cQ^\omega. \tag{7}$$

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

$$Q/d = \frac{1}{d^\lambda}\left(\frac{\beta v}{\omega c}\right)^{\frac{1}{\omega - \beta}}, \;\; \text{where } \lambda = \frac{\omega - 1}{\omega - \beta} \in [0, 1]. \tag{8}$$

Notice that if the cost is linear ($\omega = 1$), then the marginal cost of socializing 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 observation, we let the interaction between each pair of linked agents ij depend on their degrees. In particular, let the frequency of their interaction with possible competition for attention be discounted by $\frac{1}{d_i^\lambda d_j^\lambda}$. 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,

$$DC(G; q, \lambda, T) \equiv \sum_{t=1}^{T}(q\tilde{G})^t \cdot \mathbf{1}, \;\; \text{and} \;\; \forall ij, \; \tilde{G}_{ij} = \frac{1}{d_i^\lambda d_j^\lambda}G_{ij}. \tag{9}$$

When $\lambda = 0$, this is the original diffusion centrality, which assumes 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 denote a *path* between agent i and j as a sequence of distinct agents $(i_0, i_1, \ldots, i_k)$ such that $i_0 = i$, $i_k = j$ and $G_{i_l i_{l+1}} = 1$ for any $l \in \{0, \ldots, k-1\}$. Let $D(i,j)$ be the *distance* between agent i and j in the network, defined as the length of the shortest path between them. 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 A1. 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 expansive networks to be either beneficial or harmful.

6.2 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, Rodriguez-Barraquer and Tan (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, agent i and j choose their effort levels a_{ij}, a_{ji} in $[0, \infty)$ as their contributions to a joint project.[34] 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

---

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

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 addition 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, and this information travels instantly.[35]

To begin, we seek to minimize contagion of deviation to the rest of the society off the equilibrium path, which follows from Jackson, Rodriguez-Barraquer and Tan (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.

---

[35]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.

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 agents 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:

$$b(a) \leq a + \int_0^\infty e^{-rt} \delta a \, dt. \tag{10}$$

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.

$$b(a) \leq a + 2 \int_0^\infty e^{-rt} \delta a \, dt. \tag{11}$$

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 the incentive of cooperation for the triangle ijk because it can sustain a^T by itself.

6.3   A benchmark model of migration

We now return to the migration decision. Through equation (2)), we assume that the utility of a network can be expressed as a linear combination of the information capital u_i^I and cooperation u_i^C capital that i receives from G. 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 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).[36] Appendix A3.2 shows that similar results obtain when we consider a log-linear (Cobb-Douglas) utility function.

As outlined in Section 6.1, 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 (9). We derive i's cooperation capital from Corollary 1 in Section 6.2, which implies that supported links are more valuable than unsupported links:

$$u_i^C = u_1 d_i^{NS} + u_2 d_i^S, \tag{12}$$

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

$$u_i = u_0 DC_i(q, \lambda, T) + u_1 d_i^{NS} + u_2 d_i^S. \tag{13}$$

We will use this model to contrast the value of information capital against the value of

---

[36]The network game approach follows in the tradition of Ballester, Calvó-Armengol and Zenou (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, could not be realistically computed on our data. (As a point of comparison, calibration of the much simpler mode (2) takes several days to complete, even after being parallelized across a compute cluster with 96 cores). See also Guiteras, Levinsohn and Mobarak (2019) for a related structural approach to dealing with network inter-dependencies.

cooperation capital, so we replace the parameters (u_0, u_1, u_2) by (π^I, π^C, α) and rewrite the overall utility:

$$u_i = \pi^I DC_i(q, \lambda, T) + \pi^C \left(d_i + \alpha d_i^S \right). \tag{14}$$

Substituting (14) into the original migration decision (1), we have

$$
\begin{aligned}
\pi^{I,d} DC_i(G^d; q, \lambda, T) &+ \pi^{C,d} \left(d_i(G^d) + \alpha^d d_i^S(G^d) \right) \\
&> \pi^{I,h} DC_i(G^h; q, \lambda, T) + \pi^{C,h} \left(d_i(G^h) + \alpha^h d_i^S(G^h) \right) + \varepsilon_i. \tag{15}
\end{aligned}
$$

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 (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.

6.4   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 (15). 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 $T = 3$.[37] Since a very large number of combinations of possible

---

[37]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

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.[38]

Estimation appears to be well-behaved. For instance, Figure A12 shows the home and destination utility values for all 270,000 individuals, using the parameterized version of model (15). Most of the true migrants (blue dots) have a predicted destination utility that exceeds their home utility; most of the true non-migrants (red dots) have a higher home utility. In aggregate, the calibrated model correctly classifies roughly 70% of the migration events.

To provide more intuition for the model estimation process, Figure 9 shows the estimation plots for λ; similar plots for the remaining five parameters are shown in Figure A11. 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 model (15), 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 5.[39] Moreover, when we take the *simulated* migration decisions $\widehat{M_{ihdt}}$ from the parameterized structural model, and estimate the equivalent of model (3) with $\widehat{M_{ihdt}}$ as the dependent variable, the regression results, presented in Table A15, are qualitatively the same as the regression results using the actual migration decision M_{ihdt} (Table 2). The only notable difference is the effect of unique friends of friends in the destination network, which becomes significantly negative in Table A15 and was insignificant in Table 2. This shows that when the rivalry parameter λ is optimally chosen for the structural model, the average

---

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.

[38]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.

[39]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 3.1 for definitions). In this linear regression (no fixed effects), $R^2 = 0.975$.

Figure 9: Model calibration results for λ
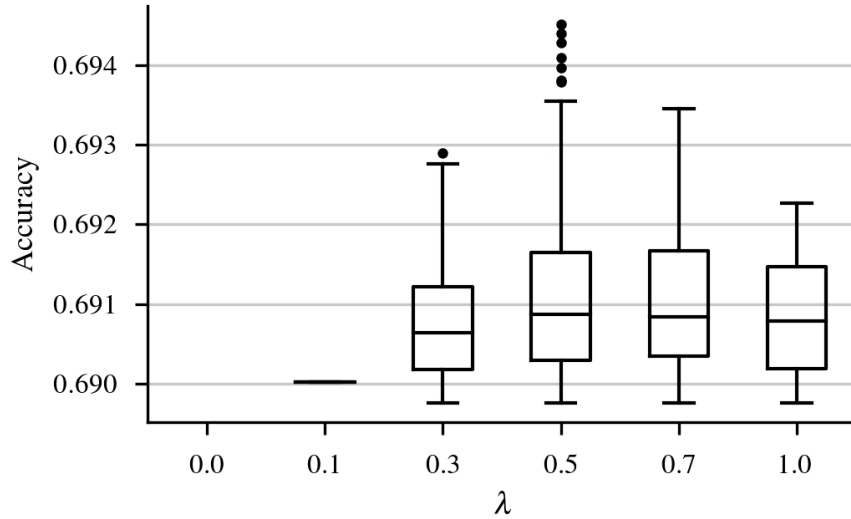
Notes: Figure shows the marginal effect of varying λ when calibrating Model (15). The full model has 7 parameters $(\lambda, \alpha^d, \alpha^h, \pi^{I,d}, \pi^{C,d}, \pi^{I,h})$; roughly 50,000 different parameter combinations are tested. The top percentile (by accuracy) of these combinations are selected. Each box and whisker plot represents the accuracy distribution within that top percentile, for each value of λ tested.

effect of one's second-neighborhood becomes negative.

6.5 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 Figure 9. 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 10, which shows the distribution of predicted utility from u_i^I and u_i^C for each of the individuals used to

43

Figure 10: Calibration results: 'information' and 'cooperation' utility
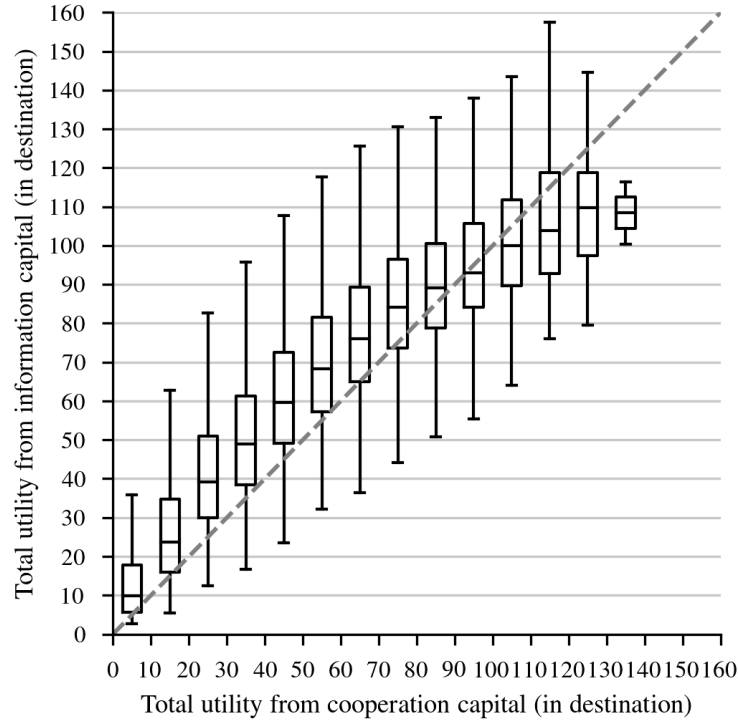
Notes: Figure shows the distribution of predicted utility from 'information' capital and 'cooperation' capital (i.e., equation 2) for 270,000 migrants and non-migrants.

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 5, 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 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 reestimate a version of model (15) where the rivalry coefficient is fixed at $\lambda = 0$. In other words, we use the original diffusion centrality (without *lambda*) 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; as shown in Figure A13, the bulk of the distribution lies far below the 45-degree line, where $u_i^I < u_i^C$.

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 A11. In particular, $\alpha^d = 5$ implies that one supported link in the destination is six

44

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 5.3, but only directly tested through structural estimation.

As a final step, Appendix A3 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:

$$u_i(G^d) > u_i(G^h) + \tau + \varepsilon_i. \tag{16}$$

Separately, instead of the linear form of (15), we consider a Cobb-Douglas utility function which implies a log-linear combination of information capital and cooperation capital. Equation (15) becomes

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

Results in Appendix A3 show that the key qualitative results persist under these alternative specifications of model (15).

7 Conclusion

Social networks play a critical role in economic decisionmaking. This paper studies the decision to migrate in order to understand the value of social networks. Relative to prior work on the topic, our data provides uniquely granulr visibility into the structure of social networks and the migration events they precipitate.

There are two main sets of findings. The first are specific to migration, and perhaps even to internal migration in Rwanda. These results establish several new stylized facts. Perhaps most surprising, we find that most migrants are *not* drawn to places where their social net-

works are expansive and diffuse. Our structural results suggest that this aversion may stem from the fact that migrants feel competition for the attention of their well-connected friends. By contrast, migrants respond strongly to the interlinkages of their friend and kinship networks, and are consistently drawn to networks that are interconnected and embedded. Such a finding is consistent with recent evidence that risk sharing and favor exchange play an important role in the migration decision (Munshi and Rosenzweig, 2016, Morten, 2019). But we also find that the notion of the "average migrant" can be a misleading generalization. Our data reveal rich heterogeneity, and we find that different types of migrants — including repeat, long-term, and short-distance migrants — value different properties of social networks differently.

