CESifo WORKING PAPERS

6568
2017

July 2017

Cities and the Structure of Social Interactions: Evidence from Mobile Phone Data

Konstantin Büchel, Maximilian von Ehrlich

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 6568*
Category 1: Public Finance

Cities and the Structure of Social Interactions: Evidence from Mobile Phone Data

Abstract

Social interactions are considered pivotal to agglomeration economies. We explore a unique dataset on mobile phone calls to examine how distance and population density shape the structure of social interactions. Exploiting an exogenous change in travel times, we show that distance is highly detrimental to interpersonal exchange. Despite distance-related costs, we find no evidence that urban residents benefit from larger networks when spatial sorting is accounted for. Higher density rather generates a more efficient network in terms of matching and clustering. These differences in network structure capitalize into land prices, corroborating the hypothesis that agglomeration economies operate via network efficiency.

JEL-Codes: R100, R230, D830, D850, Z130.

Keywords: social interactions, agglomeration externalities, network analysis, spatial sorting.

Konstantin Büchel
University of Bern
Department of Economics and Center for Regional Economic Development
Schanzeneckstrasse 1
Switzerland – 3001 Bern
konstantin.buechel@vwi.unibe.ch

Maximilian von Ehrlich
University of Bern
Department of Economics and Center for Regional Economic Development
Schanzeneckstrasse 1
Switzerland – 3001 Bern
maximilian.vonehrlich@vwi.unibe.ch

July 11 2017
We benefited from numerous comments by Juan Becutti, Sascha Becker, Aymo Brunetti, Fabian Gunzinger, Christian Hilber, Blaise Melly, Diego Puga, Elisabet Viladecans-Marsal, Alex Whalley and participants at the IEB research seminar in Barcelona, the Verein für Socialpolitik Meeting in Tellow, the Meeting of the Urban Economics Association in Minneapolis, the CRED-Workshop and Swisscom Brown-Bag in Bern, the SERC Conference in London, the Applecross Conference, and the SSES Annual Congress in Lausanne. A very special thanks is due to Swisscom AG for providing the facilities and data to conduct this research project; we are particularly indebted to Dr. Imad Aad, who accompanied the project for two years. We also want to thank search.ch, comparis.ch and Meta-Sys, for providing data on travel times, usage statistics of messenger apps, and rental prices.

1 Introduction

In the last decades the share of urban population has increased continuously. According to the World Bank (2014), more than 50 percent of the world population lives in cities producing an over-proportional share of global GDP. Cities are regarded as engines of productivity growth since high population density facilitates the social exchange of knowledge and information.[1] Beyond productivity gains, social interactions may directly contribute to a person's wellbeing. Rapid innovation in communication technologies brought forward the hypothesis that the importance of geographical proximity for social exchange and accordingly the dividends of population density have declined. In order to understand why we observe surging urbanization despite rapid technological innovation, we believe it is essential to study how population density shapes the structure of social interactions. In particular, not only the mere quantity of interactions but also the quality and efficiency of social networks is decisive for economic outcomes (see Jackson et al., 2017). Empirical work uncovering these mechanisms has so far been impeded by the lack of comprehensive individual-level data which is necessary to analyze the structure of social interactions.

Using anonymized mobile phone calls allows us to gain insights into this question. We study the relation between population density and social interactions in order to test fundamental assumptions underlying agglomeration dynamics discussed in the literature (c.f. Duranton and Puga, 2004). Our rich dataset covers about 15 million phone calls and text messages per day, collected over a period of 12 months. This allows us to examine the interplay between local characteristics and social interactions as we not only observe communication patterns but also location information derived from transmitting antennas and billing data. Based on this information and concepts from the network literature (c.f. Jackson, 2008), we investigate three main questions: *First*, how does geographical distance impact social interactions? *Second*, what is the relation between population density and the size of an individual's social network? *Third*, does population density affect the quality / efficiency of social interactions in terms of matching quality, clustering and network perimeter?

To answer these questions we evaluate the the role of distance in link formation models and complement this analysis with estimates about the causal impact of population density on micro-level network measures. The sorting of individuals with specific characteristics can distort the results of both approaches. We therefore analyze how link formation is affected by an exogenous change in travel time, triggered by a substantial revision of public transport schedules. In addition, we use individuals who permanently relocate (movers)

[1] This notion reflects one of the classic agglomeration forces described by Alfred Marshall (1890). Knowledge spillovers and innovation in cities feature prominently in seminal work by Jacobs (1969), Lucas (1988), and Glaeser et al. (1992).

to back out time-constant unobservables and identify density-related externalities. The latter identification strategy relates to approaches quantifying the earning advantages of cities (see Combes et al., 2008; De la Roca and Puga, 2017).

We show that distance is highly detrimental to social interactions, despite epoch-making progress in communication technologies. Contrary to the conventional assumption, this does not translate into larger networks in cities compared to the periphery. Density-related externalities rather arise in terms of network efficiency, namely better matching quality, lower clustering, and smaller distance costs. These findings are in line with a search strategic perspective and with the biological/anthropological literature on social groups size. By relating the quality of social networks to land prices we provide an assessment of the monetary value of efficient social networks. We demonstrate that differences in the structure of social networks can explain a significant share of the difference in land prices between cities and peripheral regions.

In recent years, significant progress was made in quantifying the magnitude of agglomeration advantages while there is relatively little evidence about their causes (for an overview see Combes and Gobillon, 2015). We focus on the mechanisms behind and uncover the causal effects of distance and population density on social interactions. These findings are derived from a novel source of information about local economic processes – mobile phone data – which we show to be very useful for further questions in urban and regional economics.[2] Below, we discuss our main findings with reference to the related literature.

Related Literature. Models that incorporate knowledge and learning spillovers as an agglomeration force typically assume that distance is costly for social interactions. The widespread adoption of information and telecommunication technologies popularized the "death-of-distance" argument (e.g. Cairncross, 2001), which raises the intriguing question of whether these technologies will fundamentally change the structure of cities (see Ioannides et al., 2008) or even make them obsolete. As argued by Gaspar and Glaeser (1998), the crucial question is whether face-to-face meetings and phone calls are substitutes or complements. We demonstrate that the social interactions recorded by mobile phones are surprisingly localized, with more than 60 percent of ties occurring between individuals that reside within less than 10 km distance of each other. Our causal estimates provide evidence that face-to-face meetings and telecommunication are complements and thus contradict the death-of-distance hypothesis. This relates to research about local adoption of Internet technologies (Forman et al., 2005) and the regional consequences of the spread of the Internet. Blum and Goldfarb (2006) show that physical distance may even impact

[2] Analyzing data from Rwanda, Blumenstock et al. (2015) show that information about mobile phone usage provides a good proxy for wealth and income.

consumption of online goods due to the formation of local tastes.[3] Forman et al. (2012) establish that the Internet benefits high-income and high-population places rather than reducing regional disparities.

