



G²LM|LIC Working Paper No. 60 | January 2022

Migration and the Value of Social Networks

Joshua Blumenstock (U.C. Berkeley) Guanghua Chi (U.C. Berkeley) Xu Tan (University of Washington)









G²LM|LIC Working Paper No. 60 | January 2022

Migration and the Value of Social Networks

Joshua Blumenstock (U.C. Berkeley) Guanghua Chi (U.C. Berkeley) Xu Tan (University of Washington)

G²LM|LIC

c/o IZA – Institute of Labor Economics Schaumburg-Lippe-Straße 5–9 53113 Bonn, Germany Phone: +49-228-3894-0 Fax: +49-228-3894-510 Email: glm-review@iza.org







G²LM|LIC Working Paper No. 60 | January 2022

ABSTRACT Migration and the Value of Social Networks*

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

JEL Classification:

O15, R23, D85, Z13, O12, C55

Keywords:

networks, migration, social networks, social capital, big data, development

Corresponding author:

Joshua Blumenstock U.C. Berkeley E-Mail: jblumenstock@berkeley.edu

^{*} We thank Oriana Bandiera, Lori Beaman, Francis Bloch, Margherita Comola, Ben Faber, Marcel Fafchamps, Andrew Foster, Matthew Jackson, David Miller, Kaivan Munshi, Mark Rosenzweig, Sebastian Sotelo, Yves Zenou, as well as seminar participants at Stanford, Brown, Paris School of Economics, U.C. Berkeley, the BREAD/CEPR Conference on Development Economics, the NSF Conference on Network Science in Economics, the Economic Demography Workshop, and the Barcelona GSE Summer Forum for helpful feedback. We are grateful for financial support from GLM-LIC Grant C2-RA4-205 and UW RRF #65-7397. This research was approved by the U.C. Berkeley Committee for Protection of Human Subjects, protocol #042196-001.

1 Introduction

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

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

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

¹Classic papers documenting the 'prevailing' view 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). Papers that highlight the potential deterrent effect of larger networks include Calvó-Armengol (2004), Calvó-Armengol and Jackson (2004), Wahba and Zenou (2005) and Beaman (2012).

²For 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

approximation of the size of a migrant's social network, but do not reveal how more nuanced network structures influence the migration decision. Higher-order network structure plays a critical role in decisions about employment, education, health, finance, product adoption, and the formation of strategic alliances.³ Yet, the role of such network structure in migration has not been systematically studied.

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

The empirical analysis links each individual's migration decisions over time to the evolving structure of their social network. For instance, we use these data to confirm the longstanding hypothesis that people move to places where they know more people; conversely, individuals are less likely to leave places where they have larger networks. While these results are expected, an advantage of our setting is that we can observe exactly how migration likelihood varies with network size. The relationship is monotonic and approximately linear, such that the probability of migration roughly doubles as the number of contacts in the destination doubles. Superficially, this result diverges from a series of studies that predict eventual negative externalities from network size, as when members compete for information and opportunities (Calvó-Armengol, 2004, Calvó-Armengol and Jackson, 2004, Beaman, 2012, Dagnelie, Mayda and Maystadt, 2019).

We then focus on developing a systematic understanding of how *higher-order* network structure — i.e., the connections of the migrant's connections (and their connections' connections, and so forth) — influences the decision to migrate. The purpose of this analysis is to understand whether, ceteris paribus, individuals are more likely to migrate to places

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

³For example: Granovetter (1973), Burt (1992), and Karlan et al. (2009) provide examples of how higherorder 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.

where their social networks have particular network topologies. A stylized version of our approach is shown in Figure 1: we are interested in understanding whether, for instance, individual A is more likely to migrate than individual B, where both A and B know exactly two people in the destination and three people at home, and the only observable difference between A and B is that B's contacts are connected to each other whereas A's contacts are from two disjoint communities.

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

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

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

To better understand this 'surprising' result, we document considerable heterogeneity in the migration response to social network structure. In particular, we find that the negative effect of extensive networks is 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 competition for attention (as in Dunbar, 1998, Banerjee et al., 2012). We also find that while the *average* migrant is not drawn to locations where her friends have more friends (as in G_3), such structure does attract several less common types of migrants. In particular, repeat migrants (who have previously migrated from their home to the destination), longterm 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 extensive.

To summarize, this paper makes two main contributions. First, it provides a new empirical perspective on the determinants of migration in developing countries (cf. Lucas, 2015). In this literature, many scholars have noted the important role that social networks play in facilitating migration. Early examples in the economics literature include Rees (1966) and Greenwood (1969); a large number of subsequent studies document the empirical relationship between network size and migration rates.⁵ More recently, Munshi and Rosenzweig

⁵Examples include Montgomery (1991), Borjas, Bronars and Trejo (1992), Munshi (2003), McKenzie and Rapoport (2010), Dolfin and Genicot (2010), Beaman (2012), Patel and Vella (2012), Bertoli, Fernández-Huertas Moraga and Ortega (2013), Bertoli and Ruyssen (2018). Two recent papers use phone data to link spatial mobility and social networks. Büchel et al. (2020) use data from a Swiss cellphone operator to link migration decisions to phone calls, and document a similar relationship between network size and migration as the one we note in Section 4.1. Barwick et al. (2019) show that migrant flows in a Chinese city correlate

(2016) document that the fear of losing social network ties may prevent profitable migration, while Morten (2019) shows that the act of migration can change social relationships and risk sharing. Kinnan (2019) theorizes about the two-way inter-connections: migration of one individual can make other network members better off if that individual has a new source of income, but others may be worse off if the act of migration improves the outside opportunity for that person or makes it easier to hide income. This paper builds on this line of work by exploiting a new source of data to establish a more nuanced set of stylized facts about networks and migration — highlighting, in particular, the value migrants place on interconnected networks, and substantial heterogeneity in how different types of migrants value networks.

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

2 Data and Measurement

To study the empirical relationship between networks and migration, we exploit a novel source of data that contains detailed information on the migration histories and evolving social networks of roughly one million individuals in Rwanda. These data capture all mobile

with call volume between regions, and link this information flow to improved labor market outcomes. Both papers focus primarily on how on network size relates to migration, whereas our focus is on the role of higher-order network structure, conditional on network size.

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

⁷Jackson, 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).

⁸Early 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).

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.

To provide intuition, the network of a single migrant, in the month before migration, is shown in Figure 2. This particular migrant (the green dot) had 20 unique contacts in the month prior to migration, 7 of whom were in his home district (blue dots), four of whom were in the destination district (red dots), and the remainder were in other districts (grey dots). Friends of friends are depicted as hollow grey circles.⁹ This section describes how we use these data to observe the structure of each individual's social network over time (Section 2.1) and to extract each individual's complete migration history (Section 2.2). Section 2.3 discusses limitations of these data.¹⁰

2.1 Modeling social network structure with mobile phone data

A central goal of network theory is to understand how the structure of a social network affects the social capital that an agent obtains from that network. In Appendix A1, we develop a model of how an individual's social capital relates to the *topology* of the network, i.e., to the configuration of connections between individuals. The model highlights two distinct types of social capital, also emphasized in the migration literature, which influence the decision to migrate: *information capital* and *cooperation capital*.

Information 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. Jackson (2020) describes this as the "ability to acquire valuable information and/or spread it to other people through social connections" (p.4). The network's ability to transmit information is closely associated with specific network topologies. In particular, efficient information gathering typically requires an *extensive* subnetwork such that one person is linked to many others via short network paths (cf. Granovetter, 1973). For instance, Jackson and Wolinsky (1996) provide an early measure of information capital as decay centrality, where each agent

⁹Throughout, 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.

¹⁰All 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 transactions involving numbers with more than 200 unique contacts in a single month (this represents the 95th percentile).

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. 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 individual to other.

In both the decay and diffusion centrality measures, information capital increases with more friends, friends of friends, friends of friends of friends, and so on. In subsequent analysis, we study how migration decisions relate to the size of an individual's second-degree neighborhood (or unique *friends of friends*, not counting direct contacts of the individual) and third-degree neighborhood (unique friends of friends of friends, excluding first- and second-degree connections). Formally, let the network G be a matrix with $G_{ij} = 1$ if i and j are connected and otherwise $G_{ij} = 0$ (including $G_{ii} = 0$). A path between i and j is an ordered sequence of agents $(ii_1i_2...i_hj)$ such that any two adjacent agents are connected. The distance between i and j, denoted as d(i, j), is the length of the shortest path between i and j.¹¹ Then, agent i's k^{th} -degree neighborhood is defined as:

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

Cooperation capital. 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. This corresponds closely to the notion of *favor capital* in Jackson (2020), defined as the network's "ability to exchange favors and transact with others through network position and repeated interaction and reciprocation" (p.315).

Cooperation capital is linked to different network topologies than information capital. In particular, a consistent set of results has shown that group enforcement is strong and cooperation is efficient when local subnetworks are tightly *interconnected* (see references in footnote 7). In particular, Ali and Miller (2016) model a dynamic game of repeated cooperation and find that a clique network (a completely connected network) generates more cooperation and higher average cooperation capital than any other networks; Jackson, 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 agents of the link share at least one common friend. Our empirical analysis thus highlights two related measures of network interconnectedness: *network support*, the probability that a friend has one or more common friends; and *network clustering*, the probability that two friends are

¹¹If there is no path between *i* and *j*, then $d(i, j) = \infty$.

connected to each other.¹² Formally,

$$Support_{i}(G) \equiv \frac{\#\{j : G_{ij} = 1 \& (G^{2})_{ij} \ge 1\}}{\#\{j : G_{ij} = 1\}}$$
(2)

Summary. As is evident in Figure 2, the topological structure of social networks can be complex. In Appendix A1, we develop a structural model of how social capital aggregates in social networks, which draws a sharper contrast between topologies that provide information capital and topologies that provide cooperation capital. In the regression analysis that follows, we rely on two simplified network statistics that capture the essence of these different topologies: % Friends with common support, a proxy for cooperation capital, defined by (2); and Unique friends of friends, a proxy for information capital, defined by $D_i^2(G)$ from (1). We will also show results pertaining to *i*'s degree centrality, $D_i^1(G)$, which simply counts the number of unique individuals with whom each person communicates. We will also 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).

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

2.2 Modeling migration

Internal migration in Rwanda. Internal migration is a prominent feature of most developing countries. According to the United Nations Population Division (2013), there are

¹²Support is undefined when an individual has no friend and it is always zero when an individual has only one friend.

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

an estimated 762 million internal migrants in the world. Yet, survey-based data on internal migration are notoriously unreliable, particularly in developing countries where many migrations are temporary (Deshingkar and Grimm, 2005, McKenzie and Sasin, 2007, Carletto, de Brauw and Banerjee, 2012, Lucas, 2015).

Our empirical analysis focuses on internal migration in Rwanda, a small agricultural economy in East Africa. Rwanda has high rates of poverty, estimated by the National Institute of Statistics of Rwanda (2012) to be 56.7% in 2005 (the beginning of the period we study) and 44.9% in 2010-2011; extreme poverty likewise fell from 35.8% in 2005 to 24.1% in 2010-2011. More broadly, this period was one of strong economic growth and poverty reduction (World Bank Group, 2017). The World Bank attributes 32 percent of the reduction of poverty to agricultural production and commercialization and 27 percent to nonfarm employment, with the share of nonfarm employment increasing from 8 percent in 2002 to 26 percent in 2011 (World Bank Group, 2015).

While fewer than 4% of Rwandan residents are born abroad, internal migration in Rwanda is common. According to the National Institute of Statistics of Rwanda (2014), roughly 20% of the resident population has experienced a lifetime migration, with similar migration rates for men and women. As with many predominantly agricultural societies, the most frequent type of internal migration in Rwanda is from one rural location to another (Lucas, 2015). For instance, the World Bank estimates that between 2005-2011, roughly two thirds of all migrants went to rural destinations; less than 20% of migrants were from rural to urban areas (World Bank Group, 2017). Beginning in 2013 (shortly after the period we study), the Government of Rwanda's economic development strategy (EDPRS II) explicitly prioritized urbanization and developing secondary cities as poles of growth (World Bank Group, 2017).