The second set of results speak more generally to the utility that social networks provide to individuals embedded in those networks. 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 difusion have seen remarkable success. Such models imply a prominent (albeit highly stylized) narrative that the primary function of networks is to diffuse information about economic opportunities (cf. Rees, 1966, Ioannides and Datcher Loury, 2004). But the patterns revealed by our data are hard to reconcile with these models, and instead point to a model of network utility where repeated cooperation, and rivalry in information diffusion, play a more prominent role.

More broadly, we are hopeful that this study can illustrate the potential for novel sources of network data to provide deeper insight into how individuals derive utility from their social networks. Such data capture incredibly rich structure that reveal hitherto unobserved correlations between networks and consequential economic decisions. Through a combination of rich descriptives and structural estimation, we see great potential for future work aimed at understanding the value of social networks.

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Breza, Emily, Arun G Chandrasekhar, Tyler H McCormick, and Mengjie Pan. 2017. "Using Aggregated Relational Data to Feasibly Identify Network Structure without Network Data." National Bureau of Economic Research Working Paper 23491.

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

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

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

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

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

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

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

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

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

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

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

Dahl, Michael S., and Olav Sorenson. 2010. "The migration of technical workers." *Journal of Urban Economics*, 67(1): 33–45.

Davies, Paul S., Michael J. Greenwood, and Haizheng Li. 2001. "A conditional logit approach to US state-to-state migration." *Journal of Regional Science*, 41(2): 337–360.

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

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

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

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

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

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

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

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

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

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

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

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

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

Hare, Denise. 1999. "'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*, 35(3): 45–72.

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

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

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

Jackson, Matthew O. 2018. "A Typology of Social Capital and Associated Network Measures." *Social Science Research Network* SSRN Scholarly Paper ID 3073496, Rochester, NY.

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

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

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

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

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

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

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

Kempe, David, Jon Kleinberg, and Éva Tardos. 2003. "Maximizing the Spread of Influence Through a Social Network." *KDD '03*, 137–146. New York, NY, USA:ACM.

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

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

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

Knapp, Thomas A., Nancy E. White, and David E. Clark. 2001. "A Nested Logit Approach to Household Mobility." *Journal of Regional Science*, 41(1): 1–22.

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

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

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

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

Lucas, Robert E. B. 2015. "Chapter 26 - African Migration." In *Handbook of the Economics of International Migration.* Vol. 1 of *Handbook of the Economics of International Migration,* , ed. Barry R. Chiswick and Paul W. Miller, 1445–1596. North-Holland.

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

Mcfadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in econometrics.* , ed. P Zarembka, 105–142. Academic Press.

McFadden, Daniel L. 1984. "Econometric analysis of qualitative response models." *Handbook of Econometrics,* 2: 1395–1457.

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

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

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

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

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

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

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

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

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

National Institute of Statistics of Rwanda. 2014. "Migration and Spatial Mobility." Kigali, Rwanda.

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

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

Revelt, David, and Kenneth Train. 1998. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *The Review of Economics and Statistics*, 80(4): 647–657.

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

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

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

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

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

Appendices - For Online Publication

A1 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,

$$DC_{ij}(G'; q) - DC_{ij}(G; q)$$

$$\geq q\left(\frac{1}{d_i(d_j-1)} - \frac{1}{d_i d_j}\right) + q^2\left(\sum_{h \in N_j \setminus \{k\}}\left(\frac{1}{d_i(d_j-1)^2 d_h} - \frac{1}{d_i d_j^2 d_h}\right) - \frac{1}{d_i d_j^2 d_k}\right)$$

$$\geq q\left(\frac{1}{d_i(d_j-1)} - \frac{1}{d_i d_j}\right) - q^2\frac{1}{d_i d_j^2 d_k}$$

$$= q\frac{1}{d_i(d_j-1)d_j} - q^2\frac{1}{d_i d_j^2 d_k}$$

$$> 0.$$

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:

$$DC_{ij}(G'; \lambda, q) - DC_{ij}(G; \lambda, q)$$

$$= q\left(\frac{1}{d_i^\lambda(d_j-1)^\lambda} - \frac{1}{d_i^\lambda d_j^\lambda}\right) - q^2\frac{1}{d_i^\lambda d_j^{2\lambda} d_k^\lambda} + q^2\sum_{h \in N_j \setminus g_k}\left(\frac{1}{d_i^\lambda(d_j-1)^{2\lambda} d_h^\lambda} - \frac{1}{d_i^\lambda(d_j)^{2\lambda} d_h^\lambda}\right)$$

$$+ q^2\sum_{l \in N_j \cap N_k}\left(\frac{1}{d_i^\lambda(d_j-1)^{2\lambda}(d_l-1)^\lambda} - \frac{1}{d_i^\lambda(d_j)^{2\lambda} d_l^\lambda}\right). \tag{18}$$

Clearly, each of the four terms in (18) 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:

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

  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:

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

- …

- 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:

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

55

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   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 (2). 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 $G_{ij}^s = G_{ji}^s = 1$ 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):

$$u_i(\mathbf{a}, G) = \pi a_i - \frac{a_i^2}{2} + \phi \sum_{j=1}^{n} G_{ij} a_i a_j + \alpha \sum_{j=1}^{n} G_{ij}^s a_i a_j. \tag{19}$$

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.[40] 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, $\alpha \sum_{j=1}^{n} G_{ij}^s a_i a_j$. This is motivated by the theory results in Section 6.2 and the empirical results in Section 5 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 6.1, $DC(G; q, \lambda, T)$ when $q = \phi$, $\lambda = 0$ and $T \to \infty$. In particular, ϕ can be viewed as the information passing probability q. The equilibrium action of agent i depends on the entire network structure, including her indirect

---

[40]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.

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 (19) has a unique (and interior) Nash equilibrium in pure strategies given by:*

$$\mathbf{a}^* = \pi(\mathbf{I} - \phi G - \alpha G^s)^{-1}\mathbf{1}. \tag{20}$$

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

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

This leads to

$$a_i^* = \pi + \phi \sum_{j=1}^n G_{ij} a_j^* + \alpha \sum_{j=1}^n G_{ij}^s a_j^*. \tag{21}$$

In the matrix form: $\mathbf{a}^* = \pi\mathbf{1} + \phi G\mathbf{a}^* + \alpha G^s\mathbf{a}^*$, which leads to the solution in (20).

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

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

We omit the details of the analogous proof, which can be found in Bramoullé, Kranton and D'amours (2014) and Jackson and Zenou (2015).

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

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

$$= a_i^* \left(\pi + \phi \sum_{j=1}^n G_{ij} a_j^* + \alpha \sum_{j=1}^n G_{ij}^s a_j^* \right) - \frac{a_i^{*2}}{2}.$$

By equation (21), $u_i(\mathbf{a}^*, G) = (a_i^*)^2/2$, which by equation (20) 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 6.1. For example, Ballester, Calvó-Armengol and Zenou (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.

A3   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 1), 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 2). 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 (1) and (2), to test the robustness of the calibration results in Section 6.5.

A3.1   Fixed migration costs

We first allow for the migration decision (equation 1) to include a fixed threshold (cost) τ, in addition to the idiosyncratic error ε_i:

$$u_i(G^d) > u_i(G^h) + \tau + \varepsilon_i. \tag{22}$$

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 (22) is calibrated with the data, the main observations in Section 6.5 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 A14. 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.

Second, the total utility from information capital and cooperation capital contribute relatively the same amount to an agent's total utility from the network. This can be seen most clearly in Figure A15, The bulk of the distribution of u_i^I and u_i^C lies around the 45-degree line, which is where $u_i^I = u_i^C$.

The calibration sensitivity plot for the new parameter, τ, is shown in the middle-right panel of Figure A14. 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.

A3.2   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 (15) becomes

$$\pi^{I,d} \log DC_i(G^d; q, \lambda, T) + \pi^{C,d} \log \left(d_i(G^d) + \alpha^d d_i^S(G^d) \right)$$
$$> \pi^{I,h} \log DC_i(G^h; q, \lambda, T) + \pi^{C,h} \log \left(d_i(G^h) + \alpha^h d_i^S(G^h) \right) + \varepsilon_i. \qquad (23)$$

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 get a 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. As a result, we want to confirm the main takeaways are robust, although we do not expect all the parameterizations are exactly the same.

We find that the main observations in section 6.5 persist. The log-linear model correctly predicts 68.6% of the migration decisions, which is close to, though slightly below, the accuracy of the model in the text, which is 69.5%. The parameterization plots for $< \lambda, \alpha^d, \alpha^h, \pi^{I,d}, \pi^{C,d}, \pi^{I,h} >$ are shown in Figure A16. 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

$\alpha^h = 10$.

Figure A17a shows the extent to which information capital and cooperation capital contribute to the agent's total utility from the network. Cooperation capital contributes roughly twice as much 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). It's worth to note that it remains the case that when λ is optimally parameterized, the information capital contributes significantly more to total utility than when we remove the possibility for rivalry by setting $\lambda = 0$. This contrast can be seen by comparing the left ($\lambda = 0.5$) and right ($\lambda = 0$) panels of Figure A17. 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.

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;
$\rightarrow$ Calculate *overall daily modal districts*;
$\rightarrow$ Calculate *overall monthly modal districts*;
Step 2 calculate the *hourly modal districts*;
if *tie districts exit* **then**
    if *overall daily modal districts can resolve* **then**
        return the district with larger occurance number;
        else
            if *overall monthly modal districts can resolve* **then**
                return the district with larger occurance number
            end
        end
    end
end
Step 3 calculate the *daily modal districts*;
if *tie districts exit* **then**
    if *overall daily modal districts can resolve* **then**
        return the district with larger occurance number;
        else
            if *overall monthly modal districts can resolve* **then**
                return the district with larger occurance number
            end
        end
    end
end
Step 4 calculate the *monthly modal districts*;
if *tie districts exit* **then**
    if *overall monthly modal districts can resolve* **then**
        return the district with larger occurance number;
    end
end