Micro-founded models of urban agglomeration have focused on the assumption that the quantity of social interactions increases with local population density.[4] For instance, Glaeser (1999) formalizes the theory that individuals acquire skills by interacting with each other. As cities are more densely populated than the hinterland, they facilitate more meetings and thus accelerate the social learning process. Sato and Zenou (2015) model social interactions and their impact on employment outcomes. They propose that city residents maintain larger networks than rural residents, enabling them to acquire more information on the labor market, which reduces job search frictions and unemployment. We show that the positive effect of cities compared to the hinterland vanishes once targeted sorting of individuals is accounted for.[5]

Another strand of literature argues that cities do not necessarily increase the quantity of social interactions but rather improve their quality / efficiency. In the model of Berliant et al. (2006), agents possess differentiated types of knowledge. The effect of cities on the number of social interactions then becomes twofold, as densely populated areas increase the number of random meetings but also make agents more selective regarding matching quality. Hence, while cities do not necessarily affect the number of social interactions, their quality in terms of knowledge complementarity should improve with increasing population density. Abel and Deitz (2015) study data on job searching of college graduates and find that larger and thicker labor markets indeed improve the matching between job advertisements and applicants' qualifications. To the best of our knowledge, no study to date has assessed the matching hypothesis with respect to social interactions. We formulate two tests, one relying on a network formation model, and the other analyzing the social adjustment process among movers. Both approaches show that urban residents indeed benefit from higher quality matches compared to people living in the hinterland.

Borrowing from the network literature, the level of clustering / triangular relations is an additional dimension of efficiency that is sometimes assumed to vary regionally. Granovetter (1973) famously argues that weak ties are often more valuable in terms of information provision than strong ties. He formally defines a weak tie as a social relation

[3] A recent study by Levy and Goldenberg (2014) uncovers similar patterns for email traffic and online social media contacts.

[4] Empirical studies by Charlot and Duranton (2004) and Schläpfer et al. (2014) support this hypothesis. However, neither can isolate the causal impact of density from non-random sorting.

[5] Burley (2015) shows for the German Socio-Economic Panel that population density is only positively correlated with an index of social interactions if person specific characteristics are ignored. Based on US survey data, Brueckner and Largey (2008) obtains negative correlations between density and social interactions. Other factors that have been shown to impact the level of social interactions are homeownership (Hilber, 2010) and racial fragmentation (Alesina and La Ferrara, 2000; Brueckner and Largey, 2008).

between two agents who have no overlap in their personal networks. In contrast, strong ties involve triangular relations that bring about redundancies in the process of information diffusion. Sato and Zenou (2015) claim that cities not only increase the number of social interactions – as discussed above – but also give rise to a disproportionally high number of weak tie relations that are more valuable in the job market. We find that personal networks in cities indeed tend to be characterized by lower levels of clustering and thus have a higher fraction of weak ties. This finding suggests that cities may facilitate the diffusion of information, although the average number of social interactions is not necessarily larger than in more sparsely populated areas.

The following section elaborates on the main concepts. Section 3 introduces the data while Section 4 explains the empirical identification strategy. Section 5 discusses the main results. Section 6 provides an assessment of the value of social interactions and Section 7 concludes.

2 Cities and Social Interactions: Main Concepts

We consider a directed network with N nodes each representing a unique phone customer which we denote by $i \in \mathcal{N} = \{1, ..., N\}$. Each customer has a place of residence, r, which is assigned either on the municipality or postcode level. The number of nodes at location r is N_r, and so with R denoting the total number of different residences, $N = \sum_r N_r$ holds. Finally, $\mathcal{N}_r$ is the set of individuals living in location r.

A link between nodes i and j is denoted by $g_{ij} = 1$, while the absence of a link is marked as $g_{ij} = 0$. The network can then be characterized by a pair $(\mathcal{N}, \mathcal{G})$ where $\mathcal{G} = [g_{ij}]$ is a $N \times N$ adjacency matrix. We assume that rational agents i and j establish a link if the net surplus from doing so is positive (c.f. Graham, 2014). This yields a random utility model of the form

$$g_{ij} = \mathbf{1}\left(X_{ij}'\eta + \nu_i + \nu_j + U_{ij} \geq 0\right), \tag{1}$$

where X_{ij} is a vector of dyad attributes (i.e. pair specific characteristics), ν_i and ν_j denote agent specific characteristics, and U_{ij} is a randomly distributed component of link surplus. We are particularly interested in the role of dyad attributes, which we divide into three groups: geographical distance or travel time (T_{ij}), the number of friends i and j share ($F_{ij} = \sum_{k=1}^{N} g_{ik}g_{jk}$), and matching ($m(\nu_i, \nu_j, \delta)$). Higher levels of $m(\cdot)$ increase link surplus, which is why we refer to it as matching *quality*. In particular $m(\cdot)$ absorbs the spread between Q individual characteristics of agent i and j, $|\nu_i - \nu_j|$, which – depending on the specific attribute $q \in Q$ – may be positively (i.e. $\delta_q > 0$) or negatively correlated

(i.e. $\delta_q < 0$) with matching quality. Based on these considerations we define the vector X_{ij} as

$$X'_{ij}\eta = \eta_1 \cdot T_{ij} + \eta_2 \cdot F_{ij}(\mathcal{G}) + \eta_3 \cdot m(\nu_i, \nu_j, \delta). \tag{2}$$

If link-surplus is indeed a function of these three dyad-specific factors, this may have important consequences for the network topography across rural and urban areas. Provided that distance is costly for social interactions, regional differences in population density may determine the *size* of an agent's social network. This is of interest, because social contacts can foster the diffusion of information, promote trust and thereby lower transaction costs, and facilitate learning from peers (Granovetter, 2005; Gui and Sugden, 2005; Jackson, 2014) in addition to having intrinsic value for a person's well-being (Burt, 1987). We further focus on matching and common friends (or clustering), as they have implications for a *network's efficiency*: Matching reflects the quality of a specific contact, which incorporates various dimensions such as productivity enhancing skill complementarity, or shared interests (e.g. Berliant et al., 2006). Clustering governs the informational value of a link, since contacts who share a common friend introduce redundancies and are therefore less valuable in the information diffusion process (Granovetter, 1973). In return, sharing mutual contacts fosters cooperative and pro-social behaviour, because the triangular relation can act as a reputational control and retaliation device (Jackson, 2014).

Network Size and Degree Centrality. The size of an individuals social network, which we measure based on *degree centrality*, is formally defined as

$$D_i(\mathcal{G}) = \#\{j : g_{ij} = 1\}. \tag{3}$$

Degree centrality reflects the number of distinct peers with whom agent i interacts socially and therefore the number of sources that potentially forward valuable information. Typically, urban economic theory presumes that cities provide a favourable environment for social interactions and support larger network sizes than rural communities. The underlying argument hinges on the assumption that the costs of social interactions increases with distance. Let us abstract from the matching spread, $m(\nu_i, \nu_j, \delta)$, as well as triangular ties, $F_{ij}(\mathcal{G})$, and focus on the relationship between distance and population density. A stylized argument is as follows: On weekdays an agent i needs to keep her travelling costs low, and she therefore has random encounters only with people in her municipality, $j \in \mathcal{N}_r$. At the weekend, however, the radius of the agent's actions is unbounded, so that she might form ties with people living outside her place of residence, $k \notin \mathcal{N}_r$. Since people spend more time in their residence's vicinity, the probability to acquire social contacts among neighbors, $P_r = P(g_{i,j\in\mathcal{N}_r} = 1)$, is larger than for the rest of the population, that

is $P_r > P_{-r} = P(g_{i,k \notin \mathcal{N}_r} = 1)$. In the outlined example, the size of a person's social network positively depends on the population living in the neighborhood, N_r, so that cities support a larger degree than rural municipalities, i.e.

$$D_i = N_r \cdot P_r + (N - N_r) \cdot P_{-r} \qquad \text{with } \frac{\partial D_i}{\partial N_r} > 0. \tag{4}$$

While this intuitive rationale is appealing, it may be challenged from two angles, namely from biological/anthropological and search strategic points of view.