The push and pull factors driving internal migration in Rwanda have varied over the last few decades. The 1994 genocide and surrounding conflict were major drivers of internal migration in the 1990s, but conflict has been far less common since 2000. While the National Institute of Statistics of Rwanda (2014) did not collect data on migration motives, their analysis of patterns of urban and rural migration by gender "suggests that males mainly migrate toward urban areas for employment purposes while women tend to move shorter distances, either for marriage or agricultural purposes" (p. 7). Likewise, a series of reports from the Famine Early Warning System highlights the role that agriculture and construction play in driving labor migration, but also emphasizes the unpredictability of this demand (FEWS NET Rwanda, 2014).

The most comprehensive quantitative study of internal migration in Rwanda, conducted

by the World Bank Group (2017) and based on nationally representative household survey data from 2014 (EICV4), points to push and pull factors that are common in many African countries: "Employment and density — or their absence — are the primary push and pull factor for migration in Rwanda. Rural-to-urban migration is fueled by skills, with highereducated rural dwellers seeking better opportunities and higher returns to education in moredensely populated urban areas. The lack of land is another important push factor driving rural migration to the cities, as well as better access to public services and infrastructure. The opposite is also true, with the lack of skills and education, the high cost of living in the city, and, in secondary cities, the lack of access to public services, all pushing people to move from urban to rural areas... Rural-to-rural migration is fueled by higher population density in the origin district, with less-educated migrants moving to other low-density rural districts where they hope to farm, primarily in the Eastern Province" (page 12).

Measuring migration with mobile phone data. We use mobile phone data to provide quantitative insight into the migration patterns of mobile phone owners in Rwanda. Every time a person uses a mobile phone in Rwanda, the phone company records the approximate location of the subscriber at the time of the event. We use these logs to reconstruct each individual's history of migration, building on prior methodological work developing methods to infer migration from mobile phone data (cf. Blumenstock, 2012, Lai et al., 2019).

Full details of these methods are described in Appendix A3. To summarize, we use the phone data to infer the district of residence d of every individual i in each calendar month t (Figure 3). From this sequence of residential locations, we then determine whether or not each individual migrates in each month. Following Blumenstock (2012), we say that a migration occurs in month t+1 if three conditions are met: (i) the individual's 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).

Using this approach, 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, 23.2% travel from rural to urban districts and 23.4% travel from urban to rural districts.

An advantage of the data is that they can provide a 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 large 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 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.¹⁴

2.3 Data limitations and validation

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

Non-representative population. During the period under study (early 2005 to early 2009), mobile phone penetration rose from roughly 5% to 22% (estimates based on the number of subscribers who appear in our dataset). During this time, mobile phone subscribers in Rwanda were not representative of the larger Rwandan population; survey evidence suggests they were significantly wealthier, older, better educated, and are more likely to be male (Blumenstock and Eagle, 2012). While this non-representativeness limits the external validity of our analysis, survey evidence suggests that the population of phone owners and the population of migrants have similar demographic characteristics.¹⁵ More importantly, our empirical specifications are designed to limit the scope for patterns of phone ownership, including trends in mobile penetration over time, to bias our results — see the discussion of omitted variable bias and shift-share analysis in Section 3.

¹⁴According to the 2012 census: 9% of Rwandans are living 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).

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

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

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

Related, we measure migration based on the movement of phones, rather than with traditional survey-based instruments. Prior work suggests that patterns of migration inferred from mobile phone data broadly match inferences drawn from other sources — this includes work in Rwanda using the same dataset as in this paper (Blumenstock, 2012, Williams et al., 2013), neighboring countries in East Africa (Wesolowski et al., 2013*a,b*, Pindolia et al., 2014), as well as other low-income (Bengtsson et al., 2011, Lu et al., 2016, Lai et al., 2019) and wealthy nations (Lenormand et al., 2014). In our context, the aggregate patterns of population flows that we calculate from the mobile phone data between 2005 and 2009 are broadly similar to those reported in the 2012 Rwandan census, but there are discrepancies between the two measurements. For instance, Figure A1 compares estimates of internal migration from the phone data (red bars) to those from the census (blue bars), as reported by National Institute of Statistics of Rwanda (2014, p.29). These inconsistencies could be due to the non-representativeness of phone owners, to differences in how the two instruments

define migration,¹⁶ or to the fact that mobile phone data is a more sensitive instrument for detecting human mobility than the typical census questionnaire.

3 Identification and estimation

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

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

Simultaneity

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

We address this concern in three principal ways. First, we study the effect of lagged network characteristics on the current decisions of migrants. Specifically, we relate the migration decision made by individual i in month t to the structure of i's social network

¹⁶Our algorithm defines a migrant as someone who remains in one district for 3 or more months and then moves to another district for 3 or more months. The Rwandan census does not capture this type of short-term migration; we instead show the census estimates of internal recent migrants, which are defined as "a person who moved to his/her current district of residence five years or less prior to the census."

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

These techniques reduce, but do not eliminate, the potential for simultaneity. In particular, a migrant might plan her migration many months in advance of migration, and in that process could change her higher-order network structure — for instance by asking a friend to make new friends on her behalf, or by encouraging two friends to talk to each other. To gauge the extent to which this might bias our results, we run several empirical tests, and find little evidence of such anticipatory behavior. For instance, Figure 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 A_2 indicates that the total number of contacts increases over time, but there is no sudden spike in the months before migration; Figure A2b shows analogous 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.¹⁷

Omitted Variables

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

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

¹⁷These tests are similar in spirit to those used by Büchel et al. (2020), who regress migration decisions in t on social network structure aggregated over the period from t - 6 until t - 4. Using Swiss cellphone data, the paper finds an increase in direct calls to the destination beginning in t - 3; it does not present analysis on anticipatory changes in higher-order structure prior to migration.

¹⁸For 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 origindestination-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.

Identification

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

The fixed effects we include significantly reduce the scope for omitted variables to bias our estimates of the effect of network structure on migration, but they do not eliminate such bias entirely. If, for instance, origin-destination wage differentials are individual-specific, our fixed effects will not help. This might occur if carpenters' networks in a particular district grew more interconnected over time (relative to carpenter network growth in other districts) than farmers' networks in that district (again relative to farmers' networks in other locations), and if migration rates of carpenters to that district are higher for reasons unrelated to network structure. Likewise, our use of lagged network structure reduces, but does not eliminate, the likelihood that a migrant would first decide to migrate and then modify his network accordingly.

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

Estimation

Formally, for an individual i considering a move from home district h to destination district d in month t, we wish to estimate the effect of a vector of (s-lagged) network characteristics

 $Z_{ihd(t-s)}$ on the migration decision M_{ihdt} , where M_{ihdt} is a binary variable equal to 1 if the individual 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, we estimate:

$$M_{ihdt} = \beta' Z_{ihd(t-s)} + \pi_{hdt} + \mu_i + \sum_k \eta_k \mathbb{1}(D_{id(t-s)} = k) + \epsilon_{ihdt}$$
(3)

where π_{hdt} are (home district * destination district * month) fixed effects and μ_i are individual fixed effects. We also condition on *i*'s degree centrality in the destination $D_{id(t-s)}$ using a set of fixed effects η_k to control for factors that are invariant across all people with the same number of contacts in the destination. The vector β contains the coefficients of interest, which indicate the average effect of network properties $Z_{ihd(t-s)}$ on the probability of migration. This vector of network characteristics may contain characteristics of the destination network $Z_{id(t-s)}$ as well as characteristics of the home network $Z_{ih(t-s)}$.¹⁹ Except as noted, specifications use cluster robust standard errors, two-way clustered by individual and by home-destination-month (Cameron, Gelbach and Miller, 2011). Alternative treatments of the standard errors are discussed in Section 4.4.

When estimating (3), we include one observation for each potential destination d of each individual i in each month t. We define a 'potential destination' as any non-home district $d \neq h$. Thus, the regression includes at most 26 observations for each individual i in each month t. Individuals are only considered in months where they can be classified as a migrant or a non-migrant in that month, and when the relevant measures of network structure can be calculated from the CDR. When an individual is classified as "other" (see Section 2.2), or when network support is undefined (see Section 2.1), those observations are excluded from the regression.²⁰

We consider the linear model to be our reference specification, since it can be estimated quickly on a very large dataset, even with computationally demanding fixed effects (subsequent robustness tests will include over 600 million fixed effects). However, the linear model is misspecified, since in any month the set of choices the individual faces are not independent.

¹⁹In specifications that include higher-order characteristics of the home network $Z_{ih(t-s)}$, we likewise condition on *i*'s degree centrality at home with a set of fixed effects that control for factors that are invariant across all people with the same number of home contacts $H_{ih(t-s)}$. This adds to (3) an additional term $\sum_{j} \eta_j \mathbb{1}(H_{ih(t-s)} = j)$. ²⁰Network support is undefined or always zero when *i* has fewer than 2 contacts in the destination; those

²⁰Network support is undefined or always zero when i has fewer than 2 contacts in the destination; those individual-month-destination observations are excluded when estimating (3). Results are qualitatively similar when network support is set to zero in these observations and they are included in the regression.

We therefore also present results estimated using a discrete choice model.

In the discrete choice setting, each individual *i* faces 27 mutually exclusive choices in month *t*, one for each district *d* in Rwanda (including the home district *h*). We assume the indirect social capital *i* would receive from being in *d* is a function of fixed characteristics of *d* in the month the choice is being made (π_{dt}) , individual characteristics (μ_i) , and a vector of choice-specific attributes that may also differ across individuals $(Z_{id(t-s)})$:

$$U_{idt} = \mathbb{1}(d \neq h) \left[\beta_1' Z_{id(t-s)}\right] + \mathbb{1}(d=h) \left[\gamma + \beta_2' Z_{id(t-s)}\right] + \pi_{dt} + \mu_i + \epsilon_{idt}$$

$$\tag{4}$$

Our focus is on the influence of *i*'s network $Z_{id(t-s)}$, which the above formulation allows to differ for home networks (d = h) and destination networks $(d \neq h)$. Thus, β_1 and β_2 indicate the influence of home and destination networks, respectively. The parameter γ captures the average tendency for individuals to not migrate.

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

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

which can be estimated with a conditional logit model (using \tilde{U} to denote U without the disturbance term ϵ).²¹ We omit μ_i from \tilde{U} because it does not vary across the set of choices faced by i in month t. As with the linear model, we will also present conditional logit results that condition on i's degree centrality in the destination by adding a vector of η_k fixed effects to (4), just as in (3).²²

4 Results

Table 2 summarizes the main results from estimating the linear 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 higher-order structure, after controlling for the immediate contacts of

²¹We use the approach described by Eaton, Kortum and Sotelo (2012) and Sotelo (2019) to estimate the multinomial model by Poisson pseudo maximum likelihood, using the ppmlhdfe package to facilitate estimation with high-dimensional fixed effects (Correia, Guimarães and Zylkin, 2019).

²²In specifications that include higher-order characteristics of the home network $Z_{ih(t-s)}$, we likewise condition on *i*'s degree centrality at home with a set of fixed effects that control for factors that are invariant across all people with the same number of home contacts $H_{ih(t-s)}$.

the individual (i.e., after including the "degree fixed effects"). The second and third rows of Panel A highlight two main results which we will revisit below: migrants are more likely to go to places where their destination networks are more interconnected; but they are *not* more likely to migrate to destinations where their contacts have a large number of contacts.

Where the first column of Table 2 separately estimates the "push" and "pull" forces of networks on migration (cf. Hare, 1999), Table 3 jointly estimates both effects, to allow for a more direct comparison. Comparing the coefficients on Destination Degree and Home Degree in the first two columns, for instance, we find 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.

Results from estimating the multinomial logit model (4) are presented in Table 4, and are broadly consistent with those of the linear model.²³ Migrants are more likely to likely to migrate to destinations where their networks are interconnected, but are *less* likely to move to destinations where their contacts have a large number of contacts. Comparing home and destination networks, the coefficients associated with destination contacts are 5 to 6 times as large as those associated with home contacts.²⁴

In the subsections below, we disaggregate these results, showing how the average effects in Table 2 mask considerably heterogeneity for different types of social networks and for different types of migrants. We study this heterogeneity using variations of the linear model (3).