Algorithm 1: Home location assignment

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

if $D_1 == D_2$ *AND* $D_3 == D_4$ **then**

    if $D_2 == D_3$ **then**

        if $D_4 == Kigali$ **then**

            | migration type is **urban resident**

        end

        else

            | migration type is **rural resident**

        end

    end

    else

        if $D_4 == Kigali$ **then**

            | migration type is **rural to urban**

        end

        else

            if $D_1 == Kigali$ **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
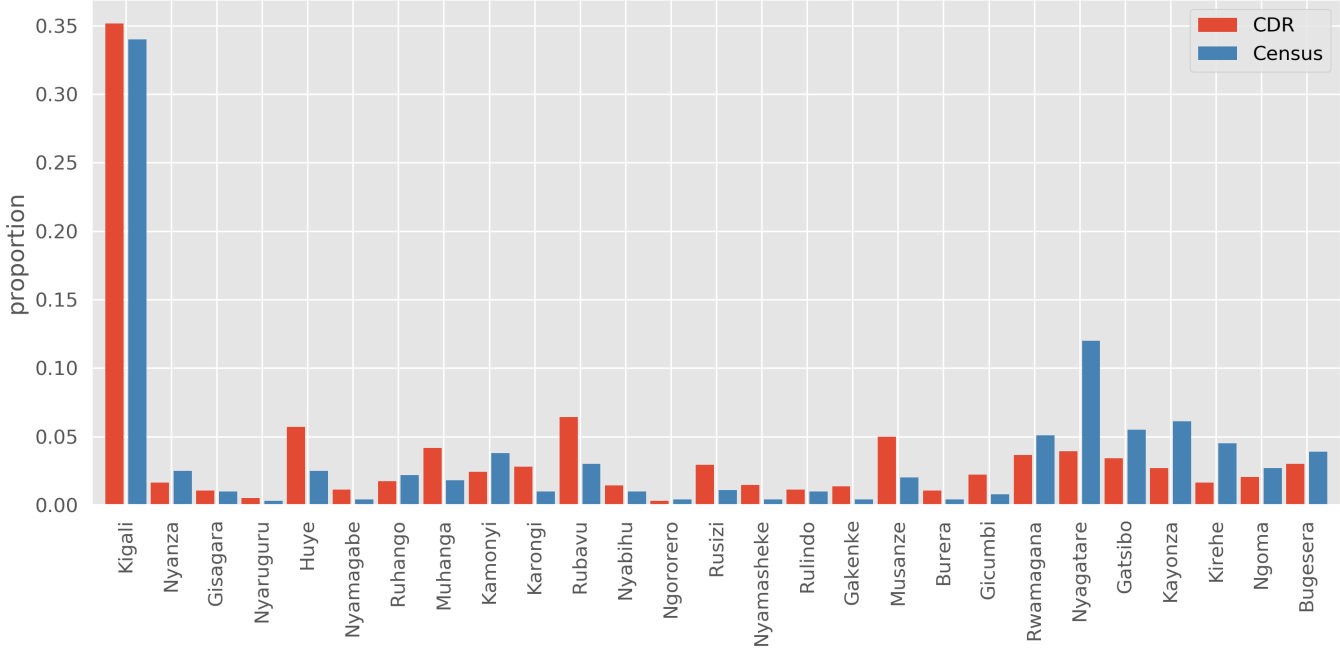
Notes: Figure shows the proportion of migrants to each district in Rwanda. Red bars indicate the proportion inferred from the mobile phone data; Blue bars indicate the proportion calculated from 2012 Rwandan census data, as reported by National Institute of Statistics of Rwanda (2014).

Figure A2: Network structure of migrants

(a) Number of contacts

(b) Number of calls
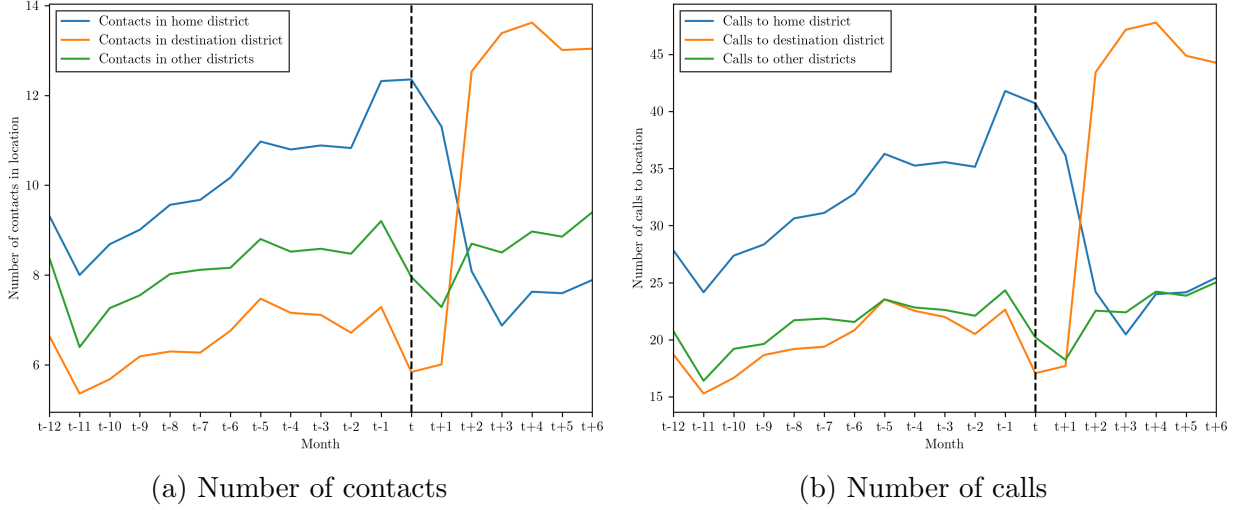
Figure A3: Network structure of non-migrants

(a) Percent of contacts
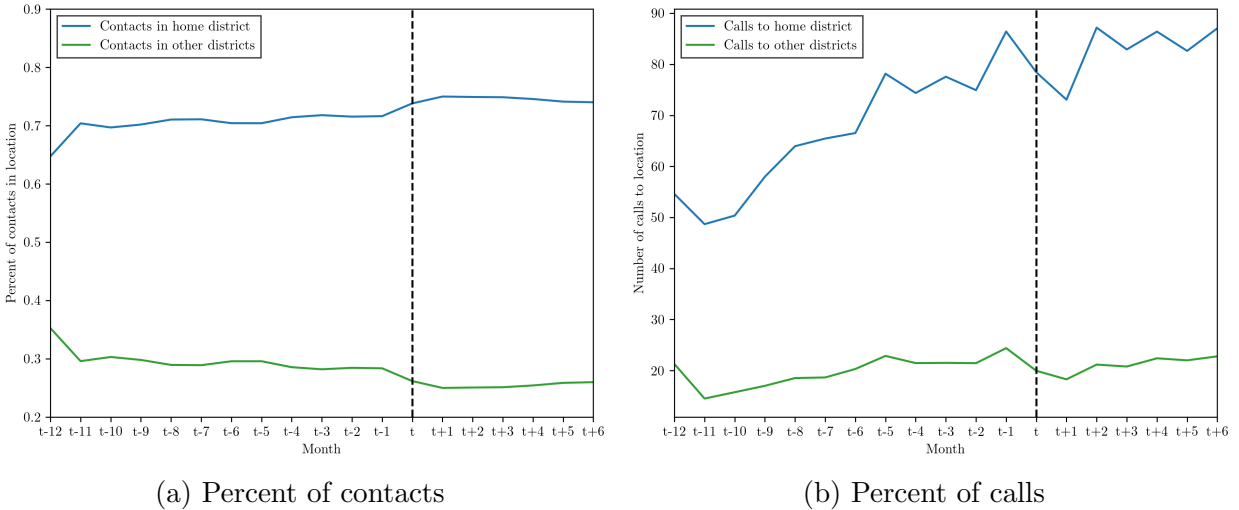
(b) Percent of calls

Notes: Top figures shows how the network connections of migrants evolves over time, in each of the 12 months before and 6 months after migration. These are similar to Figure 4, except that instead of showing the *percent* of calls to each location, Figure plots the *number* of unique contacts in each location and Figure A2b indicates the number of *phone calls* to each location. Bottom figures show equivalent figures for *non*-migrants, as a sort of placebo test. For non-migrants, the index month t is sampled from the same distribution of months in which actual migrations occur).

Figure A4: Number of friends of friends, before and after migration (migrants)
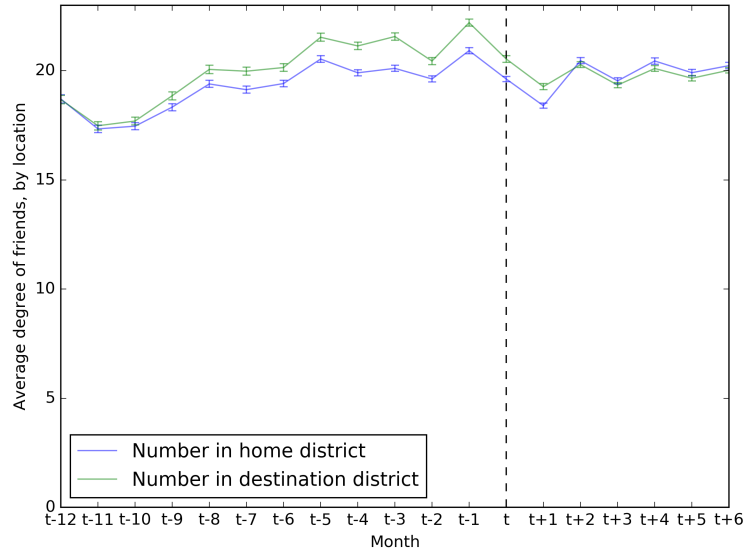
Figure A5: Percent of friends with common support, before and after migration (migrants)
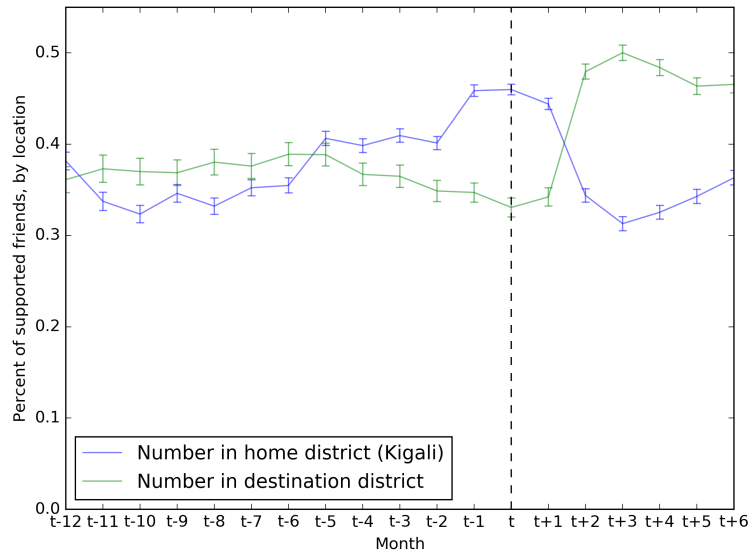
Notes: Top figure shows total number of friends of friends migrants have in their home district and their destination district, in each of the 12 months before and 6 months after migration. Bottom figure shows the percent of the mgirants friends who have a common friend.