In evolutionary biology, Dunbar (1992) has famously advocated and popularized the *social brain hypothesis*.[6] According to this hypothesis there is an upper limit to group size that is set purely by cognitive constraints. For humans, Dunbar (1993) calculates the upper limit to lie between 100 and 230 social contacts, citing anthropological studies on modern hunter-gatherer societies as evidence that support his prediction.[7] In consideration of the manifold results corroborating the social brain hypothesis, one may note that the size of a person's network is fundamentally restricted by congenital factors. Because the population of practically all Swiss municipalities exceeds the range for the limit of network size as calculated in the literature, the number of social interactions may be independent of regional differences in population density.

In equation (4) a random encounter between two persons is equivalent to establishing a link. We now add another layer: After meeting a potential contact, agents can either accept or reject to form a link based on the other person's characteristics. Since forming a link consumes time and cognitive capacity, this introduces a quality-quantity trade-off. Consequently, it may be optimal to reject some potential contacts to wait for a better match. Hence, from a *search strategic* perspective, higher population density may impact network size only marginally, but it may allow for higher selectivity along dyad-specific characteristics. This has important consequences for the analysis of social networks across different regions. Even if densely populated areas improve social networks, the advantages may not be in terms of size but in terms of efficiency. In this respect, *matching quality* between agents i and j, $m(\nu_i, \nu_j, \delta)$, is of key interest, as it determines how well their interests correspond or how fruitful the intellectual exchange between them is. Once we add the strategic component of weighing between quality and quantity to the above mechanics, we would expect a positive effect of population density on matching quality, or network size, or both.

[6] The hypothesis challenges the field's traditionally dominant view that brains evolved to address ecological problem-solving tasks, such as foraging. Instead the social brain hypothesis attributes the growth in primates' brain sizes to the computational demands of their increasingly complex social systems.

[7] Recent studies explore this hypothesis by analyzing patterns among adults' brains, cognitive ability, and the size of their social networks (Powell et al., 2012; Stiller and Dunbar, 2007) or by exploiting social media user statistics (Dunbar, 2015).

Perimeter of Social Interactions and Within-Degree. The previous line of reasoning also has implications for the perimeter of a person's network. If distance is costly when maintaining a link, one would rather form a tie with a neighbor than with an identical person living far away. This implies that cities allow people to be more selective regarding the travel distance to their contacts. Put differently, one may expect that urban residents can recruit their contacts within a relatively narrow perimeter whereas rural residents prefer to widen the search radius with the objective of improving their network's quality. To analyze these claims, we examine the degree within an individual's neighborhood or *within-degree*, formally defined as

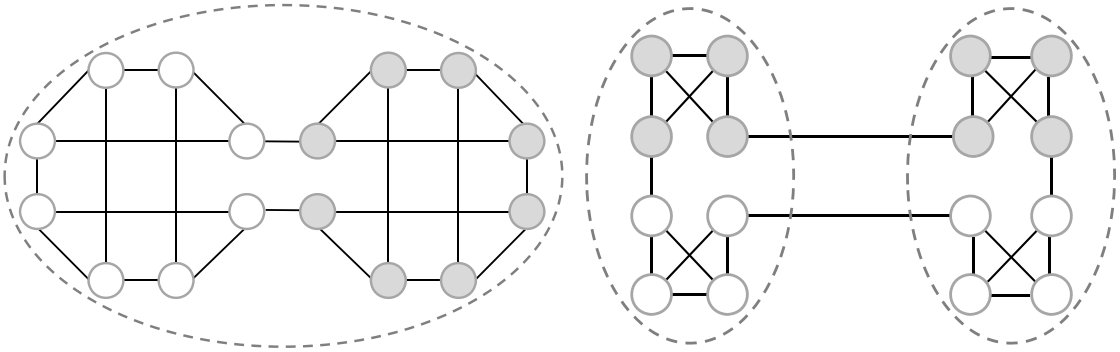
$$DW_i^r(\mathcal{G}) = \#\{i, j \in \mathcal{N}_r : g_{ij} = 1\}. \tag{5}$$

Of course, negligible distance costs would wipe out any differences between cities and rural areas. Costs related to distance may indeed be of secondary importance for a person with naturally few social interactions, whereas highly sociable persons may benefit more from densely populated areas, as recently formalized in a paper by Helsley and Zenou (2014). Consequently, differences in network size may simply be observable due to the sorting of highly sociable types into cities, because they gain disproportionally from low distance costs per contact.

Clustering. Clustering is an important network characteristic as it can provide insights into reciprocity and information diffusion. On the one hand, high clustering strengthens reputational concerns and with it the enforcement of social norms and cooperation (e.g. Ali and Miller, 2009), or risk-sharing (e.g. Ambrus et al., 2014). On the other hand, Granovetter (1973) highlights the importance of local bridges for passing on information. An individual with high clustering introduces redundancies in the network, which are inefficient in terms of information diffusion. The *clustering coefficient* for node i is given by

$$C_i = \frac{\sum_{j,k\, j\neq k} g_{jk}}{\sum_{j,k,j\neq k} g_{ij} g_{ik}}, \tag{6}$$

and measures whether an individual's contacts form a tightly knit group ($C_i \to 1$) or are completely separate from each other ($C_i \to 0$). How does population density relate to clustering? There are two potential channels, one mechanical and the other as a consequence of differing preferences. Figure 1 illustrates the mechanical rationale: Panel (a) shows a city with 16 agents, eight white and eight grey. All agents socially interact with three other agents, preferably of the same type. Panel (b) represents a peripheral region with lower population density, therefore the 16 agents are equally split between two municipalities. As in the city, all individuals have a degree of three. Importantly,



(a) **City:** Average Degree=3, Matching Rate=0.833, Average Clustering=0, Average Path Length=2.73

(b) **Periphery:** Average Degree=3, Matching Rate=0.833, Average Clustering=0.5, Average Path Length=3.2

Figure 1: Clustering in Cities and the Periphery – An Illustrative Example

travelling between the two municipalities is costly, therefore agents prefer to form links with their neighbors. Since every person has only three neighbors of the same type, the network ends up tightly clustered. In contrast, the city makes clustering less likely, because each urbanite can choose among seven agents of the same colour.[8] Thus, low density locations should tend to display higher clustering, simply because residents of these areas face a substantially smaller set of suitable contacts in their direct vicinity compared to urban residents. In addition to this purely probabilistic relation between density and clustering, incentives for forming links with friends of friends, $F_{ij}(\mathcal{G})$, could be different in cities than in rural areas. Agents face a trade-off in terms of efficient information exchange (i.e. low clustering) and benefits due to stronger reciprocity (i.e. high clustering). The optimal balance may vary regionally due to factors that assign a higher weight to reciprocity or information diffusion. For instance, high quality local institutions may substitute for reciprocity or a dynamic labor market environment may support the value of information diffusion. In addition, clusters may facilitate simultaneous interactions with multiple persons, allowing for larger networks given a certain time constraint. If people living in rural neighborhoods have more geographically dispersed social networks, clusters of friends could be a strategy to mitigate travel costs.[9]

[8] In the way the example is drawn, the average clustering in the city equals 0, while it amounts to 0.5 in the periphery. As a consequence, the average path length in the city (=2.73) is lower than in the periphery (=3.2), which accelerates the diffusion of information.