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

The first coefficient in Table 2 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. This relationship is disaggregated in Figure 5a, which shows how the average migration rate varies by degree centrality at destination (i.e., the number of unique contacts of the individual). We observe that, for instance, roughly 4% of individuals with 10 contacts in a potential destination d in month t - 2 migrate to that location at t.

This figure also provides some intuition for the role of fixed effects in our identification strategy. The average migration rates depicted Figure 5a are likely confounded by omitted

²³Note that the sign of the coefficients associated with home network characteristics (bottom four rows) should be interpreted differently with the multinomial logit (Table 4) than with the linear model (Table 3): With the linear model, the coefficient indicates how the likelihood of *leaving* home changes with $Z_{ih(t-s)}$; with the multinomial logit, the coefficient indicates how the likelihood of *choosing* home changes with $Z_{id(t-s)}$.

²⁴Table A2 presents multinomial logit results more comparable to Panel A of Table 2, by looking at the choice of destination among actual migrants (i.e., estimated on the subset of individual-months in which the individual is observed to migrate).

variables. For instance, people 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 shows the relationship between migration and degree centrality after conditioning on a series of increasingly restrictive fixed effects. Figure A6a shows results with no fixed effects; Figure A6b includes destination district fixed effects (which most immediately addresses the Kigali concern described above); Figure A6c replaces destination fixed effects with more stringent destination-origin-month fixed effects (π_{hdt}); Figure A6d adds individual fixed effects (μ_i).²⁵ In all figures, the qualitative relationship is remarkably unchanged. Individuals with more contacts in a destination are more likely to migrate to that destination. This relationship is positive, monotonic, and approximately linear with slope of unity.

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

4.2 Higher-order network structure

We next examine how the higher-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 extensiveness described in Section 2.1.

Network 'interconnectedness'

Across all potential migrants, we earlier observed in Table 2 that the interconnectedness of the individual's destination network had a positive and significant effect on migration — suggesting that, on average, networks like G_2 in Figure 1 are more attractive than networks like G_1 .²⁶

$$M_{ihdt} = \sum_{k} \eta_k \mathbb{1}(D_{id(t-s)} = k) + [\pi_d +][\pi_{hdt} +][\mu_i +]\epsilon_{ihdt}$$

²⁵Specifically, Figure A6 shows the η_k coefficients from estimating the following equation, with and without the fixed effects shown in brackets:

 $^{^{26}}$ The fact that people are more likely to go to places where their networks are interconnected may seem natural, 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.

Figure 6 disaggregates this average effect, showing how migration rates vary with different levels of network interconnectedness. The left panels (a and c) show how the average migration rates varies with network support, the measure of interconnectedness defined by equation (2). The right panels (b and d) condition on a rich set of fixed effects to help with causal identification, by plotting the β_k coefficients from the following specification:²⁷

$$M_{ihdt} = \sum_{k} \mathbb{1}(D_{id(t-s)} = k) \left[\eta_k + \beta'_k Z_{ihd(t-s)}\right] + \pi_{hdt} + \mu_i + \epsilon_{ihdt}$$
(5)

The right panels of Figure 6 reinforce the prior finding that people are systematically drawn to places where their networks are more interconnected: all coefficients in Figure 6b are (weakly) positive, and all coefficients in Figure 6d are (weakly) negative. The figure also adds a level of nuance that would not be possible with traditional survey-based data. For instance, the fact that the β_k coefficients in Figure 6b are generally increasing indicates that as the potential migrant has more direct contacts in the destination, the value of interconnections between those contacts increases. Appendix Figure A7 finds qualitatively similar results when using network clustering, instead of network support, as a measure of interconnectedness.²⁸

Network 'extensiveness'

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

Of course, the number of friends of friends a person has is largely determined by the

²⁷This specification is directly analogous to the original linear model (3), but it now estimates the effect of network support $Z_{ihd(t-s)}$ on migration separately for each unique value of network size k.

 $^{^{28}}$ 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).

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 (5)). 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. We also observe that the number of friends of friends at home matters more for people with fewer direct contacts — by the time an individual has a very large number of direct home contacts, their contacts' contacts matter less.

The surprising result is Figure 7b, which indicates that the likelihood of migrating does not generally increase with the number of friends of friends in the destination, after conditioning on the number of friends. The friend of friend effect is positive for people with 1-3destination contacts, but negative for people with > 4 destination contacts. Averaged over all migrants, this effect is negative and insignificant (row 3 of Tables 2, 3 and 4). This result is difficult to reconcile with standard models of information diffusion(e.g., Kempe, Kleinberg and Tardos, 2003, Banerjee et al., 2013). Indeed, much of the literature on migration and social networks suggests that, all else equal, individuals would be more likely to migrate if they have friends with many friends, as such networks would provide more natural conduits for information about job opportunities and the like.

We run a large number of empirical tests to convince ourselves that this pattern is not an artifact of our estimation or measurement strategy — several of these are described in Section 4.4. However, the data consistently indicate that the average migrant is no more likely to go to places where she has a large number of friends of friends. This is perhaps most transparent in Figure A8, which shows the distribution of the count of friends of friends for all migrants and non-migrants with exactly 10 friends in the potential destination. Among this sample, it is apparent that, on average, non-migrants have more friends of friends in the destination networks than migrants.

4.3 Heterogeneity and the 'friend of friend' effect

The effect that networks have on the "average migrant" masks considerable heterogeneity in how different types of migrants are influenced by their social networks. Tables A3-A5 disaggregate the results from Table 3 along several dimensions that are salient in the migration literature: whether the migrant has previously migrated to the destination (Table A3); whether the migrant is between adjacent districts or over longer distances (Table A4); and whether the migrant stays in the destination for a long period of time (Table A5).

Heterogeneity and unawareness of the broader network

Several patterns can be discerned from these tables, but we focus our attention on how the network "extensiveness" effect changes with these different subgroups, as that was the most unexpected of the above results. Here, we find that for certain types of migration — repeat migrations, short-distance migrations, and long-term migrations — the number of friends of friends is positively (though not always significantly) correlated with migration rates. Each of these types of migration are significantly less common than the typical migration event (a first-time, long-distance, and short term migration) — hence the statistically insignificant negative average effect observed in Table 2.

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

Strong ties, weak ties, recent migrants, and recent visits

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 20-50% larger than that of a weak destination tie; at home, the effect of a strong tie is 2-3 times as large as the effect of a weak tie. These results are shown in Table A6, which defines a strong tie as one that supports five or more communication events in the reference month (the 90^{th} percentile of communication frequency).

Recent and co-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). Coefficient estimates in column 3 of Table A7 indicate that knowing a recent migrant in the destination increases the likelihood of migration by roughly 3.5 times the amount as knowing anyone else in the destination. The effect is larger for recent migrants who arrived in the destination very recently (last month) than for recent migrants who arrived at any point prior. Such evidence is consistent with the fact that households and extended families frequently make joint labor allocation decisions (cf. Rosenzweig and Stark, 1989).

We also investigate the possibility that in-person visits to the destination could provide (perhaps complementary) information to potential migrants. Table A8 controls for a binary variable indicating whether i ever appeared in district d in the month prior to t. In column 2, a "prior visit" is defined as making or receiving a call or text message from a tower in d; in column 3, we only consider activity that occurs between 6pm and 7am, in an effort to capture overnight visits. There is a strong correlation between such visits and migration (the effect is 3-5 times as large as the effect of an additional direct contact in the destination). Controlling for these in-person visits does not change the qualitative role that networks play in shaping migration, but it does noticeably attenuate the effect of support decreases by 35-50%; and the effect of friends of friends decreases by 2-30%), suggesting the in-person experience might substitute for network connections. Controlling for in-person visits has little effect on the influence of home network structure.

In the above three tests, we find evidence of the many factors that influence migration: having strong social ties to the destination, recent contacts who migrated there, and inperson visits. Most important to our main results is the fact that none of these factors dominate the migration decision: after controlling for these factors, we still see a strong association between destination support and migration, but no positive relationship with destination friends.

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

These results are presented in Figure 8, which summarizes the regression coefficients from Tables A9 and A10. The figure indicates the sign of the regression coefficient (using +/-

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

This heterogeneity is consistent with the notion, proposed by Dunbar (1998) and others, that people might have a capacity constraint in the number of friendships they can effectively support, which in turn might induce a degree of rivalry for the attention of a friend. In our context, migrants may be drawn to places where they receive their friends' undivided attention.²⁹

Beyond location-specific subnetworks

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

Results in Table A11 suggest that the main results — and in particular the negative or insignificant role of extensiveness — do not depend on restrictions on the location of intermediate connections. Column 1 runs the same specification as Table 2 (panel A, column 4), but with an expanded set of fixed effects to facilitate comparison across columns: extensiveness (friends of friends) is negative and insignificant while interconnectedness (common

²⁹Dunbar 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 interpersonal relationship can be maintained." In the migration context, Beaman (2012) and Dagnelie, Mayda and Maystadt (2019) find evidence that migrants may compete with each other for economic opportunities. See also Wahba and Zenou (2005), who empirically test the tradeoff between information and rivalry in an Egyptian labor market survey. They show that up to a certain (network) size, the network information effect dominates the competition (rivalry) effect so that network is always beneficial for finding a job. However, above a certain size, the second effect dominates the first one so that agents have less chance of finding job when network size increases.

support) is positive and significant. Column 2 shows results when *i*'s direct contact *j* is at home; Column 3 allows for *j* to be in any district other than *h* or *d*; Column 4 shows the joint regression. In all specifications, the extensiveness coefficient is negative or insignificant. The interconnectedness coefficient is positive and significant in all specifications except the second.³⁰

4.4 Robustness and identification (revisited)

Section 3 introduced our identification and estimation strategy. For our estimates to be causal requires the identifying assumption that $E[\epsilon_{ihdt}|\pi_{hdt},\mu_i,\eta_k] = 0$. In other words, we assume that the variation in higher-order network structure we observe is exogenous, conditional on the identity of the individual making the migration decision, the origin-destination-month choice being made, and the number of direct contacts the individual has in that destination in that month. Here, we discuss and test the limitations of that assumption, focusing on the two main threats to identification highlighted in Section 3: simultaneity and omitted variable bias.

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

Our identification relies, in part, on the assumption that migrants do not strategically shape the higher-order structure of their social networks after making the decision to migrate. To support this assumption, Figures A4 and A5 indicate that, even among eventual migrants, average higher order network structure remains stable in the months leading up to migration. Here, we provide additional tests of the appropriateness of this identifying assumption.

First, we increase the lag between measurement of network structure and migration. Our main specifications (e.g., Tables 2 and 3) test how migration decisions in month t (e.g., August 2008) relate to social network structure in month t-2 (e.g., June 2008). Tables A12 and A13 show that results are qualitatively unchanged when migration in t is regressed on network structure in t-6 instead. Migrants may plan migrations more than 6 months in advance, but the similarity of the results using 2-month vs. 6-month lagged networks suggests that strategic network formation is not driving our results.³¹

Second, we test a "shift-share" specification that relates migration decisions to changes in an individual's higher-order destination network structure, holding lower-order network

³⁰The negative coefficient on support in column 2 is likely due to the fact that when *i*'s home network includes an additional contact *j* who provides support to a destination connection *k*, this symmetrically increases the number of supported links *i* has at home.

³¹Results, available on request, show qualitatively similar effects when using lags between 2 and 6 months.

structure fixed. Specifically, we define an early period t_0 (e.g., 6 months prior to migration) and a late period t_1 (e.g., 2 months prior to migration), and measure the change in the higherorder network structure of each individual *i* between t_0 and t_1 . In these specifications, we "freeze" the set of *i*'s direct contacts at t_0 , in an effort to reduce the endogenous decisions that *i* makes about their direct contacts (over whom they presumably have more direct control). The identifying variation in the shift-share specification comes from changes in the contacts of *i*'s "frozen" contacts.

