CENTRO STUDI LUCA D'AGLIANO

WWW.DAGLIANO.UNIMI.IT

CENTRO STUDI LUCA D'AGLIANO
DEVELOPMENT STUDIES WORKING PAPERS

N. 353

October 2013

The Formation of Migrant Networks

*Margherita Comola**
*Mariapia Mendola***

* Paris School of Economics (University Paris 1 Panthéon-Sorbonne)
** University of Milan Bicocca and Centro Studi Luca d'Agliano

ISSN 2282-5452

The formation of migrant networks*

Margherita Comola† and Mariapia Mendola‡

June 10, 2013

Abstract

This paper provides the first direct evidence on the determinants of link formation among immigrants in the host society. We use a purposely-designed survey on a representative sample of Sri Lankan immigrants living in Milan to study how migrants form social links among them and the extent to which this network provides them material support along three different dimensions: accommodation, credit, job-finding. Our results show that both weak and strong ties are more likely to exist between immigrants who are born in close-by localities at origin. The time of arrival has a U-shaped effect: links are more frequent between immigrants arrived at the same time, and between long-established immigrants and newcomers. Once the link is formed, material support is provided mainly to relatives while early migrant fellows are helpful for job finding.

JEL codes: J15; D85; C45

Keywords: Migration, Networks, Sri-Lanka, Milan

---

*We are deeply indebted to Raffaele Vacca and the fieldwork team for their excellent work of data collection. We have benefited from useful comments from participants to the 6th International Conference on Migration and Development. Financial support from CISEPS-Bicocca and Centro Studi Luca d'Agliano (LdA) is gratefully acknowledged. All remaining errors are our own.

†Paris School of Economics (Université Paris 1 Panthéon-Sorbonne). Email: margherita.comola@univ-paris1.fr

‡Department of Economics, Università di Milano Bicocca and LdA, Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy. Email: mariapia.mendola@unimib.it

1 Introduction

Interpersonal relationships have long been shown to be a key element in the functioning of imperfect markets and the economy as a whole.[1] At the same time, a growing body of research in economics and other social sciences has documented that network formation is an endogenous process with potentially uneven consequences on individual outcomes and distribution (Jackson and Rogers, 2007).

The purpose of this paper is to investigate the factors determining the formation of social networks among immigrants in the host society, and their economic function. It is well recognized that social ties are particularly important to the migrant population, since newcomers often lack skills or knowledge specific to the receiving country (*e.g.* Massey *et al.* 1999; Munshi, 2003). However, by assuming that migrants interact in groups, much of the empirical literature has relied on very indirect measures of migrant social ties since investigators typically observe neither the immigrant's social contacts, nor whether individual economic outcomes are a consequence of the specific structure of the network in place. This paper fills this gap and provides what is, to the best of our knowledge, the first systematic evidence on the internal structure of migrant social networks by analyzing the formation of dyadic links among immigrants at destination. We use unique data purposely collected by the authors on an ethnically-homogenous sample of male migrants originally from Sri Lanka and living in Milan. In particular, we have collected detailed information on all interpersonal links and episodes of material supports among sampled individuals, along with socio-economic background, time of immigration and city of origin in the native country.

Our point of departure is the idea that, within a group, individuals are likely to have different patterns of interactions and this may affect their outcomes (Goyal, 2007). The empirical evidence on the creation of links in different contexts have shown that social ties are largely shaped by personal history, interpersonal relationships and geographic proximity (*e.g.* Fafchamps and Lund, 2003; Fafchamps and Gubert, 2007). On the other hand, recent works argue that individual-level heterogeneity, reflected in differences in wealth and race for instance, plays an important role in network formation (*e.g.* Krishnan and Sciubba, 2009; Mayer and Puller, 2008). We further ex-

[1] See *e.g.* Granovetter (1985, 1995) Montgomery (1991), Jackson (2005), Goyal (2007).

plore these arguments for the community of immigrants, a sub-set of the population where personal networks are particulary relevant for social economic outcomes (e.g. Munshi, 2003; Mckenzie and Rapoport, 2010). Yet, while the existing evidence has focused on the size of the migrant community (at either origin or destination) to study the effects of migrants' network, much less is known on the internal structure of such a network and the dynamics of node formation among immigrants in the host country. It is legitimate to assume that the endogenous creation of links among newcomers in the society is substantially different from other link formation processes documented in the literature, which mainly refer to geographically stable communities (e.g. risk-sharing in villages, teenager friendships, homeless).

We examine the formation of personal relationships from the dyadic perspective, as a function of proximity and incentive factors. In our baseline definition a social link exists if the two immigrants talk to each other and know each other personally. We first analyze all interpersonal links among individuals, and later on we restrict the analysis to the subset of 'strong links' (*i.e.* the first people you would contact to ask for help or advice). We find that both types of links are more likely to exist between immigrants who are born in close-by localities in Sri Lanka and arrived in Italy at the same time. Moreover, we find a U-shaped curve describing the relationship between the difference in the time of arrival and the probability to form links, with the turning point at about 20 years difference in the time of arrival. Such a non-linear relation seems to partially support the prediction of a model of link formation based on preferential attachment (Barabási and Albert, 1999). These findings are also consistent with the argument pointed out in the migration literature that newcomers get in contact with earlier-generation migrants in order to obtain information and to move (Massey and Espinosa, 1997; Massey et al. 1987). These results stand robust when we control for demographic and economic characteristics, household composition in Sri Lanka and pre-emigration labor market status.

Second, we restrict the analysis to the sub-sample of linked dyads to investigate the extent to which the social network provides material support along three dimensions that are the most crucial for migrants: accommodation, credit, job finding. Our results suggest that material support is provided mainly to relatives. Conditional on the link being established, we find no further significance of the locality of origin. The time of arrival in

Italy appear significant for job finding only, suggesting that migrants tend to rely on previously emigrated individuals in order to be employed. These results shed new light on the long-standing claim that those who have been at destination longer are likely to provide most of the support within the network (Munshi, 2003).

This paper contributes to the literature on both the economics of social networks and migration. As for the former, the migrant population constitutes an ideal setting where to study the factors determining the formation of new links, as immigrants are typically newcomers in a novel environment where the quality and quantity of information (about the local context and other individuals' characteristics) is particularly low, thereby affecting the economic value of interpersonal links. Moreover, this paper takes - for the first time, to our knowledge, in the migration literature - a network-based dyadic approach to investigate how migrants form their social links to other fellow migrants, and how the formation of these links actually shapes interpersonal exchanges. This analysis provides new insights on how the socio-economic integration of certain immigrants may generate spillover effects depending on their position within the social network.

The rest of the paper is organized as follows. Section 2 describes the background literature. The data are presented in Section 3, while Section 4 and 5 describes the empirical strategy and the results for personal links and material support respectively. Section 6 concludes. Tables and figures are reported at the end of the paper.

2 Background literature

According to network theory, social links are formed by individuals who trade off the costs of creating and maintaining the network against the potential rewards from doing so (Jackson and Wolinsky 1996; Bala and Goyal, 2000; Genicot and Ray, 2003). The expected compensation motivating the initial costly contribution may be in the form of public goods provision, labor or production opportunities, informal insurance, credit access, and more in general mutual help (see also Kimball, 1988; Coate and Ravallion, 1993; Foster and Rosenzweig, 2001).