[9] For instance, Fischer (1982) documents that people living in peripheral areas have a higher proportion of kin ties than urban residents, which is likely to increases the clustering in an individual's network.

3 Data

The main dataset used in this paper is provided by Switzerland's largest telecommunications operator, *Swisscom AG*, whose market share is about 55% for mobile phones and about 60% for landlines (ComCom, 2015). The data comprises comprehensive *call detail records (CDR)* of all calls made and received by the operator's customers between June 2015 and May 2016. The CDRs include the anonymized phone number of caller and callee, a date and time stamp, a binary indicator for private and business customers, a code for the type of interaction recorded (e.g. call, SMS, MMS), the duration of calls in seconds, and the x-y-coordinates of the caller's main transmitting antenna. We observe finely grained information on about 15 million calls and text messages per day, covering about 9.1 million phones, of which 4.1 million are mobile phones and 2.7 million are private mobile phones.[10] Along with the anonymized CDRs, the operator also provided monthly updated *customer information* including billing address, language of correspondence (German, French, Italian, English), age, and gender. Table 1 summarizes the socio-demographic characteristics of mobile phone customers in our sample, while Table A.3 shows correlations between census data and our customer statistics for various subpopulations. This comparison suggests that the data at hand is highly representative of the Swiss population even at very local level.

The phone data are complemented by various municipal statistics provided by the Federal Statistical Office (FSO), including population figures and the degree of urbanisation as classified by EUROSTAT.[11] Figure 2 shows the regional variation in urbanisation based on the aforementioned measure. We also compute geographical distances between pairs of municipalities and pairs of postcodes using GIS software and shape files for administrative boundaries published by the Federal Office of Topography. Car driving distances (between centroids of municipalities/postcodes) and public transport travel times (for all existing pairs of stops) were obtained from *search.ch*. Descriptive statistics for all $2,322$ Swiss municipalities and $3,201$ postcodes are shown in Table A.1 in the appendix.

The anonymity of Swisscom customers was guaranteed at all steps of the analysis. We never dealt with or had access to uncensored data. A data security specialist retrieved the CDRs from the operator's database and anonymized the telephone numbers using a 64-bit hash algorithm that preserved the international and local area codes. He further removed columns with information on the transmitting antenna before making the data available. Once the anonymized data were copied to a fully sealed and encrypted Swisscom workstation, we ran the analysis on site. To utilize information on the transmitting

[10] More specifically, the data set covers 2.7 million private mobile phones, 2.1 million private land lines, 1.4 million corporate mobile phones, and 2.9 million corporate landlines.

[11] See `http://ec.europa.eu/eurostat/ramon/miscellaneous/index.cfm?TargetUrl=DSP_DEGURBA` (last access: 01.06.2016) for more information on the EUROSTAT DEGURBA measure.

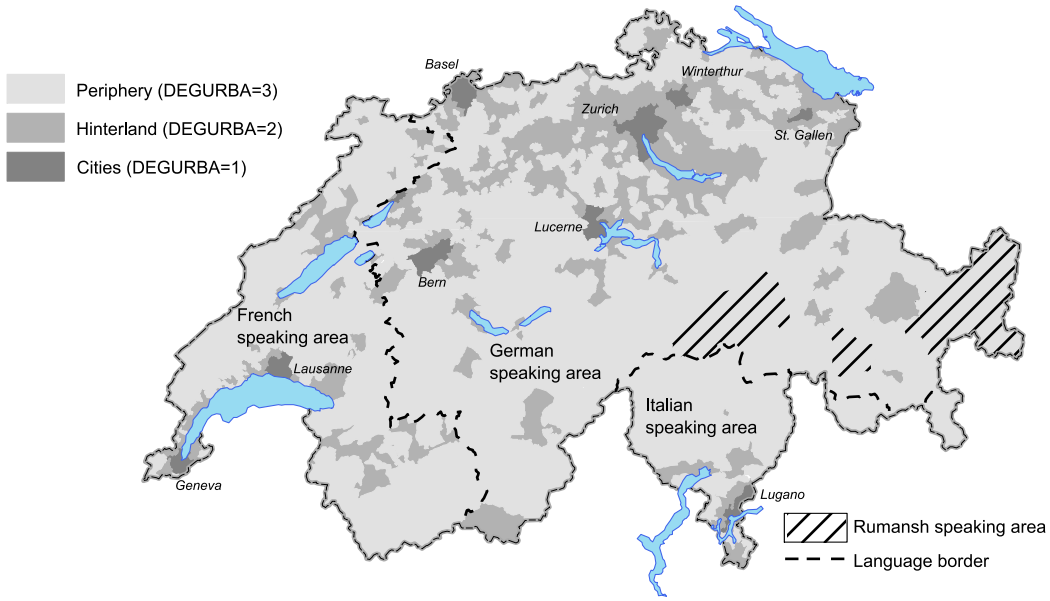


Figure 2: Degree of Urbanisation – Cities, Hinterland and Periphery

antenna we passed location scripts to Swisscom personnel who executed them for us.

Our primary aim is to observe social networks, but not every instance of phone activity reflects a social interaction in the narrower sense so that the dataset needs to be cleaned beforehand.[12] In our benchmark analysis, we filter the data as follows: *First*, we restrict the analysis to *calls* between mobile phones. Mobile phones are personal objects and are thus representative of the social network of a single person, while calls from landlines possibly resemble overlapping social networks as they are usually shared by multiple users. For the same reason, all results are based on customers who have registered only one active mobile phone number. Customers with multiple active numbers typically include corporate customers, as well as parents acting as invoice recipients for their children. *Second*, we limit the analysis to outgoing calls in order to cover intra-operator and inter-operator activity equally well and to filter out promotional calls by call centres. *Third*, calls with a duration of less than 10 seconds are considered accidental and are therefore excluded from the analysis. *Fourth*, we drop mobile phone numbers that display implausibly low or high monthly usage statistics, with a minimum threshold of 1 minute and a maximum threshold of 56 hours per month, respectively. This removes practically inactive numbers as well as phones used for commercial purposes. *Fifth*, the analysis is limited to *private mobile phones*, so that daily business calls between corporate customers do not create noise in our measures. *Sixth*, some measures require address information for both caller and callee such that inter-operator calls cannot be used in all steps of the analysis. Measures requiring

[12]For a discussion see Blondel et al. (2015).