Table A14 presents four variants of this shift-share approach. The first two columns freeze *i*'s contacts from t_0 , and measure changes in the network based on the connections of those t_0 contacts between t_0 and t_1 — even if the direct contact in t_0 is no longer a direct contact in t_1 . The last two columns measure changes in the network based on the set of *i*'s contacts who are connected to *i* in both t_0 and t_1 . All specifications set the late period t_1 at 2 months prior to migration. Odd columns set the early period t_0 at 6 months prior to migration; even columns set t_0 at 12 months prior to migration.

The results of the shift-share analysis are broadly consistent with the main results in Table 2. Increases in friends of friends that occur within *i*'s frozen-in-time contacts in the destination are insignificantly or negatively correlated with migration. Increases in support in the destination are positively associated with migration, though these estimates are not always precise. We also note that the total predictive power of changes in network structure is limited (i.e., the partial R^2 values in Table A14 are all less than 0.002; see also Table A15, which shows that changes in network structure between t - 2 and t - 6 are not predictive of migration). If migrants were systematically shaping their higher-order networks in anticipation of migration (and in advance of the t - 2 lag used in our main specification), it is likely that such behavior would better predict migration.

Evidence against omitted variable bias: Increasingly restrictive fixed effects

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

Appendix Table A16 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 η_k . Column 2 in Table A16 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).

To provide better intuition for these results, imagine we want to know whether an individual would be more likely to move to a location where his friends happened to become more connected in the months prior to migration. This is what the coefficient 0.35 in column 3 of Table A16 indicates: fixing the individual and the destination, rates of migration are higher in the months after friends in the destination become more interconnected. To make this more transparent, consider the following: We pull a random sample of 20,000 individuals who have exactly two contacts in a specific district for 4 consecutive months. We then calculate, for each person, whether those two contacts are more likely to become connected or disconnected at the end of the 4-month period (by regressing a dummy for triadic closure on a linear time trend); we then compare the migration rate in month 5 among the population whose two contacts became connected relative to the migration rate in month 5 of the population whose two contacts became disconnected. The migration rate is 2.2% in the former group, and 1.3% in the latter. In other words, when focusing on a sample who consistently have exactly two contacts in the destination, rates of migration are higher when a given individual's two contacts become more connected (over the 4-month period) than when they become more disconnected (over the 4-month period).

 $^{^{32}}$ 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.

³³This 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).

Additional tests of robustness

We perform several additional tests to check whether the main results are sensitive to different measurement strategies used to process the mobile phone data. Since these results show a very similar picture and are highly repetitive, we omit them 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' ≠ 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*.
- 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, "public" phones, and large businesses.
- Standard errors: Our preferred specification uses two-way clustered standard errors, clustered by individual (i), and by the home-destination-month combination (h * d * t). This is because we believe these are the most likely groups among which unobserved components in outcomes would be correlated. We also investigate how alternative approaches to clustering affect the results, including clustering by (a) i; (b) h * d * t; (c) h * d * t and i; (d) h * d * t and i * t; and (e) h * d * t and i * d. Relative to the standard errors shown in Table 2, which uses (c), standard errors are smaller under (a), but nearly identical under (b), (d), and (e).

Summary

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

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

5 Conclusion

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

Our analysis suggests several new stylized facts about the relationship between social networks and migration. Perhaps most surprising, we find that most migrants are *not* drawn to places where their social networks are extensive. Our results suggest that this aversion may stem from the fact that migrants may have limited information about unknown destinations, and that they may feel competition for the attention of their well-connected friends. By contrast, migrants respond strongly to the interlinkages of their friend and kinship networks, and are 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.

More broadly, these results highlight the role of social networks in consequential economic decisions, and how novel sources of data can reveal how network structure influences those decisions. In contexts ranging from product adoption (Banerjee et al., 2013) and disease transmission (Keeling and Eames, 2005) to the spread of new ideas and innovations (Rogers, 1962, Kitsak et al., 2010), simple models of information diffusion have seen remarkable success. Such models often support the stylized narrative that the primary function of networks is to diffuse information about economic opportunities (cf. Rees, 1966, Ioannides and Datcher Loury, 2004). But the patterns revealed by our data are hard to reconcile with these models, and emphasize the importance of social networks that foster repeated cooperation and community interaction.

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. 2019. "Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials." *The Review of Economic Studies*.
- Banerjee, Abhijit V., and Andrew F. Newman. 1998. "Information, the Dual Economy, and Development." *The Review of Economic Studies*, 65(4): 631–653.
- 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."
- Barwick, Panle Jia, Yanyan Liu, Eleonora Patacchini, and Qi Wu. 2019. "Information, Mobile Communication, and Referral Effects." National Bureau of Economic Research Working Paper 25873.
- Büchel, Konstantin, Maximilian V. Ehrlich, Diego Puga, and Elisabet Viladecans-Marsal. 2020. "Calling from the outside: The role of networks in residential mobility." *Journal of Urban Economics*, 119: 103277.
- **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.
- Bengtsson, Linus, Xin Lu, Anna Thorson, Richard Garfield, and Johan von Schreeb. 2011. "Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti." *PLoS Med*, 8(8): e1001083.
- 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. 2010. "Mobile divides: gender, socioeconomic status, and mobile phone use in Rwanda." 6. ACM.
- 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.
- 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.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." *Journal of Business & Economic Statistics*, 29(2): 238– 249. Publisher: Taylor & Francis _eprint: https://doi.org/10.1198/jbes.2010.07136.
- 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.
- Coleman, James S. 1988. "Social capital in the creation of human capital." American journal of sociology, 94: S95–S120.
- Comola, Margherita, and Mariapia Mendola. 2015. "The Formation of Migrant Networks." *Scandinavian Journal of Economics*, 117(2): 592–618.
- Correia, Sergio, Paulo Guimarães, and Thomas Zylkin. 2019. "ppmlhdfe: Fast Poisson Estimation with High-Dimensional Fixed Effects."
- **Dagnelie, Olivier, Anna Maria Mayda, and Jean-François Maystadt.** 2019. "The labor market integration of refugees in the United States: Do entrepreneurs in the network help?" *European Economic Review*, 111: 257–272.
- **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.
- Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. 2016. "Referral-based Job Search Networks." *The Review of Economic Studies*, 83(2): 514–546.
- Easley, David, and Jon Kleinberg. 2010. Networks, Crowds, and Markets: Reasoning About a Highly Connected World. 1 edition ed., New York: Cambridge University Press.
- Eaton, Jonathan, Samuel S Kortum, and Sebastian Sotelo. 2012. "International Trade: Linking Micro and Macro." National Bureau of Economic Research Working Paper 17864. Series: Working Paper Series.
- Edin, Per-Anders, Peter Fredriksson, and Olof Åslund. 2003. "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment." *The Quarterly Journal of Economics*, 118(1): 329–357.

- Fafchamps, Marcel, and Forhad Shilpi. 2013. "Determinants of the Choice of Migration Destination." Oxford Bulletin of Economics and Statistics, 75(3): 388–409.
- **FEWS NET Rwanda.** 2014. "Rwanda Food Security Outlook: July to December 2014." Famine Early Warning Systems Network Food Security Update.
- 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. 2020. "A typology of social capital and associated network measures." Social Choice and Welfare, 54(2): 311–336.
- 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.** 2019. "Distinguishing barriers to insurance in Thai villages." *Journal* of Human Resources.
- 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.
- 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.
- Lai, Shengjie, Elisabeth zu Erbach-Schoenberg, Carla Pezzulo, Nick W. Ruktanonchai, Alessandro Sorichetta, Jessica Steele, Tracey Li, Claire A. Dooley, and Andrew J. Tatem. 2019. "Exploring the use of mobile phone data for national migration statistics." *Palgrave Communications*, 5(1): 1–10. Number: 1 Publisher: Palgrave.
- Lenormand, Maxime, Miguel Picornell, Oliva G. Cantú-Ros, Antònia Tugores, Thomas Louail, Ricardo Herranz, Marc Barthelemy, Enrique Frías-Martínez, and José J. Ramasco. 2014. "Cross-Checking Different Sources of Mobility Information." PLOS ONE, 9(8): e105184.

- 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.
- Lu, Xin, David J. Wrathall, Pål Roe Sundsøy, Md. Nadiruzzaman, Erik Wetter, Asif Iqbal, Taimur Qureshi, Andrew Tatem, Geoffrey Canright, Kenth Engø-Monsen, and Linus Bengtsson. 2016. "Unveiling hidden migration and mobility patterns in climate stressed regions: A longitudinal study of six million anonymous mobile phone users in Bangladesh." Global Environmental Change, 38: 1–7.
- Mahajan, Parag, and Dean Yang. 2017. "Taken by Storm: Hurricanes, Migrant Networks, and U.S. Immigration." National Bureau of Economic Research Working Paper 23756.
- 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. 2012. "The Evolution of Poverty in Rwanda from 2000 to 2011: Results from the Household Surveys (EICV)." Kigali, Rwanda.
- 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.
- Pindolia, Deepa K., Andres J. Garcia, Zhuojie Huang, Timothy Fik, David L. Smith, and Andrew J. Tatem. 2014. "Quantifying cross-border movements and migrations for guiding the strategic planning of malaria control and elimination." *Malaria Journal*, 13(1): 169.
- **Rees, Albert.** 1966. "Information Networks in Labor Markets." *The American Economic Review*, 56(1/2): 559–566.
- Rogers, Everett M. 1962. Diffusion of Innovations. Simon and Schuster.
- Rosenzweig, Mark R., and Oded Stark. 1989. "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India." *Journal of Political Economy*, 97(4): 905–926. ArticleType: research-article / Full publication date: Aug., 1989 / Copyright © 1989 The University of Chicago Press.
- Simmel, Georg. 1950. The sociology of georg simmel. Vol. 92892, Simon and Schuster.
- **Sotelo, Sebastian.** 2019. "Practical aspects of implementing the multinomial pml estimator." Mimeo.
- **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.
- **United Nations Population Division.** 2013. "Cross-national Comparisons of Internal Migration: An Update on Global Patterns and Trends." *Technical Paper No.* 2013/1.
- Wahba, Jackline, and Yves Zenou. 2005. "Density, social networks and job search methods: Theory and application to Egypt." *Journal of Development Economics*, 78(2): 443– 473.

- Wesolowski, Amy, Caroline O. Buckee, Deepa K. Pindolia, Nathan Eagle, David L. Smith, Andres J. Garcia, and Andrew J. Tatem. 2013a. "The Use of Census Migration Data to Approximate Human Movement Patterns across Temporal Scales." *PLOS ONE*, 8(1): e52971.
- Wesolowski, Amy, Nathan Eagle, Abdisalan M. Noor, Robert W. Snow, and Caroline O. Buckee. 2013b. "The impact of biases in mobile phone ownership on estimates of human mobility." *Journal of The Royal Society Interface*, 10(81): 20120986.
- Williams, Nathalie E., Timothy Thomas, Matt Dunbar, Nathan Eagle, and Adrian Dobra. 2013. "Measurement of human mobility using cell phone data: developing big data for demographic science." Citeseer.
- 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.
- World Bank Group. 2015. Rwanda Poverty Assessment. Washington, DC:World Bank.
- World Bank Group. 2017. Reshaping Urbanization in Rwanda: Economic and Spatial Trends and Proposals. Washington, DC:World Bank.

Figures



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

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



Figure 2: The social network of a single migrant

Notes: Diagram shows the social network, as inferred from phone records, of a single migrant i. Nodes represent individuals; edges indicate that two individuals communicated in the month prior to i's migration. Direct contacts of i are shown in blue (for people i's home district), red (for people in i's destination district), and solid grey (for people in other districts). Small hollow circles indicate i's "friends of friends," i.e., people who are not direct contacts of i, but who are direct contacts of i's contacts. All individuals within two hops of i are shown. Nodes are spaced using the force-directed algorithm described in Hu (2005).



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

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



Figure 4: 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.