Figure A6: Migration rate and degree centrality, controlling for different fixed effects

(a) No fixed effects

(b) Destination district fixed effects

(c) Home-destination-month F.E.'s

(d) Home-dest-month & individual F.E.'s
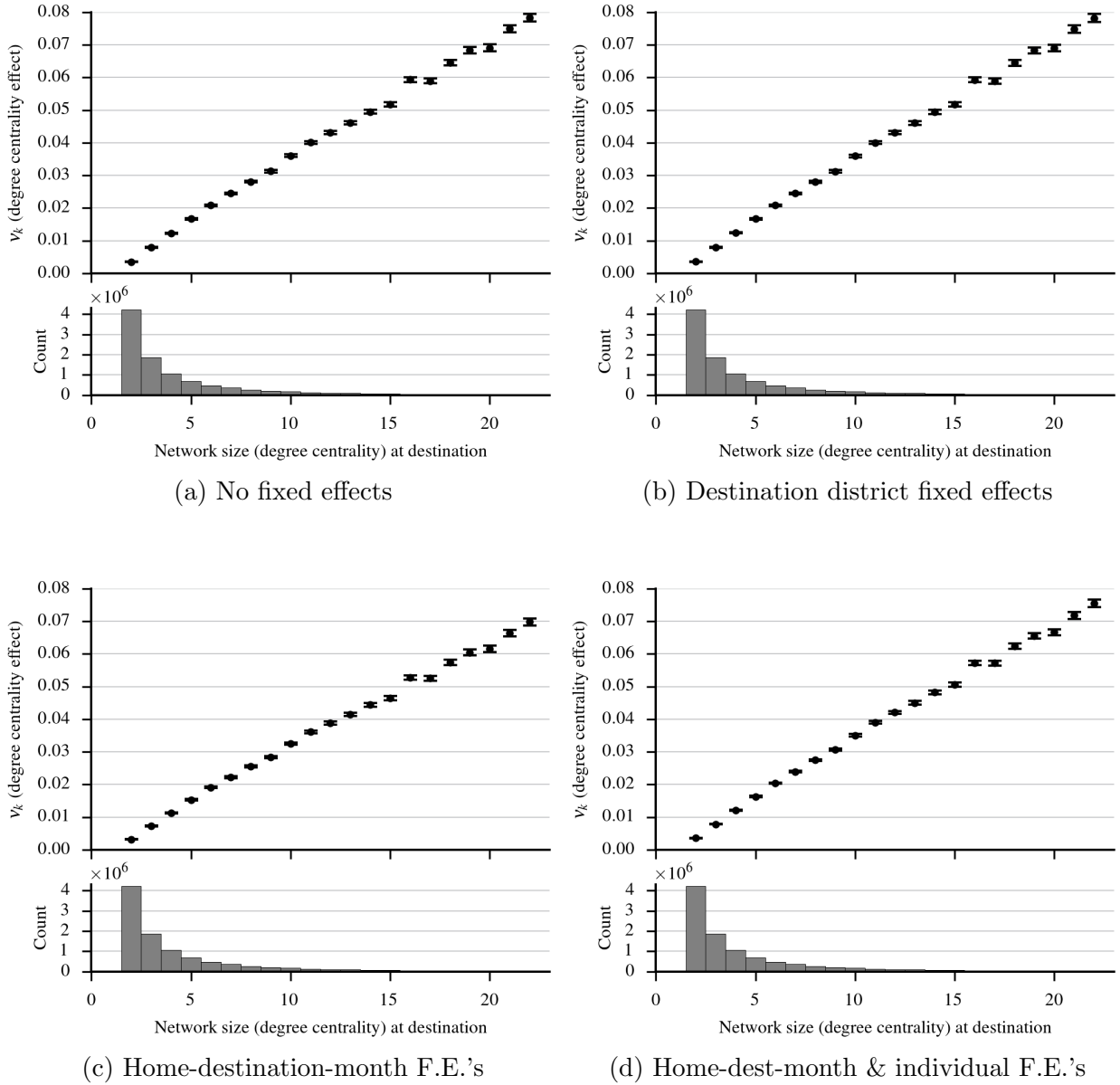
Notes: Each figure shows the fixed effect coefficients estimated from a regression of migration on separate fixed effects for each possible destination network size (see Section 5.1). Figure subtitle indicates any other fixed effects included in the specification. Error bars indicate 95% confidence intervals, clustered by individual.

Figure A7: Relationship between migration rate and clustering
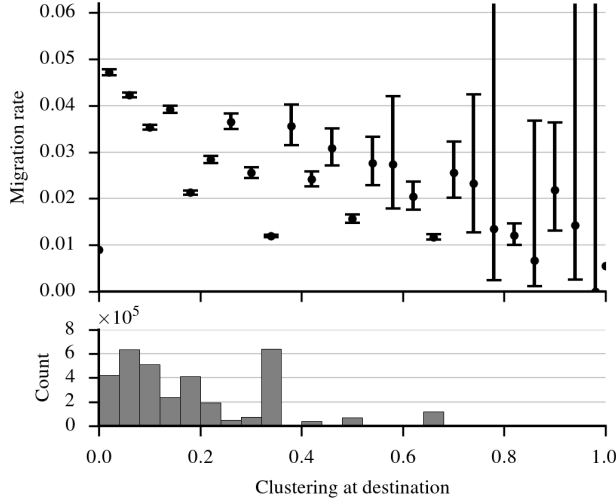
(a) Clustering at Destination
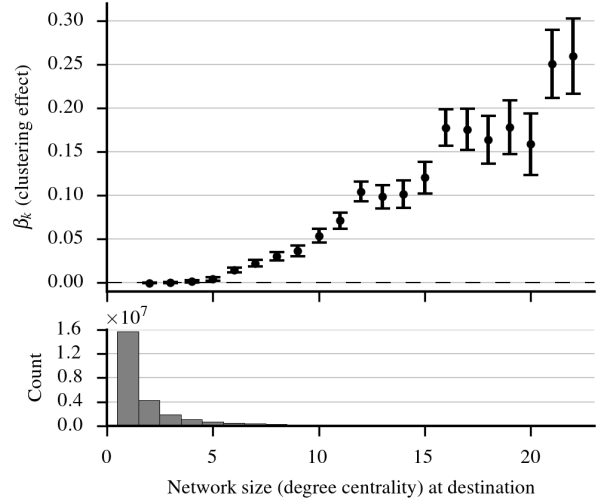
(b) Clustering at Destination, by Degree
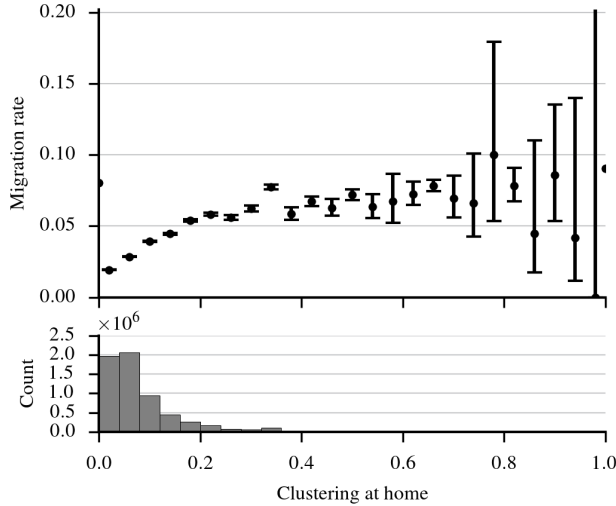
(c) Clustering at Home
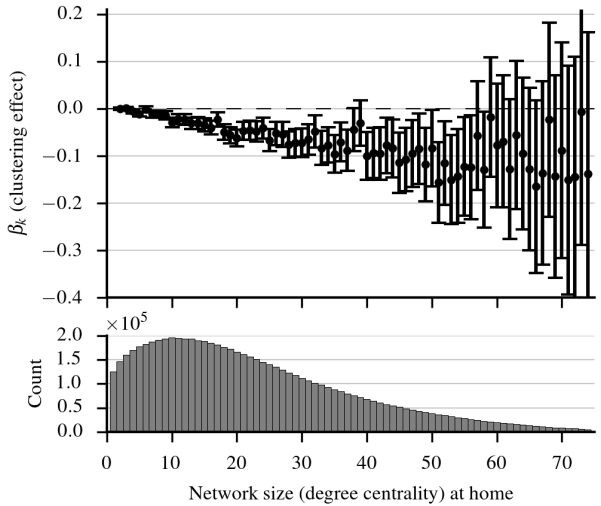
(d) Clustering at Home, by Degree

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

Figure A8: Migration rate and home friends of friend in destination
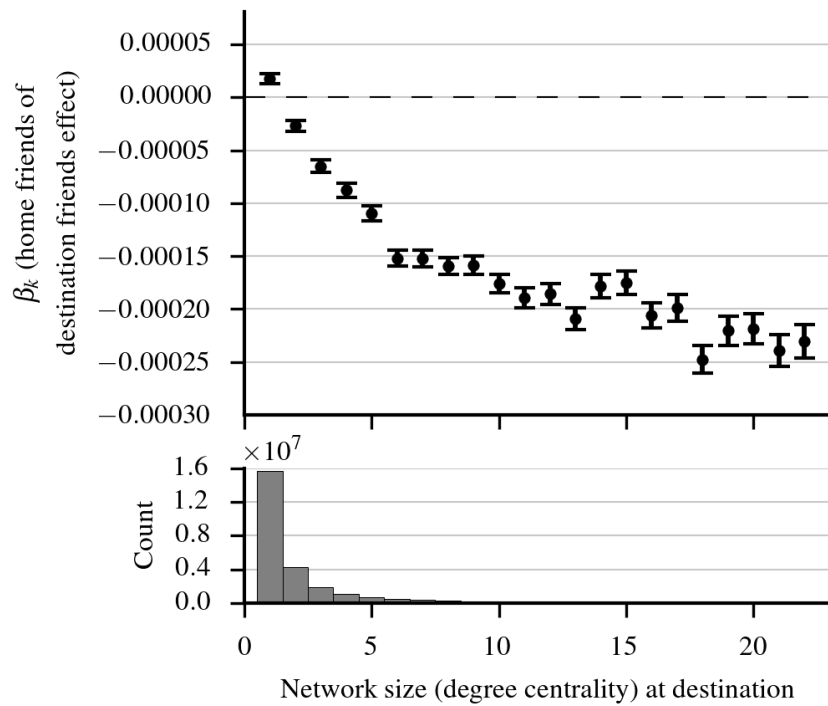

Notes: Figure shows the β_k values estimated with model 4, i.e., the correlation between migration and unique friends (at home) of friends (in the destination) for individuals with different numbers of friends (in the destionation), after conditioning on fixed effects — see Section 5.2. Error bars indicate 95% confidence intervals, clustered by individual.