Several researchers have argued that interpersonal links are more likely to be formed on the basis of assortative matching, *i.e.* between proximate individuals rather than (geographically or socially) distant fellows. On the other hand, in many economic situations of interest, costs and benefits related to link formation have been assumed to increase with distance, either geographic or social. The most common example is when social networks serve a risk-sharing purpose, as gains from risk pooling are assumed to be largest between economic agents with different endowments. Several empirical studies have been testing which variables predict the creation of a link for different economic outcomes in both developed and developing contexts. Among them, Fafchamps and Gubert (2007) show that interpersonal relationships among rural households in the Philippines are mainly determined by proximity factors rather than being the result of purposeful diversification of income risk. By contrast, Mayer and Puller (2008) show that, even after controlling for a variety of measures of socioeconomic background and ability, factors predicting the formation of social links among students on university campuses in the US are related to individual characteristics such as race. Finally, in a recent study Fafchamps *et al.* (2010) control for individual differences and find that a pure network proximity effect has a positive impact on the formation of co-authorship links among economists over a twenty year period.[2]

However, academic research communities, as well as university campuses or traditional village economies, may be particularly restricted and favorable environments where the quantity and quality of information about individual characteristics are relatively high. On the other hand, the degree to which social networks are able to convey (good quality) information, and hence the factors determining their formation, within groups in a less favorable conditions is not unambiguous *a priori*. Since disadvantaged groups may be forced to rely on family, friends and fellows in case of need, the economic value of interpersonal links will be high. At the same time though, social networks may not be able to carry relevant resources or create opportunities for valuable face-to-face interactions in alien contexts, as they may exacerbate existing deprived situations (Calvo-Armengol and Jackson, 2004).[3]

[2]There are other important contributions in the empirical literature on social networks (*e.g.* De Weerdt 2004; Udry and Conley 2010; Fafchamps and Lund, 2003). In particular, Krishnan and Sciubba (2009) and Comola (2010) have documented the role of the connection structure of the network, along with individual characteristics, in shaping the formation of links in rural Ethiopia and Tanzania respectively.

[3]For example, Green *et al.* (1999) show that the use of informal job search strategies, such as using personal contacts like friends or relatives during a job search, results

We explore this issue by studying the formation of social networks within a community of migrants in the host society.

Immigrants typically live, especially in the initial period of settlement and integration, in an environment where public information is hardly available and hence may rely on informal network-based resources to access production and socio-economic opportunities. The importance of social links for the migrant population has been established by a large literature in different social sciences (*e.g.* Tilly, 1990; Massey *et al.* 1999; Winters *et al.* 2001). In particular, it has been shown that migrant networks decrease settlement costs of chain-migrants and grease information flows for job search at destination (Massey and Espinosa, 1997; Orrenious, 1999; Mckenzie and Rapoport, 2010; Genicot and Dolfin, 2010). Similarly, they serve to relax credit constraints (Mckenzie and Rapoport, 2007) and can increase the economic returns to migration. By using retrospective data on Mexico, Munshi (2003) studies job networks among Mexican migrants in the U.S. - measured as the proportion of individuals at destination who belong to a common community at origin - and show that more established migrants help newcomers to be employed and to hold an higher paying occupation.

Mainly due to data limitations, though, most studies use indirect or aggregate measures of migrant social ties across different immigrant groups or over time, ignoring the unobserved individual heterogeneity *within groups*. On the other hand, looking at variation in social connections within a group is key to understand differences in individual migrant behavior. To the best of our knowledge, as of today there exists no empirical studies that directly examine the way pairwise social links are formed among immigrants and the mechanisms through which they exert their purposes. We address this lacuna by employing some popular empirical approaches in the economics of social networks which have yet to be exploited in the migration literature.

in lower-paid jobs for Hispanics, whereas this strategy results in higher paying jobs for whites. Similarly, Kahanec and Mendola (2009) show that in Britain “ethnic networks,” measured by the interactions between individuals of the same ethnic minority, do not play a significant role in facilitating paid employment, while mixed or non-ethnic social networks do.

3 Data and setting

Our study is based on a unique survey covering a sample of co-ethnic migrants originally from Sri Lanka and living in the city of Milan, designed and conducted by the authors between December 2011 and February 2012. In our benchmark model, the sample consists of 5460 dyads based on 105 individual interviews to male Sinhalese immigrants older than 18 years of age.

The Sinhalese are Sri Lanka's ethnic majority, one of the largest immigrant populations in Europe, in Italy in particular.[4] The focus on one homogenous ethnic group is crucial in the study of networks formation among immigrants. This is because if the analysis was based on different ethnic communities, the effect of ethnic variability on the relevant relationships would be likely to hide and confound the effects of variability across individual characteristics of interest.[5] On a similar line of reasoning, our sample includes only male adult migrants, therefore excluding any existing and significant variation in social network formation by gender.

The sampling frame of our survey has been carefully designed as to overcome the common problem of interviewing (regular or irregular) immigrants in a host society, and to obtain a representative sample of a particularly hard-to-trace segments of the population.[6] The sample size has been deliberately

[4]In official statistics, the Sinhalese cannot be distinguished from Tamils, Sri Lanka's second ethnic group, since both Sinhalese and Tamil immigrants are recorded as Sri Lanka nationals. Nevertheless, it is well known that Italy has not been among the main destinations of the Tamil diaspora since the 1980s. More permissive legislations on political asylum have attracted the Tamil emigration towards other western countries, such as the United Kingdom, France and Canada. On the contrary, Italy has been one of the favorite destinations for the Sinhalese migration, which was more difficult in other European or American countries having stricter legislation on labor immigration. Therefore, unlike in other European countries, in Italy official statistics on Sri Lanka nationals can be considered a good approximation of the size of the Sinhalese population in Italian cities.

[5]A number of characteristics vary across ethnic groups, *e.g.* language, religion and other original cultural traits, the history of immigration in the host country or city, the contexts of exit from the country of origin and reception in the host country. Characteristics of this sort are extremely important to the relationships of interest, and especially affect migrants' willingness and ability to establish social relations in the country of origin and in the country of destination, as well as their propensity to acquire values and behaviors of the host society.

[6]Prior to the beginning of the survey different neighborhoods in Milan have been canvassed in order to identify the target population. The actual recruitment of survey respondents has been done by setting public information stands in a set of likely hangout places of Sri Lankan in Milan, distributed across the city. Each stand was set in a preselected location for one day only, with the target of attracting passing-by Sri Lankans *via*

kept small because of the design and scope of our study, which imposes a stringent trade off between quantity and quality of elicited network information as explained in what follows. Our main goal was to map as accurately as possible all the interpersonal links within the sampled population, avoiding response bias, inaccuracy and fatigue. At the same time, our estimation samples are comparable in size to the risk-sharing data from Tanzania which have been object of numerous articles (e.g. De Weerdt, 2004; De Weerdt and Dercon, 2006; De Weerdt and Fafchamps, 2011; Vandenbossche and Demuynck 2012), to the risk-sharing data from Philippines by Fafchamps and Lund (2003), and to the data on communication among Indian farmers in Comola and Fafchamps (2013).

In all previous network surveys with dyadic information, in order to elicit the links respondents were first invited to give an open list of partners' names, and these names were afterward traced back to the identity of other survey respondents (Fachamps and Lund, 2003; Calvó-Armengol, Patacchini and Zenou, 2009; Banerjee *et al.* 2012). This strategy, which is the most time-efficient to collect dyadic data, has two shortcomings: first, while it certainly picks up the strong links within the sampled community, it may not track satisfactorily the acquaintances of secondary importance from the respondent's perspective, on which we are particularly interested in. Second, it may be a source of bias if respondents tend to list a limited number of partners because they are fatigued by a burdensome questionnaire, and the distribution of links is uneven (*e.g.* the most popular memeber of the community will end up omitting most of his links because he has too many). We have proceeded in the following way instead: at the end of the questionnaire, we have confronted each respondent with the full list of survey participants and their basic information (names, city of origin in Sri Lanka, job and place of residence in Milan).[7] We have asked the respondent to go

advertisement boards written in Sinhalese. Those who stopped by were offered to leave their coordinates and participate to our remunerated survey (the interviews took place a few weeks after the recruitment). When a group of several people stopped in front of the stand, only one of them was randomly picked to participate to the survey. Great care was taken to recruit individuals from all possible residential areas in Milan such that we may claim that the target population is a representative sample of the Sinhalese community in the city (see Figure 1 and 2 for an overview of sample recruitment sites and sample migrants' residential locations, respectively).