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

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



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

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



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

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



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

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

Tables

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

Table 1: Summary statistics of mobile phone metadata

Notes: Migration statistics calculated from Rwandan mobile phone data. Column (1) is based on data from a single month; column (2) includes two years of data, potentially counting each individual more than once. "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.

| | (1) | (2) | (3) | (4) | |
|--|-----------------|-----------|-----------------|-----------------|--|
| Panel A: Destination network characteristics | | | | | |
| Degree (network size) | 3.6547*** | | | | |
| | (0.0630) | | | | |
| % Friends with common support | | 1.4813*** | | 1.4808*** | |
| | | (0.1431) | | (0.1435) | |
| Unique friends of friends | | | -0.0005 | -0.0002 | |
| | | | (0.0035) | (0.0035) | |
| Observations | 9,889,981 | 9,889,981 | 9,889,981 | 9,889,981 | |
| R^2 | 0.1851 | 0.1853 | 0.1853 | 0.1853 | |
| Panel B: Home network characteristics | | | | | |
| Degree (network size) | -0.3985^{***} | | | | |
| | (0.0145) | | | | |
| % Friends with common support | | -0.3467 | | -0.5710 | |
| | | (0.3838) | | (0.3800) | |
| Unique friends of friends | | | -0.0089^{***} | -0.0089^{***} | |
| | | | (0.0013) | (0.0013) | |
| Observations | 9,889,981 | 9,889,981 | 9,889,981 | 9,889,981 | |
| R^2 | 0.1743 | 0.1751 | 0.1751 | 0.1751 | |
| Degree fixed effects | No | Yes | Yes | Yes | |
| Individual fixed effects | Yes | Yes | Yes | Yes | |
| Home*Destination*Month F.E. | Yes | Yes | Yes | Yes | |

Table 2: Migration and social network structure - linear model

Notes: Each observation corresponds to an individual-month-destination tuple. 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*. Panel A and Panel B represent separate regressions. Coefficients in Panel B indicate how the probability of leaving home changes as the home network changes (i.e., negative coefficients indicate *i* is less likely to leave as the variable increases). Social network characteristics calculated in month t - 2. Observations are included for all individual-months where the individual is classified as a migrant or a non-migrant, and for all individual-month-destinations where degree, support, and friends-of-friends are observed both at home and in the destination. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

| | (1) | (2) | (3) |
|---------------------------------------|-----------------|-----------------|-----------------|
| Destination Degree (network size) | 4.8033*** | 3.7637*** | |
| | (0.1424) | (0.0946) | |
| Destination $\%$ friends with support | 3.7855^{***} | 1.7164^{***} | 1.0618^{***} |
| | (0.1975) | (0.1485) | (0.1424) |
| Destination friends of friends | -0.0324^{***} | -0.0059 | -0.0001 |
| | (0.0055) | (0.0037) | (0.0035) |
| Home Degree (network size) | -0.7377^{***} | -0.5089^{***} | |
| | (0.0223) | (0.0259) | |
| Home $\%$ friends with support | 8.1299*** | -6.1902^{***} | 0.2216 |
| | (0.4279) | (0.3942) | (0.3755) |
| Home friends of friends | 0.0113^{***} | 0.0059^{***} | -0.0035^{***} |
| | (0.0008) | (0.0012) | (0.0012) |
| Observations | 9,889,981 | 9,889,981 | 9,889,981 |
| \mathbb{R}^2 | 0.0214 | 0.1859 | 0.1869 |
| Degree fixed effects | No | No | Yes |
| Home*Destination*Month fixed effects | No | Yes | Yes |
| Individual fixed effects | No | Yes | Yes |

Table 3: Jointly estimated effects of home and destination network structure - linear model

Notes: Each observation corresponds to an individual-month-destination tuple. 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*. Coefficients on home network characteristics indicate how the probability of leaving home changes as the home network changes (i.e., negative coefficients indicate *i* is less likely to leave as the variable increases). Social network characteristics calculated in month t-2. Observations are included for all individual-months where the individual is classified as a migrant or a non-migrant, and for all individual-month-destinations where degree, support, and friends-of-friends are observed both at home and in the destination. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

| | (1) | (2) | (3) |
|---------------------------------------|--------------|------------|-----------|
| Destination Degree (network size) | 171.57*** | 224.46*** | |
| | (1.71) | (1.75) | |
| Destination $\%$ friends with support | 2549.03*** | 2091.76*** | 189.14*** |
| | (8.64) | (9.17) | (8.88) |
| Destination friends of friends | -1.21*** | -4.05*** | -0.38*** |
| | (0.06) | (0.06) | (0.04) |
| Home Degree (network size) | 28.09*** | 46.09*** | |
| | (0.57) | (0.63) | |
| Home $\%$ friends with support | 504.28*** | 478.95*** | -16.89* |
| | (9.17) | (9.16) | (9.30) |
| Home friends of friends | 0.46^{***} | -0.33*** | 0.20** |
| | (0.02) | (0.02) | (0.02) |
| Home district | 5.93*** | | |
| | (0.01) | | |
| Observations | 9,838,431 | 9,838,431 | 9,838,431 |
| Pseudo \mathbb{R}^2 | 0.71 | 0.71 | 0.72 |
| Degree fixed effects | No | No | Yes |
| Destination*Month fixed effects | No | Yes | Yes |
| Individual*Month fixed effects | Yes | Yes | Yes |

Table 4: Effects of home & destination network structure - conditional logit results

Notes: Each observation corresponds to an individual-month-district tuple. Each column indicates a separate regression of a binary variable indicating 1 if an individual *i* chose to live in district *d* in month *t*. Results are estimated using a conditional logit model, using social network characteristics of the location calculated in month t - 2. "Home District" is a binary variable indicating whether the destination choice in *t* is *i*'s home in t - 1. Different from the linear model, coefficients on home network characteristics indicate how the probability of choosing the home district in *t* changes as the home network changes (i.e., positive coefficients indicate the individual is more likely to stay at home as the variable increases). Degree fixed effects, as well as Destination*Month fixed effects, are interacted with the Home District fixed effect. See discussion in Section 3. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. Standard errors are clustered by individual. *p<0.1; **p<0.05; ***p<0.01.

Appendices - For Online Publication

A1 Model and structural estimation

A1.1 A model of social capital and migration

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

$$u_i(G^d) > u_i(G^h) + \varepsilon_i. \tag{6}$$

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

A1.1.1 Information capital: competition and 'extensiveness'

A robust theoretical and empirical literature suggests that the value of a social network stems, at least in part, from its ability to efficiently transmit information (see footnote 8). We build on recent efforts by Banerjee et al. (2013) to model this information capital as an information sharing process with possible loss of information. It is worth noting that Banerjee et al. (2013) study a seeding process in which an agent is injected with one unit of information, and this agent's diffusion centrality measures the impact of his information to the network. We study a receiving process in which each agent is initially endowed with one piece of information. The nature of information could be about job openings, the work environment, compensation, and so on. We seek to measure how much information an agent could receive from the network. Using the same information sharing process as Banerjee et al. (2013), we will show that the (proxy) measure we seek turns out to be diffusion centrality, because the flow of information is symmetric.³⁴

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

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

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

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

 $^{^{34}}$ Our (proxy) measure counts the sum of expected number of times an agent hears about each piece of information, similar to Banerjee et al. (2013).

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

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

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

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

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

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

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

All proofs are in Appendix A2. This result shows that when λ is small, having more neighbors of neighbors increases one's information capital, whereas when λ is large (i.e.,

close to one), having more indirect neighbors decreases one's information capital. Thus, λ allows for extensive networks to be either beneficial or harmful.

A1.1.2 Cooperation capital: support and 'interconnectedness'

Social networks also facilitate interactions that benefit from community cooperation and enforcement, such as risk sharing and social insurance. We model this dynamic following the setup of Ali and Miller (2016), which highlights the importance of *supported* relationships, where a link is supported if the two nodes of the link share at least one common neighbor (see also Jackson, 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, instead of sharing information, agent i and j now choose their effort levels a_{ij}, a_{ji} in $[0, \infty)$ as their contributions to a joint project.³⁵ Player i's stage game payoff function when partnership ij meets is $b(a_{ji}) - c(a_{ij})$, where $b(a_{ji})$ is the benefit from her partner j's effort and $c(a_{ij})$ is the cost she incurs from her own effort. We normalize the net value of effort a as b(a) - c(a) = a, and assume the cost function c is a smooth function satisfying c(0) = 0 and the following assumption.

ASSUMPTION 1. The cost of effort c is strictly increasing and strictly convex, with c(0) = c'(0) = 0 and $\lim_{a\to\infty} c'(a) = \infty$. The "relative cost" c(a)/a is strictly increasing.

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

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

³⁵The variable-stakes formulation is adopted from Ghosh and Ray (1996) and Kranton (1996).

and this information travels instantly.³⁶

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.

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

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

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

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

 $^{^{36}}$ 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.

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.

$$b(a) \le a + 2 \int_0^\infty e^{-rt} \delta a dt.$$
(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 j's incentive to cooperate in the triangle ijk because it can sustain a^T by itself.

A1.2 Complete model

We now return to the migration decision. In equation (12),

$$u_i = U(u_i^I, u_i^C), \tag{12}$$

we assume that *i*'s utility from a network contains information capital and cooperation capital $(u_i^I \text{ and } u_i^C)$; here, we further assume that the utility can be expressed as a linear combination of these two capitals. This stylized formulation is not meant to imply that u^I and u^C are orthogonal or that other aspects of the network do not weigh in the decision to migrate. Rather, this linear combination is intended to provide a simple benchmark that contrasts two archetypical properties of network structure, which we can also estimate with our data. Appendix A2.1 develops a more general model of network utility, based on a network game approach, which allows for more complex interactions among agents (for instance that an individual's utility may be affected by her position in the global network as well as her local network structure).³⁷ Appendix A2.2.2 shows that similar results obtain when we consider a log-linear (Cobb-Douglas) utility function.

As outlined in Section A1.1.1, we say that agent *i*'s information capital is proportional to their modified diffusion centrality $DC_i(q, \lambda, T)$, which is the *i*-th element of the vector in (9). We derive *i*'s cooperation capital from Corollary 1 in Section A1.1.2, which implies that supported links are more valuable than unsupported links:

$$u_i^C = u_1 d_i^{NS} + u_2 d_i^S, (13)$$

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

$$u_i = u_0 DC_i(q, \lambda, T) + u_1 d_i^{NS} + u_2 d_i^S.$$
(14)

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

$$u_i = \pi^I DC_i(q, \lambda, T) + \pi^C \left(d_i + \alpha d_i^S \right).$$
(15)

Substituting (15) into the original migration decision (6), we have

$$\pi^{I,d}DC_i(G^d;q,\lambda,T) + \pi^{C,d} \left(d_i(G^d) + \alpha^d d_i^S(G^d) \right)$$

> $\pi^{I,h}DC_i(G^h;q,\lambda,T) + \pi^{C,h} \left(d_i(G^h) + \alpha^h d_i^S(G^h) \right) + \varepsilon_i.$ (16)

Notice that we allow agents to have different weights $(\pi^{I,d}, \pi^{C,d}, \pi^{I,h}, \pi^{C,h})$ for the home

³⁷The 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, it could not be realistically computed on our data. (As a point of comparison, calibration of the far simpler model (12) 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.

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.³⁸

A1.3 Model parameterization

We use the migration decisions made by several hundred thousand migrants over a 4.5-year period to estimate the parameters of model (16). 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 = 3.³⁹ Since a very large number of combinations of possible parameters exist, we use an iterative grid-search maximization strategy where we initially specify a large set of values for each parameters, then focus and expand the search around local maxima.⁴⁰

Estimation appears to be well-behaved. For instance, Figure A12 shows the home and

 $^{^{38}}$ In particular, the 'destination' network of a given migrant is actually the 'home' network for all of the contacts that live in that destination.

³⁹When we treat q as a free parameter and estimate it via MLE, the likelihood-maximizing value of q is very close to $1/\mu_1(G)$. Banerjee et al. (2013) show that this approach to measuring diffusion centrality closely approximates a structural property of "communication centrality." However, we cannot directly estimate this latter property on our empirical network, which contains hundreds of thousands of nodes and tens of millions of edges.

⁴⁰Specifically, for each possible set of parameters $\langle \lambda, \alpha^d, \alpha^h, \pi^{I,d}, \pi^{C,d}, \pi^{I,h} \rangle$, 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.

destination utility values for all 270,000 individuals, using the parameterized version of model (16). 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 A9 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 (16), and calculate the percentage of correct classifications. The figures show the the marginal distributions over a single parameter of the accuracy for the top percentile of parameter vectors. In most cases, the likelihood function is concave around the global maximum.