Figure A9: Migrants have fewer friends of friends than non-migrants
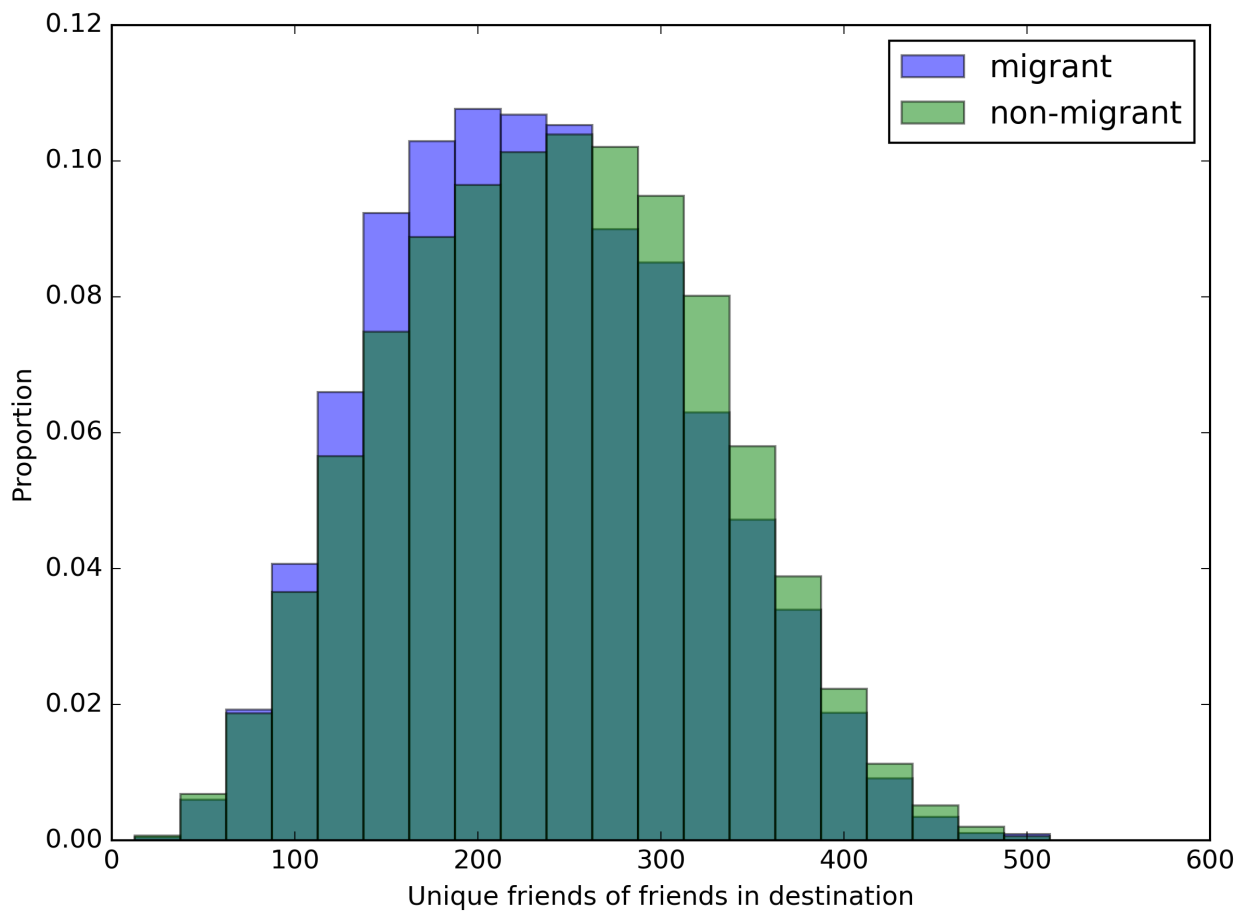
Notes: The figure focuses on all individuals who have exactly 10 unique contacts in a potential destination, and shows the distribution of the number of unique "friends of friends" in that destination. Counterintuitively, migrants have fewer unique friends of friends than non-migrants.

Figure A10: Urban and rural sectors in Rwanda
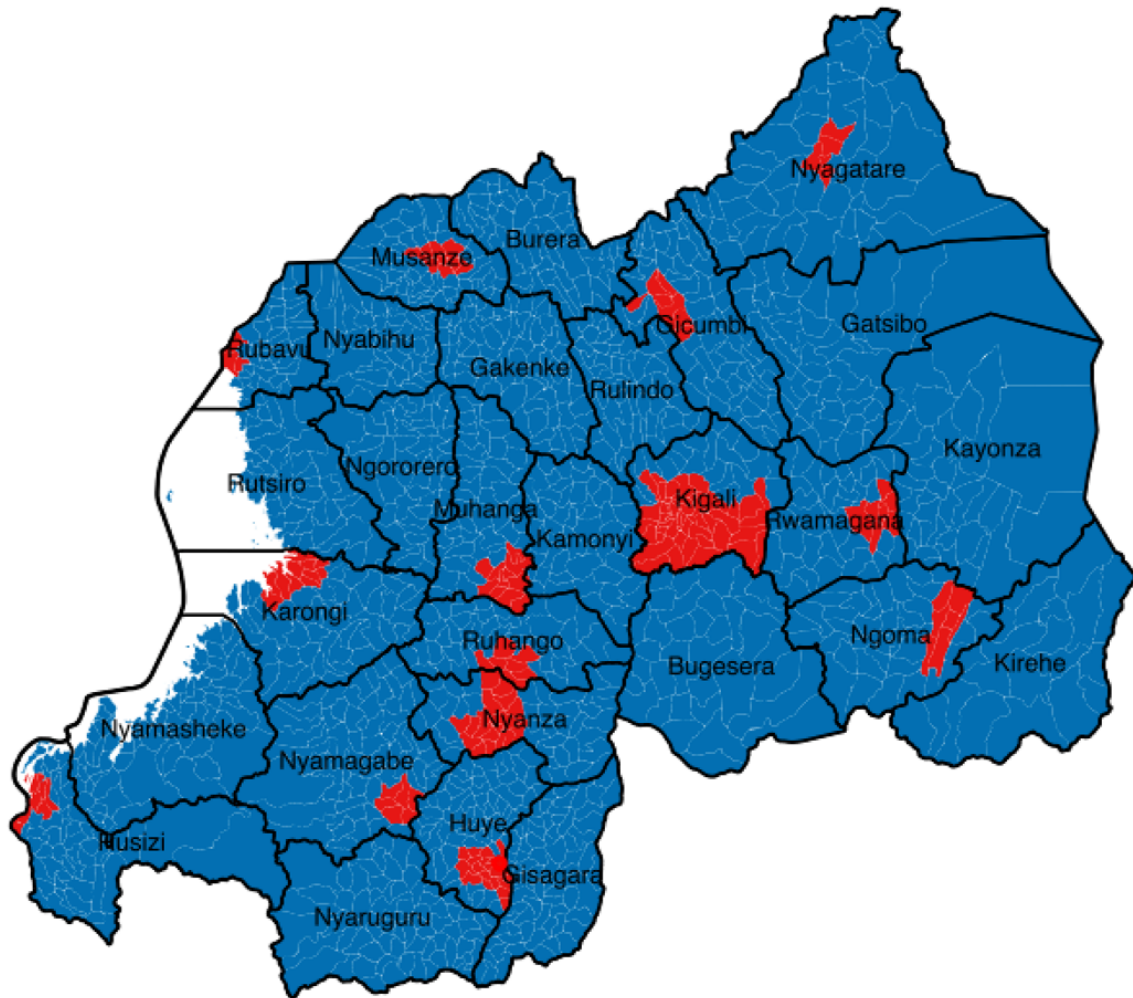
Notes: Urban zones shown in red; rural zones shown in blue. Urban and rural designations detremined using the sector boundary dataset from the website of National Institute of Statistics Rwanda, available from http://statistics.gov.rw/geodata.

Figure A11: Calibration results: marginal plots
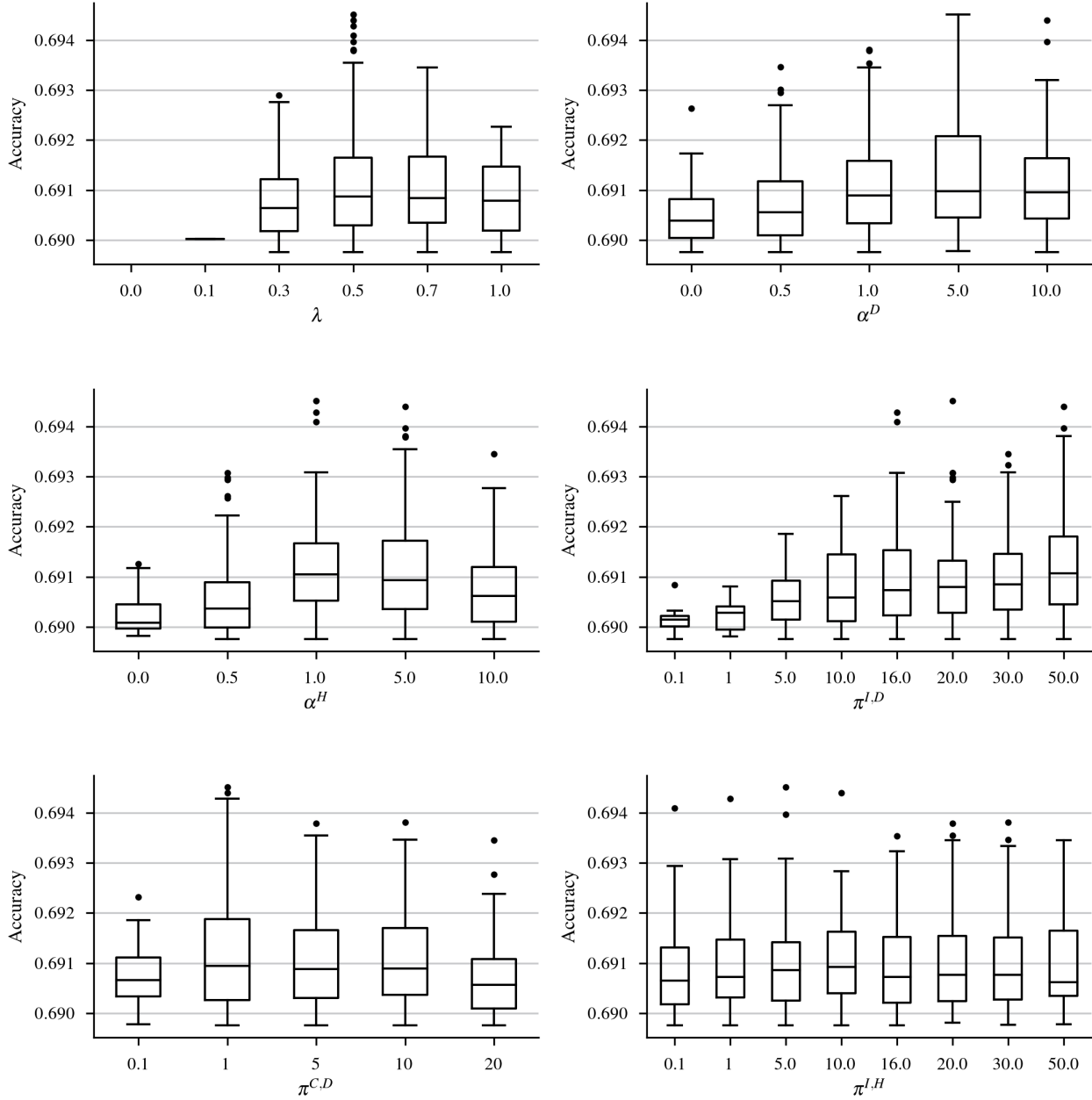
Notes: Figures show the marginal effect of varying λ, α_d, α^h and $(\pi^{I,d}, \pi^{C,d}, \pi^{I,h})$ when calibrating Model 15. Each of roughly 50,000 different parameter combinations is tested; the top percentile of simulations are used to generate this marginal plot.