Table 1: Descriptive Statistics, Private Mobile Phone Customers

	Mean	SD	N	Min	Max
Monthly Phone Usage, June 2015 – May 2016 (pooled)					
Number of Calls	111.781	109.599	10 399 549	1	10 113
Duration (Minutes)	254.970	295.609	10 399 549	2	3359
Monthly Network Characteristics, June 2015 – May 2016 (pooled)					
Degree Centrality	9.202	7.910	10 399 549	1	470
Within-Degree (15 Min. Radius)	7.067	7.231	10 399 549	0	221
Clustering Coefficient	0.092	0.132	10 248 923	0	1
Sociodemographics					
Age	34.964	13.561	866 646	20	60
Female	0.522	–	866 646	0	1
Language: German	0.681	–	866 646	0	1
Language: French	0.270	–	866 646	0	1
Language: Italian	0.043	–	866 646	0	1
Language: English	0.006	–	866 646	0	1

Notes: The table is based on the subsample of customers with phone activity in all 12 months, which we also use in the main analysis. Further filters as described in Section 3. Phone usage statistics include in- and outgoing calls. The *within-degree* measures network size within a radius of 15 minutes around an agent's residence.

location information for the callee are therefore based on intra-operator calls only, which we weight according to the operator's market share at the callee's billing address. *Finally*, we only use the first 28 days of each month to make the data easily comparable across different time periods.

These steps eliminate approximately 60 percent of the total number of calls recorded, leaving us with around 60 million calls per month that amount to a total duration of 200 million minutes (for details see Table A.2 in the appendix).

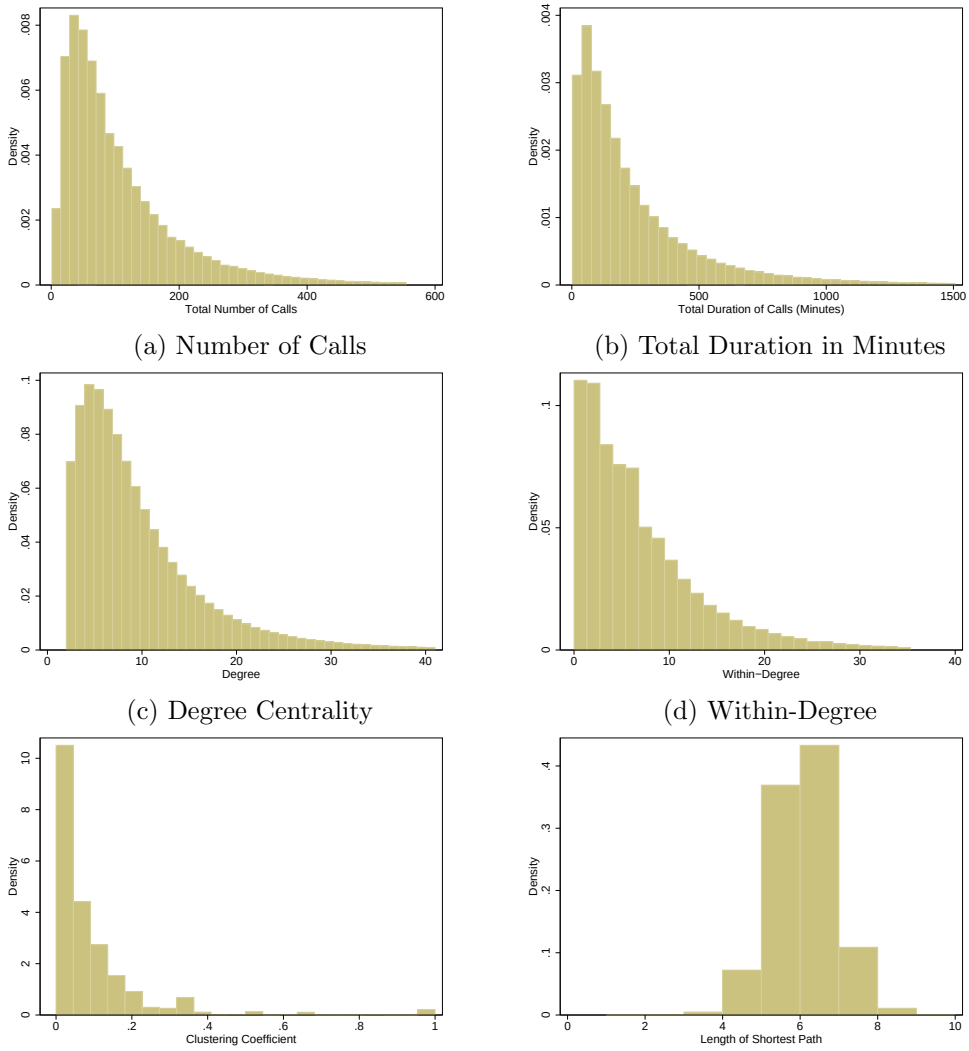
3.1 Descriptive Statistics on Phone Usage and the Social Network

Table 1 shows summary statistics on the mobile phone usage of customers aged 15 to 64 for the filtered data set.[13] The average private mobile-phone users makes about four calls per day with a cumulative duration of nine minutes. Figures 3a and 3b further show that the distributions are markedly right-skewed.

The network of private mobile phone interactions uncovered by the data exhibits characteristic features of other socially generated networks documented in the literature (Jackson and Rogers, 2007; Watts, 1999): Small diameter and short average path length between pairs, "fat tails" in the degree distribution, and substantial clustering.

To gain insights into the diameter and the average path length, we randomly select 100

[13]Due to privacy concerns, we worked with decimal age-brackets. This means that a customer aged 24 was assigned to the 20-bracket, while a customer aged 25 belongs to the 30-bracket.



(a) Number of Calls

(b) Total Duration in Minutes

(c) Degree Centrality

(d) Within-Degree

(e) Clustering Coefficient

(f) Dyad Distances

Notes: All subfigures are based on data that was filtered as described in Section 3. *(d) Within-Degree:* Number of contacts that reside within 15 min. driving time of an agent's residence. *(f) Dyad Distances:* Length of shortest paths connecting 100 randomly selected agents with every other private mobile phone user in the data.

Figure 3: Histograms of Phone Usage Statistics & Network Characteristics for June 2015

individuals and calculate the length of the shortest paths connecting every other private mobile phone user in the data. The mean path length in the sample is 5.6, with the longest path having a length of 12; the histogram plotted in Figure 3f reveals that 88 percent of dyads are separated by 6 or fewer links. This fits strikingly well with the "small-world"-hypothesis first formulated by Milgram (1967) and the early empirical evidence based on a chain letter experiment conducted by Travers and Milgram (1969).

As Figure 3c illustrates, the degree centrality distribution in our social network exhibits "fat tails", so that there are more nodes with relatively high and low degrees, and fewer

nodes with medium degrees, than one would find in a network where links are formed uniformly. The average degree centrality in our monthly data is approximately 9, with the vast majority having a degree below 20 and some hub-agents reaching network sizes of 100 links or more. As reported in other studies on social networks, the probability distribution of degree centrality is well fitted ($R^2 = 0.92$) by a power-distribution, $P(D) = cD^{-\varphi}$, with parameter estimates of $\hat{\varphi} = 3.86$ and $\hat{c} = 5.96$.