[7] Recruitment of our sample respondents has been made a few weeks before the interviews in order to have in advance a list of participating individuals, along with their basic information.

through all names on the list (with the assistance of the enumerator), and point out those who he knew personally (when requested, we provided the following explanation: "*someone who remembers your name, whose name you remember, to whom you spoke at least once*"). This piece of information was used to define whether a link exists and to trace the community network. More in detail, each adult respondent was asked to list separately the people he knew well (when requested, we provided the following explanation to clarify the concept of knowing well: "*you would personally contact them, or they would personally contact you, to ask for help or advice on important matters*") from the people that he knew, but not well. Along the paper we define the former type of links as strong links. In order to avoid an order effect (*i.e.* respondents read carefully the profile of survey participants at the beginning of the list, and then start loosing concentration because of fatigue) we have confronted different respondents with different lists where the listing order of the survey participants was randomly reshuffled. We have initially capped the number of selected participants to 110, but 5 previously selected individuals on the list either were not reached afterward for the interview or did not complete the questionnaire, which left us with a sample of 105 individual observations. For what concerns the undirected dyadic sample, we thus have $(105 \cdot 104)/2 = 5460$ observations. The network structure is remarkably connected if we consider the sampling strategy of our respondents, and displays the empirical regularities ('small world properties') commonly observed in social networks (Jackson and Rogers, 2007). Indeed, out of the 65 individuals who have at least one link, 60 of them belong to the same component (so-called giant component), and the average geodesic distance among reachable pairs is 4.4 (see Figure 3).

In addition, the dataset contains a rich set of information on the material support flowing on the network, *i.e.* whether individuals have ever exchanged help for sharing the accommodation, for finding a job, or for exchanging loans/gifts. Finally, the survey also collected detailed information on individual sample characteristics (*e.g.* demographics both in Italy and Sri Lanka, age, education, religion), asset endowment (both in Sri Lanka and in Italy), income sources, occupational status and type and intensity of social relations outside the surveyed sample, as reported by the respondents.

The timing and rhythm of their migration make the Sinhalese community a particularly suitable group for the purpose of our analysis. The Sinhalese are one of the oldest immigrant communities in Milan, in the context of relatively recent international migration flows to Italy. At the same time, immigration from Sri Lanka has been ever growing only over the last years, and is still sustained by relevant incoming immigrant flows every year.[8] As a consequence, across Sinhalese immigrants in Milan there is today high variation in years of residence, and hence high variation in variables related to socio-economic integration. On the other hand, like all immigrant minorities in Italy, the Sinhalese in Milan are mostly first-generation immigrants. More than the following generations, first-generation immigrants are in their "halfway" between origin and host society, hence in the position to choose the composition of their *fresh* personal network.[9] Moreover, Sinhalese emigration stems basically from economic reasons, not from political or ethnic persecution in the home country. It is generally a well-prepared emigration, not a sudden, forced departure from home under violent and traumatic circumstances. This kind of emigration is strongly based on migrants' co-ethnic social networks at home and in the host country, through which it is channeled and planned beforehand.

Finally, the residential distribution of the Sinhalese population in Milan is also compatible with our research questions. Census data analysis and previous ethnographic observation pointed out residential concentrations of Sinhalese immigrants in some of the peripheral neighborhoods with the highest incidence of immigrant ethnic minorities in Milan (Vacca, 2013). On the other hand, a relevant part of the Sinhalese community is known to live in some of the most central neighborhoods of Milan, with much lower a proportion of immigrant residents and much higher a socioeconomic profile of the resident population.[10] Thus, the Sinhalese community shows some degree of

[8] In the province of Milan, as of 2009, 17,250 Sri Lankan residents made Sri Lankan nationality - the ninth largest among all foreign nationalities, and the third largest among Asian nationalities (after the Filipinos and the Chinese). This numbers are constantly increasing: according to the last official statistics, coming from the applications for work permits received by the Italian Ministry of Interior on the 1st of January, 2011, the Sri Lankan nationality is the fifth overall for number of applications (the third among Asian nationalities), with 24,563 requests received by the Ministry. Knowing that Milan was the first Italian province for number of applications (it generated about 13% of total applications), we can estimate that there are a few thousands more Sri Lankan labor immigrants seeking entry (or, more typically, legalization) in the province of Milan in 2011.

[9] First-generation immigrants normally show higher overall levels of transnationalism (Itzigshon e Saucedo, 2002), as well as more variation of transnationalism degree across individuals.

[10] This is typical of a very common category of Sinhalese immigrant, those who are

residential diversity, namely a variety of individual residential outcomes in neighborhoods with different degrees of residential segregation.

4 Personal links

In this section we investigate link formation among migrants in our dyadic sample. The descriptive statistics for the estimation sample are reported in Table 1. In Subsection 4.1 we present the main results, while in Subsection 4.2 we discuss the robustness checks.

4.1 Main results

The existence of a link is based on the respondents' answers when asked to indicate those they knew among the survey participants. We first focus on the main definition of link, based on the general question on personal knowledge ("*point out the names of those you know personally*"). Undirected links leave the issue of discordant statements open: the reports of i and j about the link between them should in principle agree, but in practice they often do not. The problem is common to all empirical literature using self-reported link data, and the solution is typically to assume that a link exists if it is reported by either i or j or a combination of the two (Fafchamps and Lund, 2003; De Weerdt, 2004; Snijders, Koskinen and Schweinberger, 2010; De Weerdt and Fafchamps, 2011; Liu *et al.*, 2011; Banerjee *et al.*, 2012). For the main results of Table 2 we assume that a link between i and j exists if either of them declares so (as De Weerdt, 2004; Liu *et al.* 2011; Banerjee *et al.*, 2012), therefore every time a respondent declares to know personally another migrant we draw a link between them (this assumption will be challenged in the next subsection). This provides us with 82 undirected links among the 5460 dyads in the sample,[11] that is, 1.5% of non-zero links. We run the following dyadic linear regression:[12]

$$link_{ij} = X_{ij}^{'}\beta + \varepsilon_{ij} \tag{1}$$

employed as building caretakers or domiciliary caregivers, and are offered to live in the same building where they work.

[11] When the link is undirected only the upper-triangular part of the interaction matrix is used in the dyadic estimation.

[12] For the sake of simplicity we present in the paper the results obtained from a linear specifications (linear probability model). However, all results stand robust (for sign, significance and order of magnitude of the marginal effects) if we run a logit model.

where the unit of observation is the unique undirected dyad ij and $link_{ij} = 1$ if i and j personally know each other. The regressor set X_{ij} includes the constant and the undirected dyadic attributes.

Decisions to link are not independent of each other, since the same survey respondent is part of different dyads. Model prediction errors are therefore correlated, sometimes negatively, across observations, which invalidates inference unless standard errors are corrected to account for non-independence. All along this paper we use the dyadic clustering method first proposed by Fafchamps and Gubert (2007), which allows for arbitrary correlation of ε_{ij} with all $\varepsilon_{i.}, \varepsilon_{j.}, \varepsilon_{.i}$ and $\varepsilon_{.j}$ residuals.

In Table 2 we only include our main exogenous regressors of interest, namely the distance between the city of birth of the two migrants (which may proxy for cultural similarities), and the arrival time in Italy. The three sets of results in columns (1) to (3) correspond to three different functional specifications for the time of arrival in Italy. When the dyadic relationship is undirected, the regressors must enter in a symmetric fashion so that $X'_{ij}\beta = X'_{ji}\beta$ (*i.e.*, for an arbitrary regressor x_z the effect of x_{zi} and x_{zj} on $link_{ij}$ must be the same as the effect of x_{zj} and x_{zi} on $link_{ji}$). This is satisfied for instance if we include dyadic attributes computed from individual characteristics both in sum and in absolute difference (see among others Fafchamps and Gubert; 2007). In column (1) we include as regressors the sum of years in Italy of i and j along with their absolute difference. The former term explores whether there is a higher or lower overall propensity of link formation by earlier immigrants, and the latter term expresses whether migrants tend to form links with those who arrived in the same cohort. It has been shown, though, that long-established migrants may play a different (more significant) role in the newtork than recent migrants (Munshi, 2003). Hence, to explore the issue further in columns (2) and (3) we estimate a more general specification by allowing a single turning point in the absolute difference in the time of arrival (column 2) or a set of different thresholds captured by five different dummies (where less than 5 years absolute difference in the time of arrival is the omitted category. Column 3). Overall, results in Table 2 suggest that distance between cities of birth and time of emigration to Italy play a prominent role in explaining interpersonal relationships among migrants. The distance from the two cities of origin plays a consistently negative effect, suggesting that migrants who are born in close-by localities or coming from common-origin communities are more likely to be connected. As for the vintage of migration, while on average there seems to be a significant negative effect of the absolute difference in the time of arrival and the probability to link, we find a robust and significant U-shaped relationship between the two variables (column 2), such that newcomers tend to link between them but are also more likely to interact with immigrants arrived a long time ago. Since in column (3) the omitted category is 0-5 years, results show that there is no significant difference in the probability of having a link with someone emigrated within the same decade and over 25 years before. On the other hand, the probability of linking with someone emigrated 10 to 25 years before is significantly lower.[13] We further explore this pattern with non-parametric methods. Figure 5 shows the result of a non-parametric local regression of $link_{ij}$ obtained with a smoothing Kernel method trimming the top 1% of the independent variable. The independence variable is the continuous absolute difference in the arrival time in Italy (*e.g.,* if i arrived 16 years ago and j 9 years ago, the difference is 7). The top plot in the figure refer to the dependent variable of Table 2 (where $link_{ij} = 1$ if either migrants declares so), while the bottom plot refer to the alternative link definition of Table 4, that will be explained in the next subsection. Both plots show a neat U-shaped curve, with a long and mild decline up to 25 years difference and a sharp raise afterward. This is indeed the same effect shown by regression estimates. This result goes together with the common perception that the function of the network among migrants is the help in the migration process itself, such that newcomers are likely to interact with early-cohort fellows at destination. In relation with the stochastic network formation literature, these findings partially reconcile with a model of link formation based on preferential attachment, where older nodes have indeed more links and they also receive more links from newborn nodes (Barabási and Albert, 1999; Goyal, van der Leij and Moraga-Gonzales, 2006). Yet, we find that in the community of immigrants this process is non-linear and stronger at both tails of the migration-vintage distribution.