The structural model is largely being identified by the same variation that drives the reduced-form results. For instance, 97.5% of the variation in the total simulated utility of the destination network can be explained by the three main measures of network structure used in Section 4.⁴¹ 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 A17, 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 A17 and was insignificant in Table 2. This shows that when the rivalry parameter λ is optimally chosen for the structural model, the average effect of one's second-neighborhood becomes negative.

A1.4 Parameterization results

Estimation of the model yields several results. First, we find an optimal value of the rivalry coefficient at $\lambda = 0.5$, as shown in Figure A9. 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

⁴¹Specifically, we regress the total *simulated* utility in the destination network, using the parameterized structural model, on three 'reduced-form' properties of the individual's social network: the destination degree centrality, the number of unique destination friends of friends, and the destination network support (see Section 2.1 for definitions). In this linear regression (no fixed effects), $R^2 = 0.975$.

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 A10, which shows the distribution of predicted utility from u_i^I and u_i^C for each of the individuals used to estimate the simulation. The bulk of this distribution lies around the 45-degree line, which is where $u_i^I = u_i^C$. This result is perhaps surprising given the reduced-form results presented in Section 4, which suggest that friends of friends in the destination have an insignificant (or negative) effect on the migration decision. However, a critical difference between the reduced 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 (16) where the rivalry coefficient is fixed at $\lambda = 0$. In other words, we use the original diffusion centrality (without λ) to measure the information capital and redo the whole simulation to identify the likelihood-maximizing set of parameters. We find that information capital (as the original diffusion centrality) contributes very little to total network utility; 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 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 reducedform results described earlier. This is an important step, since the reduced form results are based on statistical properties of networks that are correlated in complex ways, which cannot be easily accounted for in a regression specification. The model parameterization also provides independent support for the presence of some degree of rivalry in information diffusion — a possibility that was suggested by the heterogeneity discussed in Section 4.3, but only directly tested through structural estimation.

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

$$u_i(G^d) > u_i(G^h) + \tau + \varepsilon_i.$$
(17)

Separately, instead of the linear form of (16), we consider a Cobb-Douglas utility function which implies a log-linear combination of information capital and cooperation capital. Equation (16) becomes

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

> $\pi^{I,h} \log DC_i(G^h; q, \lambda, T) + \pi^{C,h} \log \left(d_i(G^h) + \alpha^h d_i^S(G^h) \right) + \varepsilon_i.$ (18)

Results in Appendix A2.2 show that the key qualitative results persist under these alternative specifications of model (16).

A1.5 Proofs

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

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

decreases in λ .

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

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

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

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

Without agent k, agent j's degree decreases by one and so does any of j and k's common neighbors l. Also, agent i on longer gets indirect information from k, which is reflected as $(N_j \setminus g_k) \cup (N_j \cap N_k) = N_j \setminus \{k\}$. We have,

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

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

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

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

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

$$+ q^2 \sum_{l \in N_j \cap N_k} \left(\frac{1}{d_i^{\lambda}(d_j-1)^{2\lambda}(d_l-1)^{\lambda}} - \frac{1}{d_i^{\lambda}(d_j)^{2\lambda}d_l^{\lambda}} \right).$$
(19)

Clearly, each of the four terms in (19) 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.

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

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, $g_{ijk_1...k_h}$. Then assign the effort $a(k_mk_l|ijk_1...k_h) = a^{c=h+2}$ to each link k_mk_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 $g_{ijk'_1...k'_{h-1}}$, assign $a(k'_mk'_l|ijk'_1...k'_{h-1})$ to each link $k'_mk'_l$ to bind the inequality:

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

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 $g_{ijk''_1...k''_{h+1-\eta}}$, assign $a(k''_mk''_l|ijk''_1...k''_{h+1-\eta})$ to each link $k''_mk''_l$ to bind the inequality:

$$b(a) \le a + \int_0^\infty e^{-rt} \delta \left((h+2-\eta)a + \sum_{l \in g_{ijk_1'' \dots k_{h+1-\eta}} \setminus \{i,j,k_1'',\dots,k_{h+1-\eta}\}} a(il|ijk_1'' \dots k_{h+1-\eta}'' l) \right) dt$$

• . . .

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

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

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

A2 Extension and discussion of the model

A2.1 A network game approach

In the benchmark model, we assume the total utility each agent gets from the network is a linear combination of information capital and cooperation capital as in equation (12). 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):

$$u_i(\mathbf{a}, G) = \pi a_i - \frac{a_i^2}{2} + \phi \sum_{j=1}^n G_{ij} a_i a_j + \alpha \sum_{j=1}^n G_{ij}^s a_i a_j.$$
(20)

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.⁴² 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 A1.1.2 and the empirical results in Section 4 that an agent may derive additional utility from a supported neighbor. Second, if $\alpha = 0$, then the equilibrium action will be in proportion to the diffusion centrality in Section A1.1.1, $DC(G; q, \lambda, T)$ when $q = \phi$, $\lambda = 0$ and $T \to \infty$. In particular, ϕ can be viewed as the information passing probability q. The equilibrium action of agent *i* depends on the entire network structure, including her indirect neighbors and her supported links, and thus, this network approach allows for these network structures to jointly determine the equilibrium utility an agent gets from the network.

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

PROPOSITION 3. If $\mu_1(\phi G + \alpha G^s) < 1$, the game with payoffs (20) has a unique (and interior) Nash equilibrium in pure strategies given by:

$$\mathbf{a}^* = \pi (\mathbf{I} - \phi G - \alpha G^s)^{-1} \mathbf{1}.$$
(21)

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

$$a_i^* = \pi + \phi \sum_{j=1}^n G_{ij} a_j^* + \alpha \sum_{j=1}^n G_{ij}^s a_j^*.$$
 (22)

In the matrix form: $\mathbf{a}^* = \pi \mathbf{1} + \phi G \mathbf{a}^* + \alpha G^s \mathbf{a}^*$, which leads to the solution in (21).

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

⁴²While 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.

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 (22), $u_i(\mathbf{a}^*, G) = (a_i^*)^2/2$, which by equation (21) depends on (π, ϕ, α, G) . So in this way, we can estimate how an agent's utility depends on the interaction with neighbors ϕ , the added value of a supported link α , and his or her position in the network G.

More generally, the network game can be enriched to capture the possibilities of competition with indirect neighbors, as we modeled in Section A1.1.1. For example, Ballester, 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.

A2.2 Robustness of model calibration

Our benchmark model assumes that an individual will migrate if the total utility of the destination network exceeds the total utility of the home network (equation 6), 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 12). 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 (6) and (12), to test the robustness of the calibration results in Section A1.4.

A2.2.1 Fixed migration costs

We first allow for the migration decision (equation 6) to include a fixed threshold (cost) τ , in addition to the idiosyncratic error ε_i :

$$u_i(G^d) > u_i(G^h) + \tau + \varepsilon_i.$$
⁽²³⁾

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 (23) is calibrated with the data, the main observations in Section A1.4 persist. Full calibration plots for all parameters $\langle \lambda, \alpha^d, \alpha^h, \tau, \pi^{I,d}, \pi^{C,d}, \pi^{I,h} \rangle$ 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 45degree 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.

A2.2.2 Cobb-Douglas utility

Next, we consider a Cobb-Douglas network utility function, which can be rewritten as the total utility being a log-linear combination of information capital and cooperation capital. Specifically, equation (16) becomes

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

> $\pi^{I,h} \log DC_i(G^h; q, \lambda, T) + \pi^{C,h} \log \left(d_i(G^h) + \alpha^h d_i^S(G^h) \right) + \varepsilon_i.$ (24)

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 A1.4 persist. The log-linear model correctly predicts 68.6% of the migration decisions, which is close to, though slightly below, the accuracy of the model in the text, which is 69.5%. The parameterization plots for $\langle \lambda, \alpha^d, \alpha^h, \pi^{I,d}, \pi^{C,d}, \pi^{I,h} \rangle$ 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.

A3 Measuring migration with mobile phone data

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

First, we extract the timestamp and cell phone tower identifier corresponding to every

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

Next, we assign each subscriber to a "home" district in each month that she makes one or more transactions. Our goal is to identify the location at which the individual spends the majority of her time, and specifically, the majority of her evening hours.⁴³ The full details of this assignment procedure are given in Algorithm 1. To summarize, we first assign all towers to a geographic district, of which there are 30 (we treat the three small districts that comprise the capital of Kigali as a single district). Then, for each individual, we compute the most frequently visited district in every hour of the entire dataset (i.e., there will be a maximum of 4.5 years * 365 days * 24 observations for each individual, though in practice most individuals appear in only a fraction of possible hours). We then aggregate these hourly observations, identifying the district where each individual spends the majority of hours of each night (between 6pm and 7am). Finally, we aggregate these daily observations by identifying the district in which the individual spent the majority of nights in each month. The end result is a panel of individual-month districts.⁴⁴ 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,

 $^{^{43}}$ 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.

⁴⁴At 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.
or stayers. Individuals who do not meet these conditions are treated as "other" (and are excluded from later analysis).⁴⁵ Complete details are given in Algorithm 2.

⁴⁵Individuals 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.

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 the districts exist then

end

Step 3 calculate the *daily modal districts*;

 ${f if}$ tie districts exist then

if overall daily modal districts can resolve then

return the district with larger occurance number; else

 \mathbf{end}

end

end

Step 4 calculate the *monthly modal districts*;

if the districts exist then

 $if \ overall \ monthly \ modal \ districts \ can \ resolve \ then \\$

return the district with larger occurance number;