The clustering coefficient, which measures the tendency of linked nodes to have common neighbors, is, on average, 0.092, with more than 75 percent of the individuals in the dataset having a non-zero clustering coefficient (see Figure 3e). Considering the low network density in our data (≈ 0.00001), the observed clustering is evidently larger than in a benchmark network where links would have been generated by an independent random process.[14]

4 Identification

In order to analyze the impact of geography and location characteristics on the structure of social interactions we conduct two complementary identification strategies. The first aims to identify factors that predict the likelihood of individuals i and j forming a link and is referred to as *network formation*. In particular, this approach allows us to study the effects of distance between i's and j's place of residence on the probability that they form a link. It further enables inference on the preference for triadic relations. The presence of network overlap may influence the likelihood that i and j establish a link as the returns may be higher or lower if it involves mutual contacts. Moreover, we study whether homophily – the process of matching on common characteristics – is prevalent in the data.

The second approach, to which we refer as *network topography*, estimates the effect of local characteristics on individual-level network measures. This relates to the equilibrium outcome of network formation at different places and allows us to examine the impact of location specific attributes on the structure of social networks.

Sorting of individuals with specific characteristics can affect the results of both approaches. We address this issue by exploiting changes in the federal railway timetable to infer the causal impact of travel time on link formation and by analysing social network adjustments of movers. Studying movers allows us to obtain causal estimates of geography and population density on network topography and network formation.

[14] Note that network density is defined as the ratio of actually formed links and potential links, which is equivalent to dividing the mean degree (9.2) by the number of potential contacts of an individual (866,645).

4.1 Network Formation

We observe the social network's adjacency matrix $\mathcal{G}_t = [g_{ij,t}]$ in each month $t \in \{1, ..., 12\}$. Based on equations (1) and (2), we specify the probability that two nodes i and j form a link as

$$g_{ij,t} = \mathbf{1}(\beta g_{ij,t-1} + T'_{ij,t}\eta_1 + F'_{ij,t-1}\eta_2 + Z'_{ij}\rho + \phi_1 D_i + \phi_2 D_j + m(\xi_i, \xi_j, \delta) + U_{ij,t} \geq 0) \quad (7)$$

where vector $T_{ij,t}$ measures the distance between agent i and j based on their residence and workplace, $F_{ij,t-1}$ is a vector reflecting the number of contacts agents i and j share in common, Z_{ij} is a vector of dyad-specific time invariant covariates, D_i and D_j capture static differences in sociability based on both parties' logarithmized long-term degree, and $m(\xi_i, \xi_j, \delta)$ is a symmetric matching function of unobserved node specific heterogeneity.[15] We assume that $U_{ij,t}$ is independent and identically distributed and has mean zero such that we can estimate a linear probability model of the form:

$$g_{ij,t} = \beta g_{ij,t-1} + T'_{ij,t}\eta_1 + F'_{ij,t-1}\eta_2 + Z'_{ij}\rho + \phi_1 D_i + \phi_2 D_j + m(\xi_i, \xi_j, \delta) + U_{ij,t}. \quad (8)$$

In particular, the distance measures represented by vector $T_{ij,t}$ comprise the log travel time between agents i's and j's residence as well as a dummy for same workplace.[16] The latter equals one if they predominantly use antennas within the same 5km radius during business hours. We discretize the number of common friends, such that we obtain two dummy variables contained in $F_{ij,t-1}$: The first indicator equals one, if agents i and j share at least one common social contact, while the second indicators equals one if agents i and j share at least two common contacts.[17] The dyad-specific covariates in vector Z_{ij} include three dummy variables indicating same age, same gender and same language.

The model in (8) also accounts for matching based on unobservables as reflected by $m(\xi_i, \xi_j, \delta)$. Those that favourably match in terms of unobservable characteristics ξ feature a higher likelihood to form a link. These unobservables may bias our estimates of the cross-sectional model. If individuals with common unobservable attributes are more likely to cluster regionally and thus live closer together, our distance measure will be negatively correlated with the error term. A within-transformation will take out time invariant factors

[15] Note that the number of mutual contacts, $F_{ij,t-1}$, enters with a lag. This implies that agents form/maintain/dissolve links myopically, as if all features of the previous period's network remain fixed. Assuming this structure, eliminates contemporaneous feedback, which can be problematic for inference (see Graham, 2014).

[16] We have estimated the models also with geographical distance instead of travel time which does not qualitatively affect our results. However, due to the rugged environment in Switzerland we consider travel time as the more relevant measure.

[17] We discretize the number of mutual friends, because the continuous measure yields imprecise (yet significant) estimates. Sensitivity checks showed diminishing effects of mutual friends as mutual friends beyond two did not significantly add to the link likelihood.

that affect the matching quality, i.e.

$$\ddot{g}_{ij,t} = \beta \ddot{g}_{ij,t-1} + \ddot{T}_{ij,t}\eta_1 + \ddot{F}'_{ij,t-1}\eta_2 + \ddot{U}_{ij,t}, \tag{9}$$

where we define the within transformation for generic variable x by $\ddot{x}_t = x_t - \bar{x}$. The transformed residual, $\ddot{U}_{ij,t}$, is necessarily correlated with the lagged dependent variable, $\ddot{g}_{ij,t-1}$, because both are a function of $\bar{U}_{ij}$. Thus, OLS estimates of equation (9) are not consistent for the parameters of interest. We therefore follow Angrist and Pischke (2009) and estimate models including the lagged dependent variable but not the fixed effects, as in equation (10a.), and then compare the results to estimates obtained from a fixed effect regression without the dynamic component, as in equation (10b.):

$$\begin{aligned} &\text{a.} \quad g_{ij,t} = \beta g_{ij,t-1} + T'_{ij,t}\eta_1 + F'_{ij,t-1}\eta_2 + Z'_{ij}\rho + \phi_1 D_i + \phi_2 D_j + U_{ij,t} \\ &\text{b.} \quad \ddot{g}_{ij,t} = \ddot{T}_{ij,t}\eta_1 + \ddot{F}'_{ij,t-1}\eta_2 + \ddot{U}_{ij,t}. \end{aligned} \tag{10}$$

These two models have a useful bracketing property, that bounds the causal effect of interest. With respect to the geographical distance between two agents, we expect that the fixed effect estimates are upwardly biased, while the lagged dependent model yields a downwardly biased estimate (see Guryan, 2001; Angrist and Pischke, 2009). We also estimate equation (10a.) within a Logit framework in order to account for the dichotomous nature of the data.

A practical issue that arises with estimating the outlined network formation models is the size of the adjacency matrix that potentially includes $(2 \cdot 10^6)^2$ unique pairs of agents. It is neither computationally feasible to estimate the models based on all these pairs nor necessary for obtaining consistent estimates of the parameters of interest as is shown by Manski and Lerman (1977), and Cosslett (1981). Since we have complete information on the network we can use a stratified sample and adjust the estimates with the respective sampling weights. Our choice-based sample results from an endogenous stratified sampling scheme where each stratum is defined according to the individual responses, that is the binary values taken by the response variable $g_{ij,t}$.[18] This sampling structure requires the availability of prior information on the marginal response probabilities which is in our setting available due to the full observation of $\mathcal{G}_t$.