[13]It worhwhile noting that out of our 105 sampled individuals, we have an average of 1.6 links within the sample. Yet, 40 individuals are isolated (i.e. have no declared link within the sample). Restricting to the non-isolated individuals, the mean number of links is 2.5. The number of links seems not to be driven by the years in Milan: the raw correlation between the two variables is rather weak (0.03), as it is confirmed by Figure 4 which plots the relation between the year of arrival in Milan and the number of links.

4.2 Robustness checks

In this subsection we illustrate the robustness of the previous findings along different dimensions. First, in Table 3 we retain the last specification of Table 2, and we check the robustness of results to the inclusion of different sets of controls. In column (1) we add socio-demographic controls (in sum and absolute difference of i and j), namely age, years of education completed, and household size in Italy.[14] In column (2) we include also economic controls (still in sum and absolute difference), namely monthly net income and remittances to Sri Lanka sent in the last year (all rescaled such as 1 unit corresponds to 1000 euros). Finally in column (3) we further control for pre-emigration household and labor market condition in Sri Lanka. As for the household conditions we add the (sum and absolute difference of) the number of strict relatives of the respondent who are still living in Sri Lanka (partner, children, parents). For what regards pre-emigration labor market condition we include two dummies, namely whether both or one of the migrants was a salaried worker in Sri Lanka (rather than unemployed or self-employed). Overall, both the magnitude and the significance of main regressors in Table 3 are robust to inclusion of control variables. The controls do not appear significant in columns (1) to (3), with the exception of the age, along which we observe a high degree of homophyly (i.e. the tendency of migrants to form links with other migrants of similar age).

Second, in Table 4 we report a robustness check where we adopt a more restrictive view of the definition of links for those dyads where the report is discordant (*i.e.* i reports that he knows j but j does not report that he knows i). Facing a discordant dyad, in Tables 2 and 3 we have assumed that the link exists, that is, we have implicitly imputed all differences to under-reporting mistakes. In Table 4 we follow the Fafchamps and Gubert (2007) and Comola and Fafchamps (2013) and whenever the two reports are discordant, we assume that over-reporting and under-reporting are equally likely and we give each measurements equal weight. Operationally, this means that for each unique directed dyad ij we include two observations, namely the report of i and the report of j on the same event (the formula of the dyadic standard error is corrected to take into account this double count). The dataset now includes $(105 \cdot 104) = 10920$ observations, out of

[14]Household members in Italy include relatives (partner, children and other relatives) as well as other children and adults living under the same roof.

which we observe 1% of existing links. Note that the frequency of discordant reports is relatively small (*i.e.* from Table 3 and 4 we have passed from 1.5% to 1% of existing links), especially to other dyadic data analyzed in the network literature, for instance the widely used datasets of Nyanatoke and Add Health (Comola and Fafchamps 2013, Bramoullé Djebbari Fortin 2009; Liu et al., 2011). In our opinion, this is due to the data collection strategy of direct link elicitation (see Section 3). The first column of Table 4 only includes the baseline variables, while the other three columns integrate more and more controls (as in Table 3). Overall, Table 4 reconfirms all the results discussed for Table 3, as it can be noticed in the corresponding non-parametric plot of Figure 3 also showing a neat non-linear shape.

As a final robustness check, in Table 5 we restrict the previous analysis to the subset of links that are declared as strong by the respondent ("*point out the names of those you know well*" - in the few cases where the respondent asked for clarifications, we suggested to mention someone he would contact for help or advice on important issues), assuming under-reporting in case of discordant report as we did in Table 3. Out of the sample of 5460 dyads, we observe 47 existing links (0.86%). The results of Table 5 are very similar to what we have found in Table 3, in terms of both magnitude and significance of the coefficients. The U-shaped effect of arrival time is still present, and now the 21-25 yrs dummy is no long significant, i.e. there is no significant difference in the probability of having a strong link with someone emigrated within the same decade and over 20 years before. Overall, the results seem to suggest that the determinants of link formation among migrants remain mainly time of arrival and distance of city of origin, whether we take into consideration all links or only those personal relationships that are considered of major importance from the respondent's perspective.

5 Material support

In this section we restrict to the sub-sample of linked dyads to investigate the extent to which the social network provides material support along three dimensions that are the most crucial for migrants once they are in the host country: accommodation, credit, job finding.[15] Once the survey respondent

[15]Other important function of the migrants' network that we cannot investigate with the current setting are return migration and social inclusion.

declared to know another migrant in the sample, we have asked an additional battery of question about the nature of their relationship (whether they met in Sri Lanka before moving to Italy, whether they are blood-related) and on flows of help between them (separating help given and received). In particular, each respondent was asked whether he has ever given or received support in terms of accommodation (" *Have you ever hosted him or helped him finding accommodation in Milan?*" and "*Has this person ever hosted you or helped you finding accommodation in Milan?*"), credit ("*Have you ever given a loan or a gift to this person (in money or in kind), which was worth more than 50€?*" and "*Has this person ever given a loan or a gift to you (in money or in kind), which was worth more than 50€?*") and job finding ("*Have you ever helped this person finding a job in Milan?*" and "*Has this person ever helped you finding a job in Milan?*"). We use this pieces of information to run a set of directed dyadic regressions: for each unique dyad ij, we have two observations representing directed flows of help, namely the observation ij representing support flowing from i to j, and the observation ij representing support flowing from i to j. The estimation sample includes those dyads where at least one of the two migrants declare to know personally the other (82 dyads), which makes 164 directed dyadic observations. Note that for each flow ij we now have two reports (*i.e.* what i reports to have given to j and what j reports to have received from i) - whenever these two measurements differ, we take the non-zero report.

Table 6 reports the descriptive statistics of this directed dyadic sample. In Table 7 and 8 we present results from the linear regression:

$$support_{ij} = X_{ij}'\beta + \varepsilon_{ij} \tag{2}$$

where the unit of observation is the directed dyad ij and the dependent variable equals one if i has given support to j, the regressors X_{ij} represent a set of directed dyadic characteristics, and the error term ε_{ij} is clustered to account for dyadic dependence.[16] Note that in a directed estimation framework the regressors should *not* necessarily enter in a symmetric fashion anymore. Tables 7 and 8 are organized as follows: for each type of economic

[16]We present here the linear probability model over logit because, given the exiguous number of observations, some regressors result in perfect prediction of some of the outcomes. This is an issue that arises in every dichotomous regression analysis, such as logit or probit.

support (any support, accommodation only, credit only, job finding only) we have three specifications. In all of the three specifications we include the distance between the localities of birth in Sri Lanka, and two regressors describing the origin of the relationship, namely whether i and j are kin (i.e. blood related), and whether they are not kin but thet already knew each other from Sri Lanka. In order to investigate the effect of the time of arrival on the direct support relationship, we present three different specifications: in Column (1) we only include a dummy taking value one if the giver i arrived in Italy before the receiver j. In Column (2) we include the continous simple difference between the time of arrival (which is positive if the giver i arrived in Italy before the receiver j, and negative otherwise). In column (3) we use a set of three dummies accounting for the directed difference in arrival time, namely: whether i and j arrived in Italy within 5 years of each other, whether i arrived 6-15 years before j, whether i arrived more than 15 years before j (the omitted cathegory is j arriving more than 5 years before i).