Figure A12: Simulated balance of home vs. destination utility
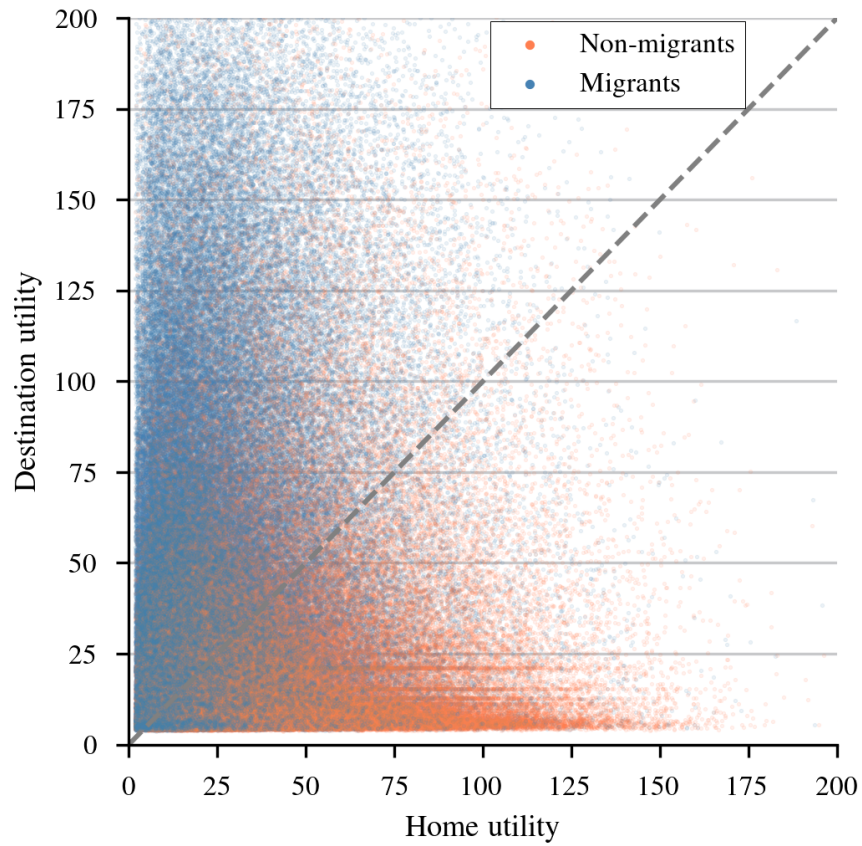
Notes: After the model is calibrated, the optimal parameters are used to calculate the total utility provided to each individual by the home network and destination network. Each dot represents one individual's combination of predicted home-destination utility. Blue (red) dots above (below) the 45-degree line are correctly classified; blue (red) dots below (above) the 45-degree line are incorrectly classified.

Figure A13: Calibration results when $\lambda = 0$: 'information' and 'cooperation' utility
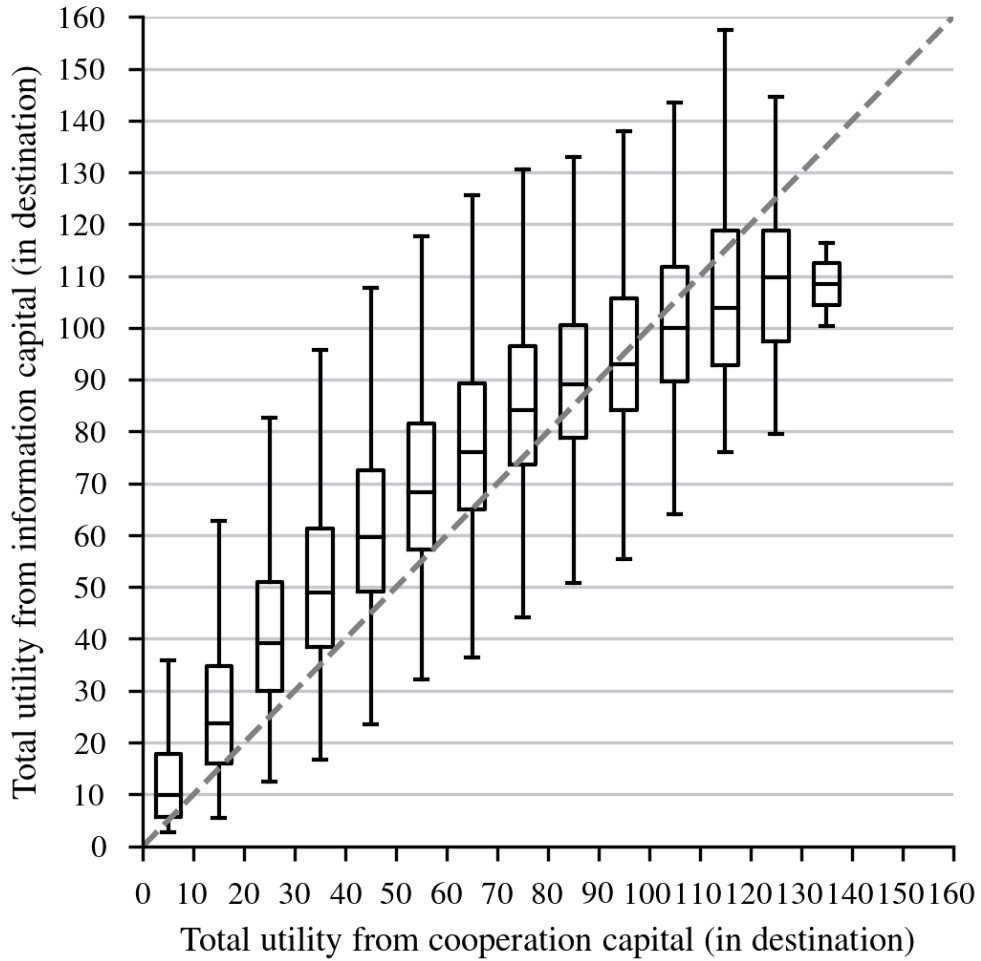
Figure A14: Calibration results (with τ): marginal plots
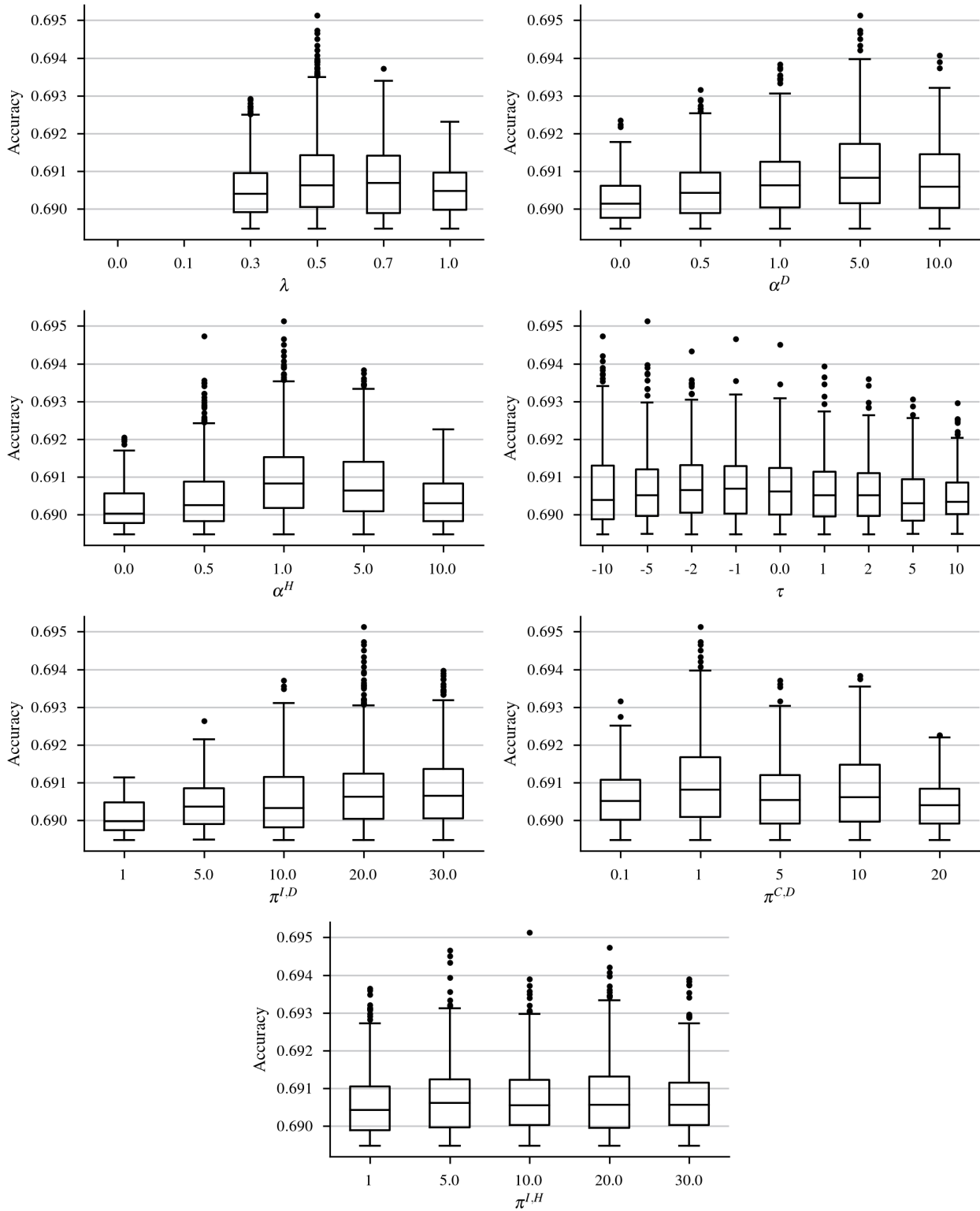
Notes: Figures show the marginal effect of varying λ, α, τ, and π when calibrating Model (22). Each of roughly 50,000 different parameter combinations is tested; the top percentile of simulations are used to generate this marginal plot.

Figure A15: Calibration results (with τ): 'information' and 'cooperation' utility

Notes: Figures show the distribution of predicted utility from 'information' and 'cooperation' (i.e., equation 2) for 270,000 migrants and non-migrants.