[18] The main motivation behind this approach is usually the possibility of oversampling rare alternatives, which can improve the accuracy of the econometric analysis but also reduce survey costs. However, in our case we undersample those dyads with $g_{ij,t} = 0$ in order to enhance computational efficiency.

4.2 Network Topography

We estimate the effect of location characteristics on the individual-level network measures formally defined in Section 2: degree, within-degree, and clustering coefficient. Below, we lay out the estimation strategy for degree centrality noting that specifications for all other network measures follow analogously.

Following the earlier notation, the econometric models involve measures of degree centrality, D_{it}, as dependent variable and location specific covariates at the place of residence denoted by L_r. Hence, we specify the benchmark model as

$$D_{ir,t} = \alpha + L'_{r,t}\beta + X'_{ir,t}\gamma + \lambda_t + \lambda^l_r + \epsilon_{ir,t}, \tag{11}$$

where $X_{ir,t}$ is a vector of individual characteristics (i.e. commuting distance, language, dummy for belonging to language minority, gender, and age), λ_t stands for month fixed effects, and λ^l_r denotes language region fixed effects. The location vector $L_{r,t}$ includes indicators for EUROSTAT's harmonized definition of functional urban areas which distinguish between the urban core, the hinterland and peripheral regions. Alternatively, we measure local density using the number of private mobile phone customers within 15 minutes travel time from the respective place. Unlike municipal population statistics this measure has the advantage that it is independent from administrative boundaries. Yet, all results are robust to using municipal population density.

In a next step, we address the issue of individual sorting on unobservables across locations. If the most sociable individuals systematically sort into high-density places, equation (11) would yield upwardly biased estimates of the density externality. Compared to the pooled OLS specification, we add an individual fixed effect in order to disentangle the density externality and the sorting effect, i.e.

$$D_{ir,t} = \mu_i + L'_{r,t}\beta + X'_{ir,t}\gamma + \lambda_t + \lambda^l_r + \epsilon_{ir,t}. \tag{12}$$

Note that this model identifies the effects on the basis of movers i.e. those who changed their place of residence between July 2015 and April 2016. These are about 147'000 individuals in the unfiltered data or 6% of the operator's private customers (see Table A.4). One concern in introducing fixed effects is that movers may differ systematically from the population. Like reported in other studies that adopt a similar identification strategy (e.g. D'Costa and Overman, 2014), movers in our data are on average younger than non-movers. Apart from age, Table A.5 shows that differences in both individual characteristics as well as phone usage behaviour and network properties are sufficiently small between both groups. Apart from sorting there may be unobserved location factors that affect network characteristics as well as population density and thereby potentially bias our estimates.

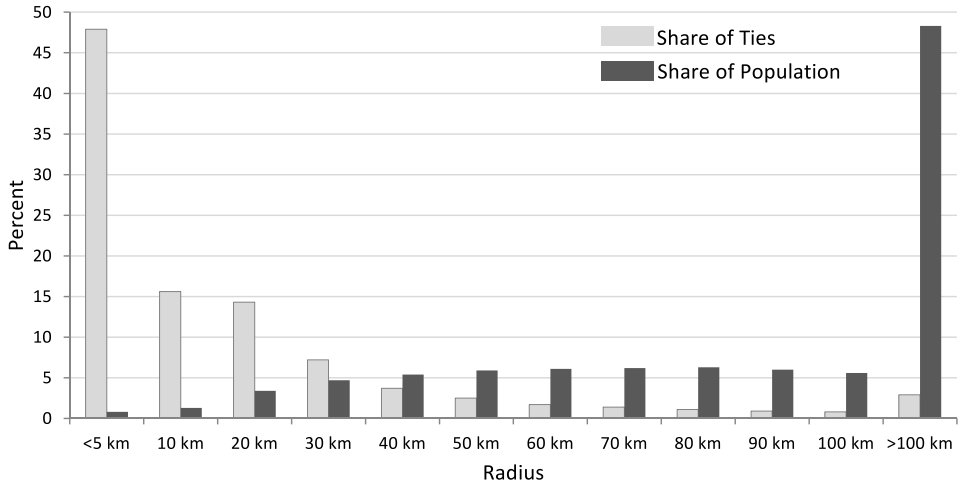
We address this concern by instrumenting population density using historical population counts and measures of soil quality (see Ciccone and Hall, 1996; Combes et al., 2010). All our results remain valid when using the instrumental variable approach which we relegate to Appendix B.8.

5 Results

In the following we present the main results for the network formation model. Our focus is on the question of whether distance is costly to social interactions. In a second step, we analyze differences in network size across regions, to test the hypothesis that cities promote social interactions. We then proceed to investigate, whether population density affects the efficiency of networks in terms of perimeter, matching quality, and clustering.

5.1 Death of Distance? The Role of Proximity in Network Formation

It is instructive to begin by looking at plain descriptives. Figure 4 plots the share of ties along the share of potential contacts by radius. Considering that almost 50 percent of bilateral ties are formed within a 5km perimeter that covers on average less than 1 percent of the population, this illustrates the rapid decline of social interactions across space.



Notes: The share of ties reflect mobile phone calls made in June 2015. The radius is calculated based on the distance between the caller's and callee's place of residence. Population statistics comprise number of mobile phone customer's by postcode.

Figure 4: Share of Social Ties and Population by Radius

Of course, this approach does not account for biases due to spatial sorting of similar types. We therefore proceed to the network formation models, outlined in the previous section. Table 2 presents the result for the linear probability model (LPM) specified in

equations (10a.) and (10b.). All coefficients were multiplied by 10'000 and therefore can be interpreted as basis points. This means that a coefficient equalling one translates to a marginal increase in $P(g_{ij,t} = 1)$ of a hundredth percentage point. The first two columns display pooled OLS estimations, the middle columns report pair fixed effects models, and the last two columns show lagged dependent variable specifications. In all models estimated, the travel time between two agents enters negatively, implying that *distance* is indeed costly when forming and maintaining a link. Columns (2), (4) and (6) reveal that tie formation is actually a convex function in distance; the log of travel time enters strongly negative, while the squared term is positive. Their relative magnitudes suggest that the negative effect of distance completely fades at approximately 90 minutes driving distance.

In addition to being neighbors, *working in the same area* also increases the likelihood that two persons form a link. The coefficient for the dummy variable "Same Workplace", which equals one if agents i and j predominantly use antennas within the same 5 km radius during business hours, ranges between 0.07 and 0.1. Hence, working in close proximity increases the probability of forming a tie by about 0.1 basis point, which is roughly ten times the estimated effect of speaking the same principal language. Thus, distance in terms of both residence and workplace are very costly to social interactions.

In order to analyze preferences for triadic closure or *clustering*, we discretize the number of common friends, such that we obtain two dummy variables: one indicating that agents i and j share at least one common social contact, and the other indicating that they share at least two common contacts. The coefficients for both "Common Contact" variables are highly significant. Column (2) shows that the probability of forming a link with another person increases by up to 22 percentage points, if one shares at least two common contacts. As one would expect, the estimates are considerably smaller in column (4), which controls for matching quality by employing dyad-specific fixed effects. Nonetheless, the additional link-surplus of 1.5 percentage points due to triangular relations – as obtained in the most conservative specification – is quantitatively substantial. Agents clearly value triadic relations, which explains the evidently non-random clustering in this network, as discussed in Section 3.1.