From results in Tables 7 and 8, the main determinant factor of any kind of support seems to be kinship, which displays a remarkably significant and large coefficient: the flows of support within the migrant community seems to be preserved within the conservative bounds of family ties. Such an effect is stronger for material support in terms of credit and accomodation rather than job-finding. Conditional on the link being established, we find no further significance of the locality of origin. On the other hand, the time of arrival in Italy appears to have a significant and positive effect on job finding only (Table 8, columns (4) to (6)), suggesting that previously emigrated individuals help newcomers to be employed. In particular, migrant fellows arrived in Italy 6-15 years earlier are those who are more likely to provide support in terms of job-finding. This does not exclude that the social links between long-established migrants and newcomers as evidenced in Table 2 to 5 serve other social purposes, for instance flows of advice and information.

Finally it has to be mentioned that, despite the small sample size, the results of Tables 7 and 8 are remarkably robusts to changes in the specification. In particular, we have performed robustness checks along two lines (results are not reported to economize on space, but are avaiable upon request): first, we have controlled for demographic, economic and pre-emigration controls (as in Tables 3 to 5). Second, we have addressed the potential selection issue by running a Heckman-like two-step dyadic selection model, where the selection equation corresponds to the specification of Column (2) - Table 2, and the outcome equations correspond to the specifications of Tables 7 and 8.[17] In both cases, we found the same results as above regarding kinship, distance and arrival time in Italy.

6 Conclusions

In this paper we carry out what is, to the best of our knowledge, the first systematic study of the determinants of link formation among immigrants in the host society. We use a purposely-designed survey on a representative sample of immigrants originally from Sri Lanka and living in Milan, which contains detailed information on all interpersonal links and material support flows among them. By taking a dyadic perspective we investigate how migrants form their links in the host society and to what extent these links exert their support function in terms of credit access, accomodation and job-finding. We find that migrants tend to interact with co-nationals who come from close-by localities in Sri Lanka and arrived in Italy either at the same time, or long before. The U-shaped relationship between the vintage of migration and the probability of (both weak and strong) link formation stands robust after controlling for a large set of demographic and economic characteristics pre- and post-migration, with a turning point at about 20 years difference in the time of arrival. On the other hand, we do not find that socio-demographic hetergeneity plays a significant role in determining the link formation, with the exception of age that shows a high degree of homophily. These results illustrate the key role played by geographic proximity at origin as well as the time of arrival at destination in shaping the network formation process among immigrants. This is consistent with the common depiction of immigrants being strongly cohesive and supportive among them in the host society (and within the same cohort) but, at the same time, being affected by information scarcity such that newcomers are also more likely to interact with long-established migrants.

We then restrict the analysis to the sample of existing links to study the extent to which the network provides the migrants with material help and

[17] Note that the model is not uniquely identified on non-linearities: since the dyadic selection equation is undirected while the dyadic outcome equation is directed, the different formulation of the regressors accounting for the arrival time serves the purpose of the exclusion restriction.

support. Conditional on the link being established, we find that interpersonal exchanges mainly flow along kinship lines, especially for what regards support in terms of credit and accomodation, while common geographic origin is no longer significant. On the other hand, distant-past migrants seem to remain significanlty helpful for newcomers to find an occupation. Our results provide rigorous evidence that it is the more established members of the network that provide most of the information and support especially in terms of job-finding, while material support may still be preserved within the bounds of family ties even within the migrant segment of the population.

References

[1] Banerjee, A., A. Chandrasekhar, E. Duflo, and M. Jackson (2012), "The diffusion of Microfinance", NBER Working Paper No. 17743.

[2] Barabási A. and R. Albert (1999), Emergence of scaling in random networks, *Science*, 286: 509-512.

[3] Bala, V. and S. Goyal (2000): "A Non-Cooperative Model of Network Formation", *Econometrica*, vol. 68 (5), pp. 1181-1231.

[4] Bramoullé, Y.; Djebbari, H.; Fortin, B. (2009) , "Identification of Peer Effects through Social Networks" , *Journal of Econometrics*, vol. 150 (1): 41-55

[5] Calvó-Armengol, A. and M. O. Jackson (2004): "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, 94(3): 426-454.

[6] Calvó-Armengol, A., Patacchini, E. and Y. Zenou (2009), "Peer effects and social networks in education", *Review of Economic Studies*, 76: 1239-1267

[7] Coate, S. and M. Ravallion (1993): "Reciprocity without Commitment: Characterization and Performance of Informal Insurance Arrangements," *Journal of Development Economics*, 40: 1-24.

[8] Comola, M. (2010), "The network structure of informal arrangements: Evidence from rural Tanzania". PSE working paper n. 2008-74.

[9] Comola, M. and M. Fafchamps (2013) "Testing Unilateral and Bilateral Link Formation", *Economic Journal*, forthcoming.

[10] DeWeerdt, J. And Dercon, S. (2006). "Risk-sharing networks and insurance against illness", *Journal of Development Economics*, vol. 81(2), pp. 337-56.

[11] Conley, T.G. and C.R. Udry, (2010): "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, vol. 100(1): 35-69.

[12] De Weerdt, J. (2004): "Risk Sharing and Endogenous Network Formation", in Insurance Against Poverty, Stephan Dercon (ed.), Oxford University Press.

[13] De Weerdt, J. and M. Fafchamps (2011): "Social Identity and the Formation of Health Insurance Networks", *Journal of Development Studies*, 47(8): 1152-1177.

[14] Fafchamps, M. and S. Lund (2003), "Risk sharing networks in rural Philippines", *Journal of Development Economics*, 71, 261-287.

[15] Fafchamps, M., and F. Gubert (2007). "The Formation of Risk-Sharing Networks", *Journal of Development Economics,* 83(2): 326-350.

[16] Fafchamps, M., Goyal, S. and van der Leij, M. J. 2010. "Matching and Network Effects," *Journal of the European Economic Association*, vol. 8(1): 203-231.

[17] Foster, A. and M. Rosenzweig, 2001. "Imperfect Commitment, Altruism, And The Family: Evidence From Transfer Behavior In Low-Income Rural Areas," *The Review of Economics and Statistics*, vol. 83(3): 389-407.

[18] Genicot, G. and Dolfin, S., (2010) “What Do Networks Do? The Role of Networks on Migration and ’Coyote’ Use” *Review of Development Economics* 14(2): 343-359.

[19] Genicot, G. and Ray, D. (2003): “Endogenous Group Formation in Risk-Sharing Arrangements,” *Review of Economic Studies*, 70 (1).

[20] Goyal, S., M.J. van der Leij and J.L. Moraga (2006), "Economics: An emerging small world", *Journal of Political Economy*, 114(2): 403-412.

[21] Goyal, S. (2007), Connections: an introduction to the economics of networks. Princeton University Press, Princeton, New Jersey.

[22] Granovetter, M. (1985), "Economic Action and Social Structure: The Problem of Embeddedness", *American Journal of Sociology*, 91(3): 481-510.

[23] Granovetter, M. (1995), Getting a Job: A Study of Contacts and Careers, 2nd edition, Chicago: University of Chicago Press.

[24] Green, G. P., Tigges, L. M. & Diaz, D. (1999), "Racial and Ethnic Differences in Job-Search Strategies in Atlanta, Boston, and Los Angeles", *Social Science Quarterly* 80(2): 263-78.

[25] Jackson, M. (2005): "The Economics of Social Networks", CALTECH Humanities and Social Sciences Division Working Paper 1237.

[26] Jackson, M. and B. Rogers (2007), Meeting strangers and friends of friends: How random are social networks?, *American Economic Review*, 97(3): 890-915.

[27] Jackson, M. and A. Wolinsky (1996), "A Strategic Model of Economic and Social Networks", *Journal of Economic Theory*, 71(1): 44-74.