end

end

Algorithm 1: Home location assignment

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

```
if 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
   \mathbf{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 count of Rwandan migrants (a) to each district and (b) from each district, according to two independent sources of data. Red bars indicate the number of migrants inferred from the mobile phone data, using the methods described in Section 2.2. Blue bars indicate the number of recent internal migrants reported in the 2012 Rwandan census data (National Institute of Statistics of Rwanda, 2014).



Figure A2: Network structure of migrants

Figure A3: Network structure of non-migrants



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 A2a 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 migrants friends who have a common friend.



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

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 4.1). Figure subtitle indicates any other fixed effects included in the specification. Error bars indicate 95% confidence intervals, two-way clustered by individual and by home-destination-month.



Figure A7: Relationship between migration rate and clustering

Notes: "Clustering" denotes the proportion of potential links between *i*'s friends that exist. In all figures, the lower histogram shows the unconditional distribution of the x-variable. Top row (a and b) characterizes the destination network; bottom row (c and d) characterizes the home network. For the left column (a and c), the main figure indicates, at each level of weighted degree, the average migration rate. For the left column (b and d), the main figure indicates the correlation between the migration rate and clustering, holding degree fixed. In other words, each point represents the β_k coefficient estimated from a regression of $Migration_i = \alpha_k + \beta_k Clustering_i$, estimated on the population of *i* who have degree equal to *k*. Error bars for a and c indicate 95% confidence intervals, using the Wilson Score interval for binomial proportions. Error bars for b and d indicate 95% confidence intervals, two-way clustered by individual and by home-destination-month. Coefficients and standard errors on b and d are multiplied by 1000 to make figures legible.



Figure A8: 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 A9: Model calibration results for λ

Notes: Figure shows the marginal effect of varying λ when calibrating Model (16). 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.





Notes: Figure shows the distribution of predicted utility from 'information' capital and 'cooperation' capital (i.e., equation 12) for 270,000 migrants and non-migrants.



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 16. 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

Notes: Figures show the distribution of predicted utility from 'information' and 'cooperation' (i.e., equation 12) for 270,000 migrants and non-migrants. It is calculated using the parameters selected by calibrating Model 16 with λ fixed at zero (i.e., no information rivalry).



Figure A14: Calibration results (with τ): marginal plots

Notes: Figures show the marginal effect of varying λ , α , τ , and π 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 A15: Calibration results (with τ): 'information' and 'cooperation' utility

Notes: Figures show the distribution of predicted utility from 'information' and 'cooperation' (i.e., equation 12) 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 (24). 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 12) for 270,000 migrants and non-migrants. The left figure is calculated using the parameters selected by calibrating Model 24. For the right figure, λ is fixed at zero (i.e., no information rivalry).

| Definition of | Total | % Ever | % Repeat migrants | % Repeat migrants | % Long-distance migrants | % Circular |
|---------------|-------------------|---------|--------------------|-------------------|--------------------------|------------|
| Migrant (k) | Individuals (N) | Migrate | (to same district) | (to any district) | (non-adjacent districts) | 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 |

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

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.

| | (1) | (2) | (3) |
|--|------------|-------------|----------------|
| Degree (network size) in destination | 500.24*** | 509.30*** | |
| | (3.38) | (3.17) | |
| % Friends with common support in destination | 1404.31*** | 1471.87*** | 281.46^{***} |
| | (13.98) | (13.96) | (11.61) |
| Unique friends of friends in destination | -7.00*** | -9.70*** | -1.67*** |
| | (0.01) | (0.01) | (0.07) |
| Observations (person-months) | 260,246 | $251,\!112$ | $251,\!112$ |
| Pseudo \mathbb{R}^2 | 0.23 | 0.30 | 0.34 |
| Degree fixed effects | No | No | Yes |
| Destination*Month fixed effects | No | Yes | Yes |
| Individual*Month fixed effects | Yes | Yes | Yes |

Table A2: Migration and destination network structure - Migrants only, conditional logit

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

| | (1) | (2) | (3) |
|--------------------------------------|-----------------|-----------------|----------------|
| Migration Frequency | Any | Repeat | First-Time |
| % Destination support | 1.0618*** | -2.7428 | 1.0934*** |
| | (0.1424) | (1.8318) | (0.0997) |
| Destination friends of friends | -0.0001 | 0.0171 | -0.0030 |
| | (0.0035) | (0.0125) | (0.0022) |
| % Home support | 0.2216 | 3.7889 | -0.7294^{**} |
| | (0.3755) | (2.6993) | (0.2911) |
| Home friends of friends | -0.0035^{***} | -0.0511^{***} | 0.0022^{***} |
| | (0.0012) | (0.0096) | (0.0007) |
| Observations | 9,889,981 | 665,780 | 9,224,201 |
| \mathbb{R}^2 | 0.1869 | 0.4383 | 0.1986 |
| Degree fixed effects | Yes | Yes | Yes |
| Home*Destination*Month fixed effects | Yes | Yes | Yes |
| Individual fixed effects | Yes | Yes | Yes |

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

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

| | (1) | (2) | (3) |
|--------------------------------|-----------------|----------------------|--------------------------|
| Migration Distance | Any | Short Distance | Long-Distance |
| | | (adjacent districts) | (non-adjacent districts) |
| % Destination support | 1.0618*** | 1.0032*** | 1.0780*** |
| | (0.1424) | (0.2570) | (0.1801) |
| Destination friends of friends | -0.0001 | 0.0042 | -0.0159^{***} |
| | (0.0035) | (0.0052) | (0.0041) |
| % Home support | 0.2216 | -0.4295 | 0.2990 |
| | (0.3755) | (0.6858) | (0.4681) |
| Home friends of friends | -0.0035^{***} | -0.0052^{**} | -0.0028^{**} |
| | (0.0012) | (0.0026) | (0.0014) |
| Observations | 9,889,981 | 3,337,184 | $6,\!552,\!797$ |
| \mathbb{R}^2 | 0.1869 | 0.3237 | 0.1972 |
| Degree fixed effects | Yes | Yes | Yes |
| Home*Destination*Month F.E. | Yes | Yes | Yes |
| Individual fixed effects | Yes | Yes | Yes |

Table A4: Heterogeneity by Distance (Adjacent districts vs. Non-adjacent districts)

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

| | (1) | (2) | (3) |
|--------------------------------------|-----------------|-----------------|-----------------|
| Migration Distance | Any | Long Stay | Short Stay |
| | | (> 12 months) | (< 6 months) |
| % Destination "support" | 1.0618*** | 0.2180*** | 0.8051*** |
| | (0.1424) | (0.0617) | (0.0954) |
| Destination friends of friends | -0.0001 | 0.0156^{***} | -0.0125^{***} |
| | (0.0035) | (0.0021) | (0.0020) |
| % Home "support" | 0.2216 | 0.0928 | 0.1442 |
| | (0.3755) | (0.2149) | (0.2574) |
| Home friends of friends | -0.0035^{***} | -0.0068^{***} | 0.0007 |
| | (0.0012) | (0.0007) | (0.0005) |
| Observations | 9,889,981 | 9,782,384 | 9,820,778 |
| \mathbb{R}^2 | 0.1869 | 0.1445 | 0.1858 |
| Degree fixed effects | Yes | Yes | Yes |
| Home*Destination*Month fixed effects | Yes | Yes | Yes |
| Individual fixed effects | Yes | Yes | Yes |

Table A5: Heterogeneity by Migration Duration (Long-term vs. Short-term)

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

| | (1) | (2) | (3) | (4) |
|--------------------------------|-----------------|-----------------|-----------------|-----------------|
| Destination "Weak tie" | 3.6077*** | 3.7190*** | 3.6771*** | 3.7849*** |
| | (0.0618) | (0.0923) | (0.0613) | (0.0920) |
| Destination "Strong tie" | 4.4319*** | 4.5117*** | 4.4074*** | 4.5034*** |
| - | (0.1218) | (0.1370) | (0.2906) | (0.2976) |
| % Destination "Support" | · · · · | 1.8786*** | × , | 1.8158*** |
| | | (0.1470) | | (0.1446) |
| Destination friends of friends | | -0.0062^{*} | | -0.0061^{*} |
| | | (0.0036) | | (0.0036) |
| Home "Weak tie" | -0.3855^{***} | -0.4813^{***} | -0.4042^{***} | -0.5021^{***} |
| | (0.0144) | (0.0248) | (0.0145) | (0.0249) |
| Home "Strong tie" | -0.7742^{***} | -0.8799^{***} | -1.4034^{***} | -1.5449^{***} |
| | (0.0279) | (0.0345) | (0.1035) | (0.1045) |
| % Home "Support" | | -6.1352^{***} | | -6.1689^{***} |
| | | (0.3786) | | (0.3799) |
| Home friends of friends | | 0.0058^{***} | | 0.0059^{***} |
| | | (0.0011) | | (0.0011) |
| Observations | 9,889,981 | 9,889,981 | 9,889,981 | 9,889,981 |
| \mathbb{R}^2 | 0.1858 | 0.1859 | 0.1858 | 0.1859 |
| 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 |

Table A6: The role of strong ties and weak ties

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. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

| | (1) | (2) | (3) |
|--------------------------------------|-----------------|-----------------|-----------------|
| Destination Degree (network size) | 3.7637*** | 3.6358*** | 3.6513*** |
| | (0.0913) | (0.0933) | (0.0895) |
| % Destination "Support" | 1.7164^{***} | 1.7326*** | 1.7847*** |
| | (0.1434) | (0.1439) | (0.1436) |
| Destination friends of friends | -0.0059 | -0.0041 | -0.0060^{*} |
| | (0.0036) | (0.0036) | (0.0036) |
| Home Degree (network size) | -0.5089^{***} | -0.5171^{***} | -0.5859^{***} |
| | (0.0250) | (0.0248) | (0.0254) |
| % Home "Support" | -6.1902^{***} | -6.1607^{***} | -6.3159^{***} |
| | (0.3806) | (0.3811) | (0.3803) |
| Home friends of friends | 0.0059*** | 0.0060*** | 0.0075^{***} |
| | (0.0011) | (0.0011) | (0.0011) |
| Recent migrant friends | | 1.1090^{***} | 12.6456^{***} |
| | | (0.3129) | (1.0365) |
| Observations | 9,889,981 | 9,889,981 | 9,889,981 |
| \mathbb{R}^2 | 0.1859 | 0.1859 | 0.1870 |
| 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 |

Table A7: The role of recent migrants and co-migrants

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 3; column (2) controls for the number of migrants that i knows, who ever migrated from h to d prior to t; column (3) controls for the number of recent migrants that i knows, who migrated from h to d in the month prior to t. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

| | (1) | (2) | (3) |
|--|-----------------|-----------------|-----------------|
| Destination Degree (network size) | 3.7637*** | 3.4338*** | 3.2097*** |
| | (0.0946) | (0.0906) | (0.0889) |
| Destination $\%$ friends with support | 1.7164^{***} | 1.1105*** | 0.8494^{***} |
| | (0.1485) | (0.1480) | (0.1437) |
| Destination friends of friends | -0.0059 | -0.0058 | -0.0041 |
| | (0.0037) | (0.0037) | (0.0036) |
| Home Degree (network size) | -0.5089^{***} | -0.5050^{***} | -0.5025^{***} |
| | (0.0259) | (0.0258) | (0.0257) |
| Home $\%$ friends with support | -6.1902^{***} | -6.4460^{***} | -6.4846^{***} |
| | (0.3942) | (0.3935) | (0.3936) |
| Home friends of friends | 0.0059*** | 0.0059*** | 0.0060*** |
| | (0.0012) | (0.0012) | (0.0012) |
| Prior visit to destination | | 10.5051*** | |
| | | (0.3216) | |
| Prior visit to destination (evening hours) | | | 14.1288*** |
| | | | (0.3002) |
| Observations | 9,889,981 | 9,889,981 | 9,889,981 |
| \mathbb{R}^2 | 0.1859 | 0.1870 | 0.1878 |
| Degree fixed effects | No | No | No |
| Home*Destination*Month fixed effects | Yes | Yes | Yes |
| Individual fixed effects | Yes | Yes | Yes |

Table A8: Migration and networks, controlling for prior visits to the destination

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 3; column (2) controls for whether i ever used their phone in d in the month prior to t; column (3) controls for whether i ever used their phone in d during the evening hours (i.e., between 6pm and 7am, as a proxy for staying overnight) in the month prior to t. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|--------------------------------------|--------------------|------------------|------------------|------------------|------------------|------------------|