Figure A16: Calibration results for log linear model: marginal plots
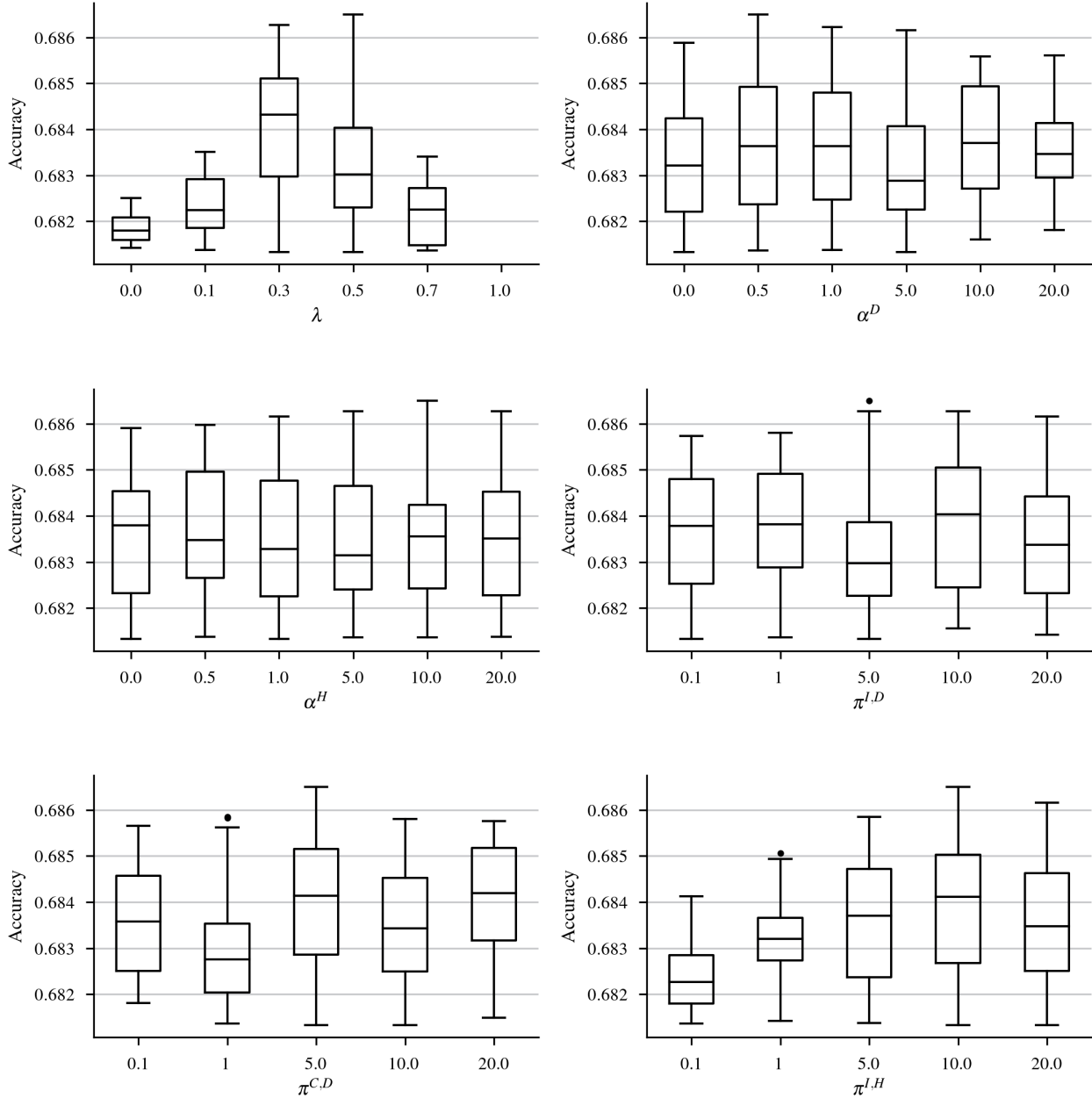
Notes: Figures show the marginal effect of varying λ, α_d, α^h and $(\pi^{I,d}, \pi^{C,d}, \pi^{I,h})$ when calibrating Model (23). Each of roughly 50,000 different parameter combinations is tested; the top percentile of simulations are used to generate this marginal plot.

Figure A17: Calibration results for log linear model: 'information' and 'cooperation' utility

(a) Rivalrous information transmission ($\lambda = 0.5$) (b) Non-rival information transmission ($\lambda = 0$)
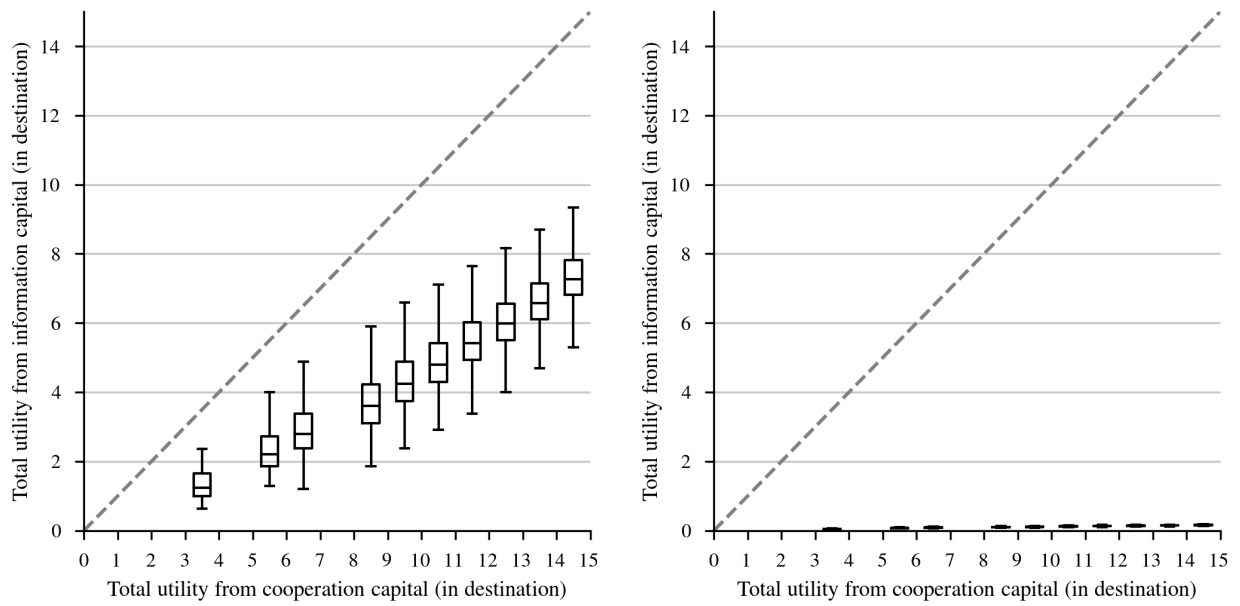
Notes: Figures show the distribution of predicted utility from 'information' and 'cooperation' (i.e., equation 2) for 270,000 migrants and non-migrants. The left figure is calculated using the parameters selected by calibrating Model 23. For the right figure, λ is fixed at zero (i.e., no information rivalry).

Table A1: Migration events observed in 4.5 years of phone data

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	935,806	34.565	11.171	21.923	23.181	18.457
2	680,267	21.634	1.933	8.244	13.828	5.934
3	518,156	13.960	0.405	2.893	9.216	2.007
6	263,182	5.294	0.000	0.192	3.547	0.128

Notes: Table counts number of unique individuals meeting different definitions of a "migration event." 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 text for details. 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. The number of individual (N) varies by row, since an individual is only considered eligible as a migrant if she is observed continuously over $2N$ consecutive months.

Table A2: Jointly estimated effects of home and destination network structure

	(1)	(2)	(3)
Destination Degree (network size)	0.0048033***	0.0037637***	
	(0.0000201)	(0.0000238)	
Home Degree (network size)	−0.0007377***	−0.0005089***	
	(0.0000060)	(0.0000107)	
Destination friends of friends	−0.0000324***	−0.0000059***	−0.0000001
	(0.0000007)	(0.0000009)	(0.0000009)
Home friends of friends	0.0000113***	0.0000059***	−0.0000035***
	(0.0000002)	(0.0000004)	(0.0000004)
Destination % friends with support	0.0037855***	0.0017164***	0.0010618***
	(0.0001088)	(0.0001130)	(0.0001146)
Home % friends with support	0.0081299***	−0.0061902***	0.0002216
	(0.0001336)	(0.0002305)	(0.0002407)
Observations	9,889,981	9,889,981	9,889,981
R^2	0.0213936	0.1858886	0.1868505
Degree fixed effects	No	No	Yes
Home*Destination*Month fixed effects	No	Yes	Yes
Individual fixed effects	No	Yes	Yes

Notes: Standard errors are two-way clustered by individual and by home-destination-month. *p<0.1; **p<0.05; ***p<0.01.

Table A3: Robustness to alternative fixed effect specifications

	(1)	(2)	(3)	(4)
Panel A: Destination network characteristics				
Degree (network size)	0.0036548***			
	(0.0000183)			
Friends of friends	−0.0000103***	−0.0000160***	−0.00000004	−0.0000002
	(0.0000007)	(0.0000007)	(0.0000008)	(0.0000009)
% Friends with common support	0.0010869***	0.0022076***	0.0028977***	0.0014808***
	(0.0001045)	(0.0001107)	(0.0001112)	(0.0001146)
Observations	9,889,981	9,889,981	9,889,981	9,889,981
Panel B: Home network characteristics				
Degree (network size)	−0.0003957***			
	(0.0000060)			
Friends of friends	0.0000021***	−0.0000109***	−0.0000165***	−0.0000110***
	(0.0000002)	(0.0000001)	(0.0000001)	(0.0000002)
% Friends with common support	0.0325365***	−0.0186718***	−0.0139236***	−0.0087495***
	(0.0001233)	(0.0001673)	(0.0001731)	(0.0002245)
Observations	9,889,981	9,889,981	9,889,981	9,889,981
Degree fixed effects	No	Yes	Yes	Yes
Home*Destination*Month fixed effects	No	No	Yes	Yes
Individual fixed effects	No	No	No	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. Standard errors are two-way clustered by individual and by home-destination-month. *p<0.1; **p<0.05; ***p<0.01.

Table A4: Robustness to alternative fixed effect specifications, part 2

	(1)	(2)	(3)	(4)	(5)
Destination friends of friends	-0.0000002	0.0000011	-0.0000064^{***}	-0.0000077^{***}	-0.0000028^{***}
	(0.0000009)	(0.0000011)	(0.0000010)	(0.0000012)	(0.0000010)
% Destination friends with support	0.0014808^{***}	0.0013719^{***}	0.0003458^{***}	0.0006663^{***}	0.0001123
	(0.0001146)	(0.0001491)	(0.0001220)	(0.0000966)	(0.0001204)
Observations	$9,889,981$	$9,889,981$	$9,889,981$	$9,889,981$	$9,889,981$
R^2	0.1853017	0.5080845	0.5952072	0.6680641	0.6332967
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. *p<0.1; **p<0.05; ***p<0.01.

Table A5: Conditional logit results

	(1)	(2)	(3)	(4)
Destination Degree (network size)	0.16427***	0.308192***	0.11818***	0.211611***
	(0.00106)	(0.002854)	(0.00114)	(0.003034)
Home Degree (network size)	-0.11931***	-0.261790***	-0.07906	-0.188931***
	(0.00114)	(0.002980)	(0.00128)	(0.003160)
Destination friends of friends		-0.005564***		-0.003503***
		(0.000108)		(0.000108)
Home friends of friends		-0.005442***		0.004055***
		(0.000112)		(0.000110)
Destination % friends with support			2.49114***	2.241620***
			(0.02788)	(0.030131)
Home % friends with support			-1.90396***	-1.57135***
			(0.01924)	(0.042690)
Home choice	6.10215***	6.114159***	6.10313***	6.082535***
	(0.01493)	(0.01514)	(0.01824)	(0.01813)
McFadden R^2	0.88563	0.88709	0.88864	0.88936
N individuals	433,782	433,782	433,782	433,782

Notes: Response variable in conditional logit is a dummy variable indicating whether inidividual i migrates from district h to district d in Jaunary 2008. Each choice represents one of the 27 districts in Rwanda (the three smaller urban districts in Kigali province are treated as a single district). Standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01.