[28] Kahanec, M. and Mendola, M. (2009), "Social determinants of labor market status of ethnic minorities in Britain", *Research in Labor Economics*, 29, 167-195.

[29] Kimball, M. 1988. "Farmers Cooperatives as Behavior Toward Risk". *American Economic Review,* 78(1): 224-232.

[30] Krishnan, P., and Sciubba, E. (2009), "Links and Architecture in Village Networks", *Economic Journal*, vol. 119(537): 917-949.

[31] Liu, X., E. Patacchini, Y. Zenou and L. Lee (2011), "Criminal Networks: Who is the Key Player?", Stockholm University's Research Papers in Economics No. 2011:7.

[32] Mayer, A. and Puller S. (2008): "The old boy (and girl) network: Social network formation on university campuses," *Journal of Public Economics*, 92, 329-347.

[33] Massey, Douglas, J. Arango, G. Hugo, A Kouaouci, A. Pellegrino and J.E. Taylor (1999), "Worlds in Motion: Understanding International Migration at the End of the Millennium," Oxford: Clarendon.

[34] Massey, D. and K. Espinosa (1997), "What's Driving Mexico U.S. Migration? A Theoretical, Empirical, and Policy Analysis." *American Journal of Sociology* 102(4): 939–99.

[35] Massey, D., R. Alarcon, J. Durand, and H. Gonzalez (1987), "Return to Aztlan: The Social Process of International Migration from Western Mexico", Berkeley: University of California Press.

[36] Mckenzie, D. and Rapoport, H., (2010), "Self-Selection Patterns In Mexico-U.S. Migration : The Role Of Migration Networks", *Review of Economics and Statistics*, vol 92(4), 811-821.

[37] Mckenzie, D. and Rapoport, H., (2007). "Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico," *Journal of Development Economics*, vol. 84(1): 1-24.

[38] Montgomery, J. (1991), "Social networks and labor-market outcomes: toward an economic analysis", *American Economic Review*, 81, 5, 1408-1418.

[39] Munshi, K. (2003), "Networks in the modern economy: Mexican migrants in the U.S. labor market", *Quarterly Journal of Economics*, 118(2), 549-599.

[40] Orrenius, P.M. (1999), "The Role of Family Networks, Coyote Prices and the Rural Economy in Migration from Western Mexico: 1965-1994." Federal Reserve Bank of Dallas Working Paper 9910.

[41] Snijders, T., J. Koskinen, and M. Schweinberger (2010): "Maximum Likelihood Estimation for Social Network Dynamics", *Annals of Applied Statistics*, 4(2): 567- 588.

[42] Tilly, C. (1990) ‘Transplanted networks,’ in Immigration Reconsidered, V. Yans-MacLoughlin (Ed.), New York: Oxford University Press, pp. 79-95

[43] Vacca, R. (2013), "Bridging across nations. The social capital of diversity, brokerage and closure in transnational migrant networks: a study on assimilation patterns in Milan and Barcelona", PhD Thesis, Department of Sociology and Social Research, Università degli Studi di Milano-Bicocca.

[44] Vandenbossche, J. and Demuynck, T. (2012). ‘Network formation with heterogeneous agents and absolute friction’, *Computational Economics*, vol. 42(1): 23-45

[45] Winters, P., A. de Janvry and E. Sadoulet (2001), "Family and Community Networks in Mexico-US migration", *Journal of Human Resources* 36(1): 159-184.

Tables and Figures

Table 1: Descriptive Statistics, Undirected Dyads

	n	dummy	mean	sd	min	max
all links, main definition	5460	yes	0.015	0.122	0	1
all links, alternative definition (T. 4)	10920	yes	0.010	0.097	0	1
strong links	5460	yes	0.009	0.092	0	1
years in Italy, abs. diff: |i-j|	5460	no	8.482	7.741	0	36
years in Italy, abs. diff, squared	5460	no	131.87	214.14	0	1296
abs. diff. arrival time: 6-10 yrs	5460	yes	0.224	0.417	0	1
abs. diff. arrival time: 11-15 yrs	5460	yes	0.116	0.321	0	1
abs. diff. arrival time: 16-20 yrs	5460	yes	0.100	0.300	0	1
abs. diff. arrival time: 21-25 yrs	5460	yes	0.050	0.217	0	1
abs. diff. arrival time: $>$ 25 yrs	5460	yes	0.042	0.201	0	1
years in Italy, sum: i+j	5460	no	17.448	11.374	2	70
distance birth cities (km)	5460	no	82.767	59.271	0	424.22
age, sum: i+j	5460	no	83.24	15.07	44	124
age, abs. diff: |i-j|	5460	no	12.432	8.778	0	41
yrs. education, sum: i+j	5460	no	10.057	2.114	3	18
yrs. education, abs. diff: |i-j|	5460	no	1.665	1.336	0	8
household size Italy, sum: i+j	5460	no	5.924	2.265	0	14
household size Italy, abs. diff: |i-j|	5460	no	1.782	1.432	0	7
remittances, sum: i+j	5460	no	6.206	5.578	0	39
remittances, abs. diff: |i-j|	5460	no	3.794	4.162	0	21
income, sum: i+j	5460	no	1.641	0.718	0	4.7
income, abs. diff: |i-j|	5460	no	0.577	0.438	0	2.5
relatives in SL, sum: i+j	5460	no	5.181	2.204	0	11
relatives in SL, abs. diff: |i-j|	5460	no	1.787	1.325	0	6
both salaried in SL	5460	yes	0.358	0.479	0	1
one salaried in SL	5460	yes	0.485	0.500	0	1

Table 2: **Undirected Dyadic Regressions, Main Results**

	(1)	(2)	(3)
distance birth cities (km)	-0.0001***	-0.0001***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)
years in Italy, abs. diff: |i-j|	-0.0008**	-0.0023***	
	(0.0003)	(0.0008)	
years in Italy, abs. diff, squared		0.0001**	
		(0.0000)	
abs. diff. arrival time: 6-10 yrs			-0.0066
			(0.0048)
abs. diff. arrival time: 11-15 yrs			-0.0147**
			(0.0058)
abs. diff. arrival time: 16-20 yrs			-0.0218***
			(0.0075)
abs. diff. arrival time: 21-25 yrs			-0.0206**
			(0.0086)
abs. diff. arrival time: $>$ 25 yrs			-0.0107
			(0.0118)
years in Italy, sum: i+j	0.0004	0.0004	0.0004
	(0.0003)	(0.0003)	(0.0003)
Constant	0.0250***	0.0290***	0.0240***
	(0.0081)	(0.0086)	(0.0080)
Observations	5460	5460	5460

Note: Dyadic standard errors in parentheses (Fafchamps and Gubert, 2007): *** p<0.01, ** p<0.05, * p<0.1

Table 3: Undirected Dyadic Regressions, with Controls

	(1)	(2)	(3)
distance birth cities (km)	-0.0001***	-0.0001***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)
abs. diff. arrival time: 6-10 yrs	-0.0058	-0.0059	-0.0057
	(0.0048)	(0.0052)	(0.0052)
abs. diff. arrival time: 11-15 yrs	-0.0140**	-0.0140**	-0.0145**
	(0.0060)	(0.0060)	(0.0061)
abs. diff. arrival time: 16-20 yrs	-0.0202***	-0.0199**	-0.0200***
	(0.0077)	(0.0078)	(0.0075)
abs. diff. arrival time: 21-25 yrs	-0.0187**	-0.0173**	-0.0168**
	(0.0087)	(0.0083)	(0.0083)
abs. diff. arrival time: $>$ 25 yrs	-0.0081	-0.0061	-0.0057
	(0.0134)	(0.0124)	(0.0123)
years in Italy, sum: i+j	0.0005	0.0004	0.0003
	(0.0004)	(0.0004)	(0.0004)
age, sum: i+j	-0.0002	-0.0002	-0.0001
	(0.0002)	(0.0002)	(0.0002)
age, abs. diff: |i-j|	-0.0005*	-0.0004*	-0.0005**
	(0.0002)	(0.0002)	(0.0002)
yrs. education, sum: i+j	0.0003	-0.0000	0.0001
	(0.0009)	(0.0011)	(0.0012)
yrs. education, abs. diff: |i-j|	-0.0011	-0.0012	-0.0011
	(0.0013)	(0.0013)	(0.0012)
household size Italy, sum: i+j	-0.0007	-0.0006	-0.0008
	(0.0010)	(0.0009)	(0.0010)
household size Italy, abs. diff: |i-j|	0.0008	0.0009	0.0011
	(0.0013)	(0.0013)	(0.0014)
remittances, sum: i+j		0.0003	0.0006
		(0.0008)	(0.0008)
remittances, abs. diff: |i-j|		-0.0005	-0.0006
		(0.0006)	(0.0006)
income, sum: i+j		0.0029	0.0019
		(0.0074)	(0.0069)
income, abs. diff: |i-j|		0.0008	0.0005
		(0.0051)	(0.0052)
relatives in SL, sum: i+j			-0.0021
			(0.0015)
relatives in SL, abs. diff: |i-j|			0.0001
			(0.0015)
both salaried in SL			-0.0049
			(0.0052)
one salaried in SL			0.0007
			(0.0047)
Constant	0.0463*	0.0445*	0.0480*
	(0.0246)	(0.0236)	(0.0245)
Observations	5460	5460	5460