| Destination friends of friends (all) | 0.0004 (0.0034) | | | | | |
| Friends of friends (strong-strong) | | 0.0175 | | | | -0.2288^{***} |
| | | (0.0193) | | | | (0.0346) |
| Friends of friends (strong-weak) | | | 0.0226^{***} | | | 0.0696^{***} |
| | | | (0.0044) | | | (0.0087) |
| Friends of friends (weak-strong) | | | | -0.0460^{***} | | -0.1103^{***} |
| | | | | (0.0135) | | (0.0166) |
| Friends of friends (weak-weak) | | | | | 0.0016 | 0.0224^{***} |
| | | | | | (0.0039) | (0.0054) |
| Observations | $10,\!089,\!959$ | $10,\!089,\!959$ | $10,\!089,\!959$ | $10,\!089,\!959$ | $10,\!089,\!959$ | $10,\!089,\!959$ |
| \mathbb{R}^2 | 0.1909 | 0.1909 | 0.1909 | 0.1909 | 0.1909 | 0.1909 |

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

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

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|--|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|----------------------------|-----------------------------|----------------------------|-----------------------------|
| Support (all) | 1.3211^{***} (0.1415) | | | | | | | | | |
| Support (sss) | | 1.6200^{*} (0.9108) | | | | | | | | 2.5375^{***} (0.9246) |
| Support (sws) | | | 6.9022^{***} (1.0226) | | | | | | | 7.6857^{***} (1.0363) |
| Support (ssw) | | | · · · · | 0.6118^{***} (0.1894) | | | | | | -0.0068 (0.2202) |
| Support (sww) | | | | · · · · | 2.7148^{***} (0.1753) | | | | | 3.0284^{***} (0.1932) |
| Support (wss) | | | | | () | -0.5028^{**} (0.2407) | | | | -0.5472^{*} (0.3257) |
| Support (wws) | | | | | | () | 0.9461^{***} (0.2694) | | | -2.5847^{***} (0.3317) |
| Support (wsw) | | | | | | | (0.2002) | -1.9032^{***} (0.3480) | | (0.3543) |
| Support (www) | | | | | | | | (0.0100) | 1.5830^{***} (0.3260) | (0.3292) |
| Strong tie | 1.3544^{***} (0.0834) | 1.3941^{***} (0.0838) | 1.3132^{***} (0.0836) | 1.3981^{***} (0.0835) | 1.3135^{***} (0.0830) | 1.4268^{***} (0.0839) | 1.3923^{***} (0.0841) | 1.4085^{***} (0.0825) | 1.4242^{***} (0.0825) | 1.2280*** (0.0886) |
| $\begin{array}{c} \text{Observations} \\ \text{R}^2 \end{array}$ | 10,089,959 0.1910 | 10,089,959 0.1910 | 10,089,959 0.1911 |

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

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

| | (1) | (2) | (3) | (4) |
|---|----------------|-----------------|-----------------|-----------------|
| Panel A: Migrant connected to friend | of friend in d | estination via | friend in desti | nation: |
| % friends with common support | 0.7882*** | | | 2.0793*** |
| | (0.1381) | | | (0.2463) |
| Unique friends of friends | -0.0016 | | | 0.0048 |
| | (0.0035) | | | (0.0035) |
| Panel B: Migrant connected to friend | of friend in d | estination via | friend at home | 2: |
| % friends with common support | | -2.7718^{***} | | -4.9650^{***} |
| | | (0.2350) | | (0.3624) |
| Unique friends of friends | | -0.0629^{***} | | -0.0618^{***} |
| | | (0.0032) | | (0.0033) |
| Panel C: Migrant connected to friend | of friend in d | estination via | friend in other | r district: |
| % friends with common support | | | 1.4941*** | 1.5658*** |
| | | | (0.2321) | (0.2764) |
| Unique friends of friends | | | -0.0142^{***} | 0.0006 |
| | | | (0.0039) | (0.0040) |
| Observations | 9,889,981 | 9,889,981 | 9,889,981 | 9,889,981 |
| \mathbb{R}^2 | 0.1872 | 0.1876 | 0.1872 | 0.1877 |
| Destination degree fixed effects | Yes | Yes | Yes | Yes |
| Home degree fixed effects | Yes | Yes | Yes | Yes |
| Degree in other districts fixed effects | Yes | Yes | Yes | Yes |
| Home*Destination*Month fixed effects | Yes | Yes | Yes | Yes |
| Home Destination Month fixed effects | | | | |

Table A11: Beyond location-specific subnetworks

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual *i* migrated from home district *h* to destination district *d* in month *t*. Panels vary the location of the intermediate connection *j* between the migrant *i* and the friend-of-friend *k* in the destination. Column (1) reproduces column 4 of Table 2, Panel A. Column 2 calculates 'friends of friends' and 'friends with support' when *j* is at home. Column 3 uses contacts *j* in any district aside from *h* and *d*. Column 4 shows the joint regression with all six regressors. Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make tables more readable. *p<0.1; **p<0.05; ***p<0.01.

| | (1) | (2) | (3) | (4) | | | | | | |
|----------------------------------|--|-----------------|-----------------|-----------------|--|--|--|--|--|--|
| Panel A: Destination network cho | Panel A: Destination network characteristics | | | | | | | | | |
| Degree (network size) | 3.0636^{***} | | | | | | | | | |
| | (0.0674) | | | | | | | | | |
| % Friends with common support | | 1.9747^{***} | | 1.9760^{***} | | | | | | |
| | | (0.1608) | | (0.1605) | | | | | | |
| Unique friends of friends | | | -0.0010 | -0.0011 | | | | | | |
| | | | (0.0045) | (0.0045) | | | | | | |
| Observations | $6,\!904,\!137$ | $6,\!904,\!137$ | $6,\!904,\!137$ | $6,\!904,\!137$ | | | | | | |
| $\frac{R^2}{2}$ | 0.1791 | 0.1793 | 0.1793 | 0.1793 | | | | | | |
| Panel B: Home network character | ristics | | | | | | | | | |
| Degree (network size) | -0.2615^{***} | | | | | | | | | |
| | (0.0167) | | | | | | | | | |
| % Friends with common support | | -1.4756^{***} | | -1.5383^{***} | | | | | | |
| | | (0.3915) | | (0.3872) | | | | | | |
| Unique friends of friends | | | -0.0052^{***} | -0.0053^{***} | | | | | | |
| | | | (0.0018) | (0.0018) | | | | | | |
| Observations | $6,\!904,\!137$ | $6,\!904,\!137$ | $6,\!904,\!137$ | $6,\!904,\!137$ | | | | | | |
| R^2 | 0.1703 | 0.1709 | 0.1709 | 0.1709 | | | | | | |
| Degree fixed effects | No | Yes | Yes | Yes | | | | | | |
| Individual fixed effects | Yes | Yes | Yes | Yes | | | | | | |
| Home*Destination*Month F.E. | Yes | Yes | Yes | Yes | | | | | | |

| Table A12: | Migration | and so | ocial r | network | structure | (6) | month | network | lag |) |
|------------|-----------|--------|---------|---------|-----------|----------|-------|---------|-----|---|
| | 0 | | | | | \ | | | | |

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

| | (1) | (2) | (3) |
|---------------------------------------|-----------------|-----------------|-----------------|
| Destination Degree (network size) | 4.6645*** | 3.1825*** | |
| | (0.1447) | (0.1039) | |
| Destination $\%$ friends with support | 4.5280*** | 2.3678*** | 1.7605^{***} |
| | (0.2171) | (0.1627) | (0.1608) |
| Destination friends of friends | -0.0446^{***} | -0.0088^{***} | -0.0021^{*} |
| | (0.0066) | (0.0049) | (0.0045) |
| Home Degree (network size) | -0.6265^{***} | -0.3528^{***} | |
| | (0.0208) | (0.0330) | |
| Home $\%$ friends with support | 3.5578^{***} | -4.6670^{***} | -0.7724^{***} |
| | (0.4182) | (0.3913) | (0.3846) |
| Home friends of friends | 0.0098*** | 0.0085^{***} | 0.0025^{***} |
| | (0.0007) | (0.0017) | (0.0018) |
| Observations | $6,\!904,\!137$ | $6,\!904,\!137$ | $6,\!904,\!137$ |
| \mathbb{R}^2 | 0.0204 | 0.1794 | 0.1800 |
| Degree fixed effects | No | No | Yes |
| Home*Destination*Month fixed effects | No | Yes | Yes |
| Individual fixed effects | No | Yes | Yes |
| | | | |

Table A13: Jointly estimated effects (6 month network lag)

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

| | (1) | (2) | (3) | (4) |
|---|-----------------|-----------------|-------------------------|----------------|
| Destination network characteristics | | | | |
| Change in $\%$ friends with common support | 5.9480*** | 5.2585*** | 0.3002 | 0.2575 |
| | (0.1673) | (0.1866) | (0.6268) | (1.1672) |
| Change in friends of friends | -0.0190^{***} | 0.0008 | -0.0151 | 0.0017 |
| | (0.0189) | (0.0136) | (0.0203) | (0.0274) |
| Observations | $6,\!483,\!525$ | $3,\!961,\!415$ | $917,\!557$ | $283,\!171$ |
| \mathbb{R}^2 | 0.1930 | 0.2082 | 0.3938 | 0.4798 |
| Partial \mathbb{R}^2 (from network changes) | 0.0011 | 0.0011 | 0.0001 | 0.0001 |
| "Early period" t_0 | 6 months | 12 months | 6 months | 12 months |
| "Late period" t_1 | 2 months | 2 months | 2 months | 2 months |
| Contacts considered | t_0 contacts | t_0 contacts | $t_0 \cap t_1$ contacts | $t_0 \cap t_1$ |
| Degree fixed effects | Yes | Yes | Yes | Yes |
| Home*Destination*Month fixed effects | Yes | Yes | Yes | Yes |
| Individual fixed effects | Yes | Yes | Yes | Yes |
| | | | | |

Table A14: "Shift share" regression results

Notes: Each column indicates a separate regression of a binary variable indicating 1 if an individual i migrated from home district h to destination district d in month t. Social network characteristics calculated based on changes in higher-order network structure between an early period t_0 and a later period t_1 . Columns 1 and 2 isolate changes in higher order structure based on the direct contacts of i in t_0 . Columns 3 and 4 isolate changes in higher order structure based on the set of contacts who are connected to i in both t_0 and t_1 . Standard errors are two-way clustered by individual and by home-destination-month. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.

| | (1) | (2) | (3) | (4) | | | | | |
|--|------------------|-----------------|-----------------|-----------------|--|--|--|--|--|
| Panel A: Change of destination network characteristics | | | | | | | | | |
| Degree (network size) | -0.5726^{***} | | | -1.1853^{***} | | | | | |
| | (0.1888) | | | (0.3587) | | | | | |
| % Friends with common support | | -0.7758 | | 1.8061^{***} | | | | | |
| | | (0.7497) | | (0.1482) | | | | | |
| Unique friends of friends | | | -0.0149^{***} | 0.0255^{***} | | | | | |
| | | | (0.0056) | (0.0095) | | | | | |
| \mathbb{R}^2 | 0.0004 | 0.0000 | 0.0002 | 0.0005 | | | | | |
| Panel B: Change of home networ | k characteristic | CS | | | | | | | |
| Degree (network size) | -0.0255^{***} | | | -0.0006 | | | | | |
| | (0.0018) | | | (0.0040) | | | | | |
| % Friends with common support | | -0.2884^{***} | | -0.0166 | | | | | |
| | | (0.0373) | | (0.0302) | | | | | |
| Unique friends of friends | | | -0.0009^{***} | -0.0009^{***} | | | | | |
| | | | (0.0001) | (0.0002) | | | | | |
| \mathbb{R}^2 | 0.0001 | 0.00000 | 0.0001 | 0.0001 | | | | | |
| Degree fixed effects | No | No | No | No | | | | | |
| Individual fixed effects | No | No | No | No | | | | | |
| Home*Destination*Month F.E. | No | No | No | No | | | | | |

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*. Independent variables are changes in network properties measured between t-6 and t-2. No fixed effects are used in this specification. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. Standard errors are two-way clustered by individual and by home-destination-month. *p<0.1; **p<0.05; ***p<0.01.

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

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

Table A16: Robustness to alternative fixed effect specifications

| | (1) | (2) | (3) | (4) | | | | | |
|--|------------------|-------------------|-----------------|-------------------|--|--|--|--|--|
| Panel A: Destination network characteristics | | | | | | | | | |
| Degree (network size) | 68.0931*** | | | | | | | | |
| | (0.0450) | | | | | | | | |
| % Friends with common support | | 172.8557^{***} | | 170.7765^{***} | | | | | |
| | | (0.4015) | | (0.4002) | | | | | |
| Unique friends of friends | | | -0.7402^{***} | -0.7033^{***} | | | | | |
| | | | (0.0035) | (0.0034) | | | | | |
| Observations | $6,\!386,\!523$ | $6,\!386,\!523$ | $6,\!386,\!523$ | $6,\!386,\!523$ | | | | | |
| \mathbb{R}^2 | 0.5968 | 0.6359 | 0.6272 | 0.6386 | | | | | |
| Panel B: Home network character | ristics | | | | | | | | |
| Degree (network size) | -11.4922^{***} | | | | | | | | |
| | (0.0197) | | | | | | | | |
| % Friends with common support | | -183.6519^{***} | | -184.6382^{***} | | | | | |
| | | (1.0150) | | (1.0159) | | | | | |
| Unique friends of friends | | | -0.0240^{***} | -0.0364^{***} | | | | | |
| | | | (0.0016) | (0.0016) | | | | | |
| Observations | $6,\!386,\!523$ | $6,\!386,\!523$ | $6,\!386,\!523$ | $6,\!386,\!523$ | | | | | |
| \mathbb{R}^2 | 0.4676 | 0.4948 | 0.4920 | 0.4949 | | | | | |
| Degree fixed effects | No | Yes | Yes | Yes | | | | | |
| Individual fixed effects | Yes | Yes | Yes | Yes | | | | | |
| Home*Destination*Month F.E. | Yes | Yes | Yes | Yes | | | | | |

Table A17: Predicted migration (from structural model) and social network structure

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. Coefficients and standard errors are multiplied by 1000 to make the tables more readable. *p<0.1; **p<0.05; ***p<0.01.