Table A6: Heterogeneity by Migration Frequency (Repeat and First-time)

Migration Frequency	(1) Any	(2) Repeat	(3) First-Time
Destination friends of friends	−0.0000001	0.0000171***	−0.0000030***
	(0.0000009)	(0.0000062)	(0.0000008)
Home friends of friends	−0.0000035***	−0.0000511***	0.0000022***
	(0.0000004)	(0.0000043)	(0.0000003)
% Destination support	0.0010618***	−0.0027428*	0.0010934***
	(0.0001146)	(0.0014071)	(0.0000920)
% Home support	0.0002216	0.0037889**	−0.0007294***
	(0.0002407)	(0.0018547)	(0.0001994)
Observations	9,889,981	665,780	9,224,201
R^2	0.1868505	0.4382679	0.1986143
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. *p<0.1; **p<0.05; ***p<0.01.

Table A7: Heterogeneity by Distance (Adjacent districts vs. Non-adjacent districts)

Migration Distance	(1) Any	(2) Short Distance (adjacent districts)	(3) Long-Distance (non-adjacent districts)
Destination friends of friends	−0.0000001	0.0000042**	−0.0000159***
	(0.0000009)	(0.0000017)	(0.0000012)
Home friends of friends	−0.0000035***	−0.0000052***	−0.0000028***
	(0.0000004)	(0.0000008)	(0.0000005)
% Destination support	0.0010618***	0.0010032***	0.0010780***
	(0.0001146)	(0.0002282)	(0.0001362)
% Home support	0.0002216	−0.0004295	0.0002990
	(0.0002407)	(0.0004260)	(0.0002933)
Observations	9,889,981	3,337,184	6,552,797
R^2	0.1868505	0.3237450	0.1972246
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. *p<0.1; **p<0.05; ***p<0.01.

Table A8: Heterogeneity by Migration Duration (Long-term vs. Short-term)

Migration Distance	(1) Any	(2) Long Stay (> 12 months)	(3) Short Stay (< 6 months)
Destination friends of friends	−0.0000001	0.0000156***	−0.0000125***
	(0.0000009)	(0.0000005)	(0.0000007)
Home friends of friends	−0.0000035***	−0.0000068***	0.0000007**
	(0.0000004)	(0.0000002)	(0.0000003)
% Destination "support"	0.0010618***	0.0002180***	0.0008051***
	(0.0001146)	(0.0000626)	(0.0000846)
% Home "support"	0.0002216	0.0000928	0.0001442
	(0.0002407)	(0.0001323)	(0.0001786)
Observations	9,889,981	9,782,384	9,820,778
R^2	0.1868505	0.1445434	0.1857658
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. *p<0.1; **p<0.05; ***p<0.01.

Table A9: Heterogeneity by destination type (Rural and Urban)

Destination Type	(1) All	(2) Rural	(3) Urban
Destination friends of friends	−0.0000001	0.0000022	−0.0000019
	(0.0000009)	(0.0000020)	(0.0000012)
Home friends of friends	−0.0000035***	−0.0000037***	−0.0000018***
	(0.0000004)	(0.0000006)	(0.0000006)
% Destination "Support"	0.0010618***	0.0009579***	0.0008771***
	(0.0001146)	(0.0001470)	(0.0001612)
% Home "Support"	0.0002216	−0.0002734	0.0002481
	(0.0002407)	(0.0003254)	(0.0003042)
Observations	9,889,981	4,236,638	5,918,664
R^2	0.1868505	0.3103749	0.2471896
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. The three districts that comprise th capital of Kigali are denoted as urban and the remaining districts are denoted as rural (see Table A10 for an alternative definition of urban and rural locations). Standard errors are two-way clustered by individual and by home-destination-month. *p<0.1; **p<0.05; ***p<0.01.

Table A10: Heterogeneity by destination type (Rural and Urban), using alternative definition of urban and rural areas

Destination Type	(1) All	(2) Rural	(3) Urban
Destination friends of friends	−0.0000001	0.0000030	−0.0000024**
	(0.0000009)	(0.0000020)	(0.0000012)
Home friends of friends	−0.0000035***	−0.0000034***	−0.0000017***
	(0.0000004)	(0.0000006)	(0.0000006)
% Destination "Support"	0.0010618***	0.0009944***	0.0009398***
	(0.0001146)	(0.0001472)	(0.0001610)
% Home "Support"	0.0002216	−0.0003122	0.0002904
	(0.0002407)	(0.0003260)	(0.0003043)
Observations	9,889,981	4,230,528	5,924,177
R^2	0.1868505	0.3101766	0.2464579
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. Urban and rural designation detremined using the sector boundary dataset from the website of National Institute of Statistics Rwanda (see Figure A10). Standard errors are two-way clustered by individual and by home-destination-month. *p<0.1; **p<0.05; ***p<0.01.

Table A11: The role of strong ties and weak ties

	(1)	(2)	(3)	(4)
Destination "Weak tie"	0.0036077***	0.0037190***	0.0036771***	0.0037849***
	(0.0000123)	(0.0000250)	(0.0000107)	(0.0000240)
Destination "Strong tie"	0.0044319***	0.0045117***	0.0044074***	0.0045034***
	(0.0000495)	(0.0000536)	(0.0001536)	(0.0001549)
Home "Weak tie"	−0.0003855***	−0.0004813***	−0.0004042***	−0.0005021***
	(0.0000050)	(0.0000108)	(0.0000049)	(0.0000107)
Home "Strong tie"	−0.0007742***	−0.0008799***	−0.0014034***	−0.0015449***
	(0.0000152)	(0.0000179)	(0.0000755)	(0.0000761)
Destination friends of friends		−0.0000062***		−0.0000061***
		(0.0000009)		(0.0000009)
Home friends of friends		0.0000058***		0.0000059***
		(0.0000004)		(0.0000004)
% Destination "Support"		0.0018786***		0.0018158***
		(0.0001138)		(0.0001133)
% Home "Support"		−0.0061352***		−0.0061689***
		(0.0002306)		(0.0002305)
Observations	9,889,981	9,889,981	9,889,981	9,889,981
R²	0.1858262	0.1859473	0.1857898	0.1859106
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 4 use a threshold of 12 calls (the 95th percentile of tie strength). Standard errors are two-way clustered by individual and by home-destination-month. *p<0.1; **p<0.05; ***p<0.01.

Table A12: 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)	0.0000004					
	(0.0000009)					
Friends of friends (strong-strong)		0.0000175*				−0.0002288***
		(0.0000104)				(0.0000202)
Friends of friends (strong-weak)			0.0000226***			0.0000696***
			(0.0000024)			(0.0000047)
Friends of friends (weak-strong)				−0.0000460***		−0.0001103***
				(0.0000048)		(0.0000072)
Friends of friends (weak-weak)					0.0000016	0.0000224***
					(0.0000011)	(0.0000017)
Observations	10,089,959	10,089,959	10,089,959	10,089,959	10,089,959	10,089,959
R²	0.1908962	0.1908965	0.1909039	0.1909041	0.1908964	0.1909380

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. *p<0.1; **p<0.05; ***p<0.01

Table A13: Disaggregating the network support effect by the strength of supported ties

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Support (all)	0.0013211***									
	(0.0001172)									
Support (sss)		0.0016200**								0.0025375***
		(0.0006504)								(0.0006531)
Support (sws)			0.0069022***							0.0076857***
			(0.0006493)							(0.0006511)
Support (ssw)				0.0006118***						−0.0000068
				(0.0001283)						(0.0001467)
Support (sww)					0.0027148***					0.0030284***
					(0.0001003)					(0.0001102)
Support (wss)						−0.0005028				−0.0005472
						(0.0003166)				(0.0003539)
Support (wws)							0.0009461***			−0.0025847***
							(0.0003165)			(0.0003400)
Support (wsw)								−0.0019032***		−0.0019038***
								(0.0003918)		(0.0003928)
Support (www)									0.0015830***	0.0012720***
									(0.0003317)	(0.0003326)
Strong tie	0.0013544***	0.0013941***	0.0013132***	0.0013981***	0.0013135***	0.0014268***	0.0013923***	0.0014085***	0.0014242***	0.0012280***
	(0.0000477)	(0.0000482)	(0.0000483)	(0.0000475)	(0.0000475)	(0.0000478)	(0.0000480)	(0.0000474)	(0.0000474)	(0.0000503)
Observations	10,089,959	10,089,959	10,089,959	10,089,959	10,089,959	10,089,959	10,089,959	10,089,959	10,089,959	10,089,959
R²	0.1909840	0.1909736	0.1909828	0.1909751	0.1910361	0.1909733	0.1909739	0.1909751	0.1909751	0.1910568

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 mont t. We show the Destination network "support" coefficient separately for geometries of different tie strengths. "SSS" (column 2) indicates the effect of network suppo 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 nomenclatu Strong ties are defined as relationships with 5 or more phone calls (the 90th percentile of tie strength) in a given month. *p<0.1; **p<0.05; ***p<0.01

90

Table A14: The role of recent migrants

	(1)	(2)	(3)
Destination Degree (network size)	0.0037637***	0.0036358***	0.0036513***
	(0.0000238)	(0.0000244)	(0.0000238)
Home Degree (network size)	−0.0005089***	−0.0005171***	−0.0005859***
	(0.0000107)	(0.0000107)	(0.0000107)
Destination friends of friends	−0.0000059***	−0.0000041***	−0.0000060***
	(0.0000009)	(0.0000009)	(0.0000009)
Home friends of friends	0.0000059***	0.0000060***	0.0000075***
	(0.0000004)	(0.0000004)	(0.0000004)
% Destination "Support"	0.0017164***	0.0017326***	0.0017847***
	(0.0001130)	(0.0001130)	(0.0001129)
% Home "Support"	−0.0061902***	−0.0061607***	−0.0063159***
	(0.0002305)	(0.0002305)	(0.0002304)
Recent migrant friends		0.0011090***	0.0126456***
		(0.0000489)	(0.0001135)
Observations	9,889,981	9,889,981	9,889,981
R^2	0.1858886	0.1859340	0.1869832
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 A2; 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. *p<0.1; **p<0.05; ***p<0.01.

Table A15: *Predicted* migration (from structural model) and social network structure

	(1)	(2)	(3)	(4)
Panel A: Destination network characteristics				
Degree (network size)	0.0680931***			
	(0.0000450)			
% Friends with common support		0.1728557***		0.1707765***
		(0.0004015)		(0.0004002)
Unique friends of friends			−0.0007402***	−0.0007033***
			(0.0000035)	(0.0000034)
Observations	6,386,523	6,386,523	6,386,523	6,386,523
R^2	0.5967755	0.6359449	0.6271628	0.6386054
Panel B: Home network characteristics				
Degree (network size)	−0.0114922***			
	(0.0000197)			
% Friends with common support		−0.1836519***		−0.1846382***
		(0.0010150)		(0.0010159)
Unique friends of friends			−0.0000240***	−0.0000364***
			(0.0000016)	(0.0000016)
Observations	6,386,523	6,386,523	6,386,523	6,386,523
R^2	0.4676148	0.4948318	0.4919757	0.4948771
Degree fixed effects	No	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes
Home*Destination*Month F.E.	Yes	Yes	Yes	Yes

Notes: Each column indicates a separate regression of a binary variable $\widehat{M_{ihdt}}$ that takes the value 1 if an individual i was *predicted* to migrate from home district h to destination district d in month t (where this prediction is based on the calibrated structural model, and determined using the actual network properties of i). Standard errors are two-way clustered by individual and by home-destination-month. *p<0.1; **p<0.05; ***p<0.01.