Note: Dyadic standard errors in parentheses (Fafchamps and Gubert, 2007): *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: **Undirected Dyadic Regressions, Alternative Link Definition**

	(1)	(2)	(3)	(4)
distance birth cities (km)	-0.0001***	-0.0001***	-0.0001***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
abs. diff. arrival time: 6-10 yrs	-0.0006	-0.0001	-0.0000	0.0001
	(0.0034)	(0.0033)	(0.0035)	(0.0036)
abs. diff. arrival time: 11-15 yrs	-0.0077**	-0.0074*	-0.0070*	-0.0074**
	(0.0036)	(0.0038)	(0.0037)	(0.0036)
abs. diff. arrival time: 16-20 yrs	-0.0129***	-0.0120***	-0.0113**	-0.0114***
	(0.0042)	(0.0045)	(0.0044)	(0.0043)
abs. diff. arrival time: 21-25 yrs	-0.0125**	-0.0115**	-0.0099*	-0.0095*
	(0.0052)	(0.0054)	(0.0052)	(0.0052)
abs. diff. arrival time: $>$ 25 yrs	-0.0078	-0.0066	-0.0047	-0.0043
	(0.0070)	(0.0079)	(0.0077)	(0.0077)
years in Italy, sum: i+j	0.0002	0.0003	0.0002	0.0001
	(0.0002)	(0.0003)	(0.0003)	(0.0003)
demographic controls	no	yes	yes	yes
economic controls	no	no	yes	yes
pre-emigration controls	no	no	no	yes
Constant	0.0148***	0.0338**	0.0330**	0.0365**
	(0.0055)	(0.0155)	(0.0152)	(0.0163)
Observations	10,920	10,920	10,920	10,920

Note: Dyadic standard errors in parentheses (Fafchamps and Gubert, 2007):
*** p<0.01, ** p<0.05, * p<0.1

Table 5: **Undirected Dyadic Regressions, Strong Links Only**

	(1)	(2)	(3)	(4)
distance birth cities (km)	-0.0001***	-0.0001***	-0.0001***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
abs. diff. arrival time: 6-10 yrs	0.0013	0.0018	0.0017	0.0018
	(0.0045)	(0.0044)	(0.0047)	(0.0047)
abs. diff. arrival time: 11-15 yrs	-0.0089**	-0.0087**	-0.0089**	-0.0092**
	(0.0038)	(0.0039)	(0.0039)	(0.0039)
abs. diff. arrival time: 16-20 yrs	-0.0109**	-0.0101*	-0.0106*	-0.0107*
	(0.0053)	(0.0055)	(0.0057)	(0.0056)
abs. diff. arrival time: 21-25 yrs	-0.0103	-0.0094	-0.0099	-0.0097
	(0.0075)	(0.0079)	(0.0084)	(0.0085)
abs. diff. arrival time: $>$ 25 yrs	-0.0074	-0.0070	-0.0073	-0.0072
	(0.0088)	(0.0093)	(0.0100)	(0.0098)
years in Italy, sum: i+j	0.0002	0.0002	0.0002	0.0002
	(0.0003)	(0.0003)	(0.0004)	(0.0004)
demographic controls	no	yes	yes	yes
economic controls	no	no	yes	yes
pre-emigration controls	no	no	no	yes
Constant	0.0143***	0.0313**	0.0310**	0.0337**
	(0.0051)	(0.0129)	(0.0137)	(0.0154)
Observations	5460	5460	5460	5460

Note: Dyadic standard errors in parentheses (Fafchamps and Gubert, 2007):
*** p<0.01, ** p<0.05, * p<0.1

Table 6: Descriptive Statistics, Directed Dyads

	n	dummy	mean	sd	min	max
support: accomodation	164	yes	0.207	0.407	0	1
support: loans/gifts	164	yes	0.250	0.434	0	1
support: job finding	164	yes	0.207	0.407	0	1
support: any	164	yes	0.372	0.485	0	1
kin	164	yes	0.073	0.261	0	1
non kin, met in SL	164	yes	0.110	0.314	0	1
i arrived first in Italy	164	yes	0.451	0.499	0	1
years in Italy, diff: (i-j)	164	no	0.000	10.468	-35	35
both i and j arrived between 5 yrs	164	yes	0.598	0.492	0	1
i arrived 6-15 yrs before	164	yes	0.146	0.355	0	1
i arrived over 15 yrs before	164	yes	0.055	0.228	0	1
distance birth cities (km)	164	no	54.150	44.391	0	197.2

Table 7: Directed Dyadic Regressions, Any Support and Accomodation Support

	support from i to j: any			support from i to j: accomodation		
	(1)	(2)	(3)	(4)	(5)	(6)
distance birth cities (km)	0.0005	0.0005	0.0004	0.0015	0.0015	0.0015
	(0.0013)	(0.0014)	(0.0014)	(0.0013)	(0.0013)	(0.0012)
i arrived first in Italy	0.0312			-0.0120		
	(0.0457)			(0.0375)		
years in Italy, diff: (i-j)		0.0016			0.0014	
		(0.0020)			(0.0017)	
both i and j arrived between 5 yrs			-0.0422			-0.0011
			(0.1170)			(0.0868)
i arrived 6-15 yrs before			0.0566			0.0733
			(0.0909)			(0.0838)
i arrived over 15 yrs before			-0.1509			-0.0843
			(0.1485)			(0.1383)
kin	0.4313***	0.4327***	0.4223**	0.4769**	0.4764**	0.4689**
	(0.1631)	(0.1629)	(0.1690)	(0.2063)	(0.2067)	(0.2131)
non kin, met in SL	0.0713	0.0710	0.0801	0.1689	0.1691	0.1746
	(0.2196)	(0.2183)	(0.2178)	(0.1788)	(0.1791)	(0.1819)
Constant	0.2918**	0.3061**	0.3383*	0.0780	0.0724	0.0695
	(0.1301)	(0.1218)	(0.1755)	(0.1012)	(0.0941)	(0.1097)
Observations	164	164	164	164	164	164

Note: Dyadic standard errors in parentheses (Fafchamps and Gubert, 2007): *** p<0.01, ** p<0.05, * p<0.1

Table 8: Directed Dyadic Regressions, Support for Credit and Job Finding

	support from i to j: credit (loans/gifts)			support from i to j: job finding		
	(1)	(2)	(3)	(4)	(5)	(6)
distance birth cities (km)	0.0007	0.0007	0.0004	-0.0000	-0.0000	-0.0002
	(0.0013)	(0.0012)	(0.0013)	(0.0008)	(0.0008)	(0.0010)
i arrived first in Italy	0.0066			0.0858*		
	(0.0559)			(0.0444)		
years in Italy, diff: (i-j)		0.0015			0.0020	
		(0.0017)			(0.0018)	
both i and j arrived between 5 yrs			-0.0792			-0.0158
			(0.1103)			(0.0908)
i arrived 6-15 yrs before			0.0605			0.1474**
			(0.0787)			(0.0735)
i arrived over 15 yrs before			-0.1614			-0.0597
			(0.1290)			(0.0901)
kin	0.4611***	0.4614***	0.4493***	0.3310*	0.3348*	0.3239*
	(0.1566)	(0.1570)	(0.1608)	(0.1841)	(0.1821)	(0.1785)
non kin, met in SL	-0.0949	-0.0949	-0.0822	0.1691	0.1682	0.1789
	(0.1362)	(0.1360)	(0.1324)	(0.1832)	(0.1809)	(0.1782)
Constant	0.1864	0.1895*	0.2502	0.1263	0.1658**	0.1672
	(0.1263)	(0.1096)	(0.1597)	(0.0776)	(0.0794)	(0.1185)
Observations	164	164	164	164	164	164

Note: Dyadic standard errors in parentheses (Fafchamps and Gubert, 2007): *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 1: Map of sample recruitment sites in Milan

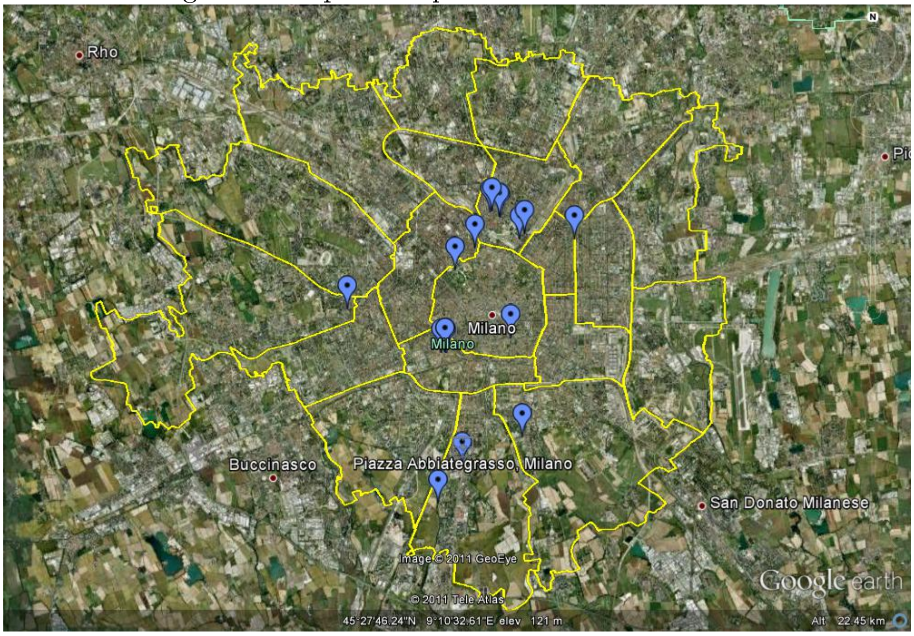

Figure 2: Map of sample migrants' residential locations in Milan

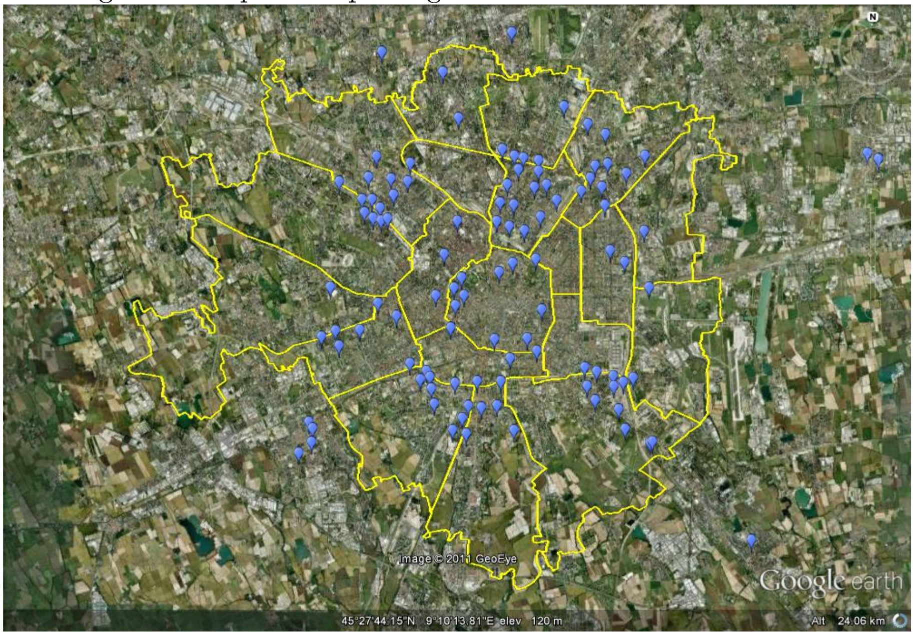

Figure 3: The network structure

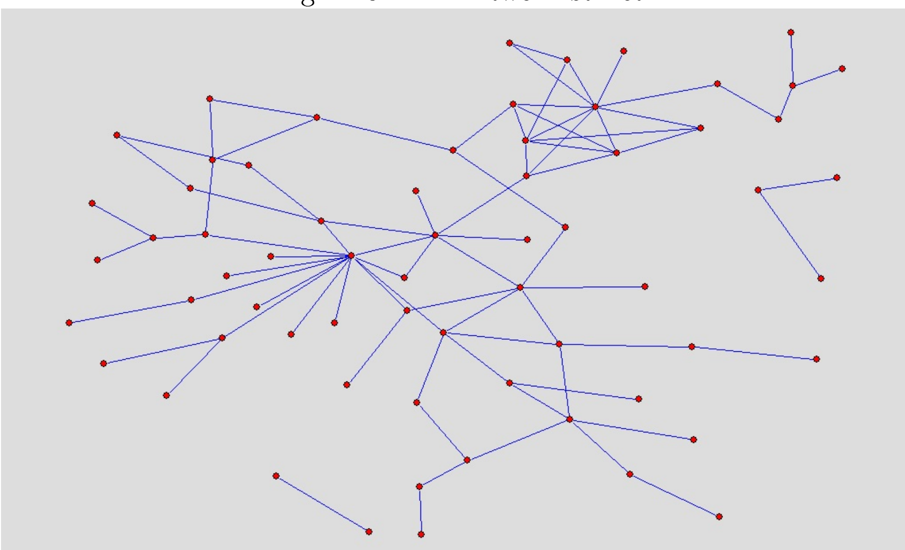

Figure 4: Plot of number of links versus year of arrival

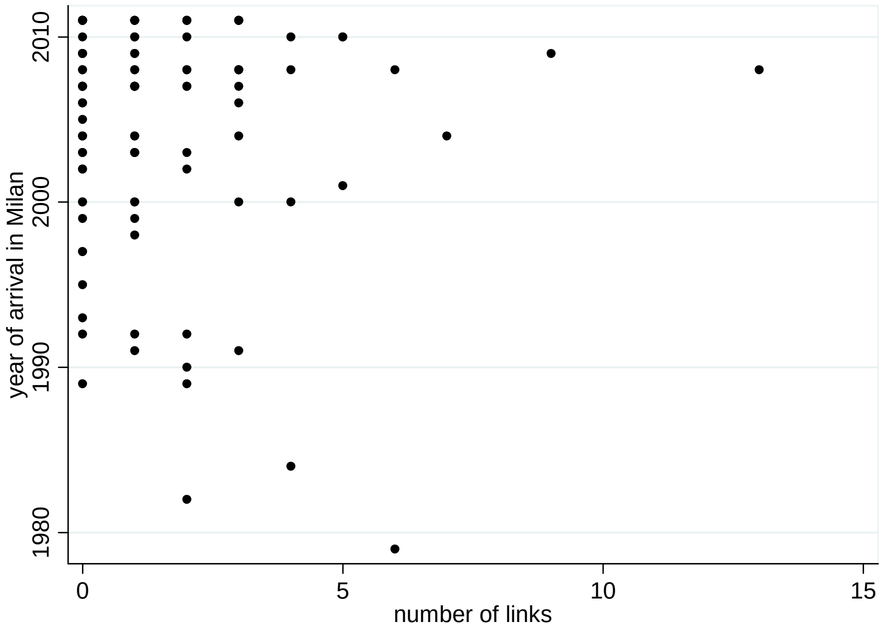

Figure 5: Non-parametric local regression on difference in arrival time