Overall *matching quality* between two agents is not directly observable, but the regressions in column (2) and column (6) account for socio-demographic (dis)similarities that are incorporated in the matching concept, namely dummies for same language, same gender and same age, as well as the absolute age difference between customers i and j. If we assume that $m(\cdot)$ is a linear and additive function, the interpretation of the estimated coefficients in terms of matching is as follows: By definition $\frac{\partial E[g_{ij}|m(\cdot)]}{\partial m(\cdot)} > 0$, therefore $sign(\hat{\rho_q}) = sign(\delta_q)$ holds. Accordingly, a positive (negative) sign not only implies an increase in the probability that two agents socially interact, but also a positive (negative)

Table 2: Network Formation

	Pooled OLS		Panel FE		Lagged Dependent Var.	
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Travel Time$_{ij,t}$)	-0.112***	-0.942***	-0.024***	-0.094***	-0.053***	-0.479***
	(0.000)	(0.053)	(0.000)	(0.019)	(0.000)	(0.024)
Ln(Travel Time$_{ij,t}$)^2		0.104***		0.010***		0.053***
		(0.006)		(0.002)		(0.003)
Same Workplace$_{ij,t}$		0.166***		0.071***		0.100***
		(0.030)		(0.002)		(0.014)
Same Language$_{ij,t}$		0.017***				0.009***
		(0.001)				(0.001)
>0 Common Contacts$_{ij,t-1}$		213.822***		11.840***		100.943***
		(10.101)		(0.928)		(4.866)
>1 Common Contacts$_{ij,t-1}$		2257.176***		145.633***		1024.429***
		(331.296)		(35.656)		(159.448)
g$_{ij,t-1}$					5231.433***	4973.641***
					(2.929)	(34.689)
Const.	0.545***	2.079***	0.135***	0.224***	0.256***	1.060***
	(0.001)	(0.114)	(0.002)	(0.038)	(0.000)	(0.052)
R^2	0.001	0.054	0.115	0.115	0.275	0.288
Further Controls	No	Yes	No	No	No	Yes
Pair FE	No	No	Yes	Yes	No	No
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Groups	–	–	2,584,869	2,582,702	–	–
Observations	30,996,082	27,238,673	30,996,082	27,238,673	28,411,817	27,238,673

Notes: We use monthly data for June 2015–May 2016. The *sample* covers movers who used their phone every month at least once. All *coefficients* of the linear probability models are multiplied by 10000, and therefore can be interpreted as basis points. *Further controls* include the degree for both agents (log), dummies for same gender and same age, as well as the absolute age difference between agents i and j. Standard errors in parentheses.
+ p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

relation in terms of matching quality. Our results unambiguously point toward homophily, which is the well documented tendency of individuals to bond with similar others (e.g. Currarini et al., 2009; McPherson et al., 2001). For instance, individuals who share the same principal language are on average more likely to form a tie than individuals with different language preferences. The same holds true for age and gender (results not shown).

The LPM results suggest that spatial proximity, the presence of common friends, and demographic similarity increase the likelihood that two individuals interact. We also estimate *Logit models* to accommodate for the binary dependent variable and check the robustness of these results. The non-linear estimates are presented in Table B.1 in the appendix and are qualitatively almost identical to the LPM results (see Figures 5a, 5b). In order to allow proximity to enter more flexibly we replace the linear/quadratic distance functions by a series of dummies for distances within 5min, 10min,...,60min. The corresponding results support the convex relationship (see Table B.2 in the appendix).

Until now, identification of the causal impact of distance on link formation rested on the assumption that matching quality is time constant, so that the issue of non-random sorting can be eliminated by analyzing movers. In a next step, we exploit a *natural ex-*

Table 3: Changes in Public Transport & Network Formation

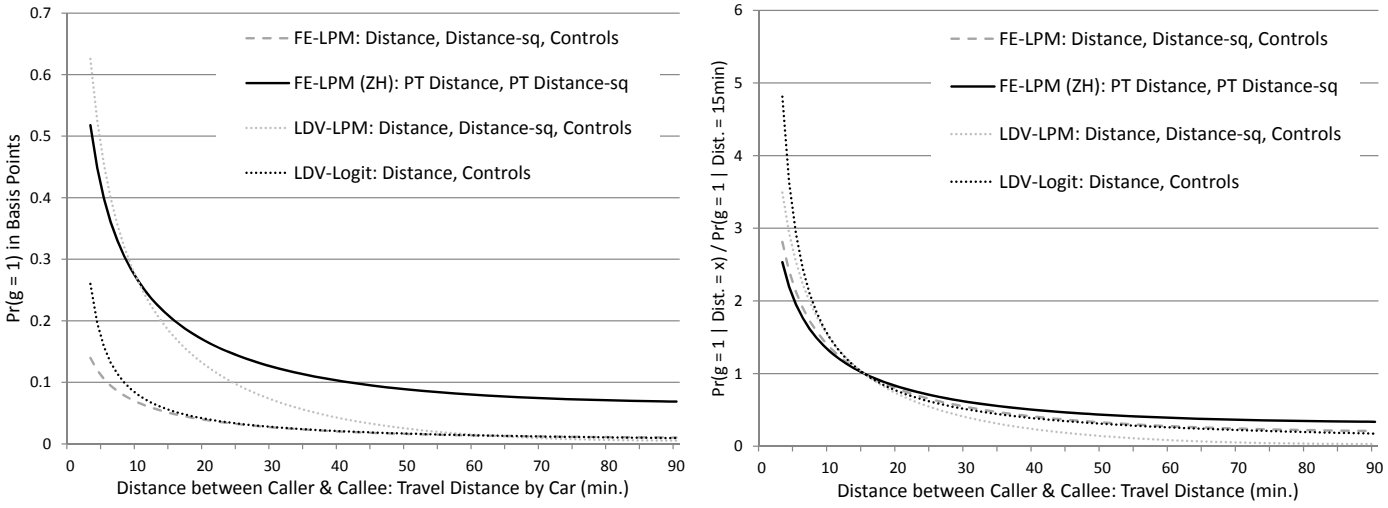
	Switzerland			Canton Zurich		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Travel Time PT$_{ij,t}$)	-0.405***	-0.008**	-0.055	-0.382***	-0.022*	-0.328**
	(0.001)	(0.003)	(0.074)	(0.001)	(0.009)	(0.121)
Ln(Travel Time PT$_{ij,t}$)^2			0.004			0.035**
			(0.007)			(0.013)
Constant	2.157***	0.099***	0.222	1.753***	0.178***	0.836***
		(0.016)	(0.198)	(0.006)	(0.039)	(0.273)
R^2	0.001	0.523	0.523	0.001	0.554	0.554
Pair FE	No	Yes	Yes	No	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Postcode Pairs	5,384,294	5,384,294	5,384,294	203,227	203,227	203,227
Observations	83,183,964	83,183,964	83,183,964	18,149,188	18,149,188	18,149,188

Notes: We use data from three-months windows prior and after the change in the public transport timetable on December 13th 2015, i.e. June 2015–August 2015 and March 2016–May 2016. The *sample* covers only non-movers (both caller and callee) who used their phone every month at least once. In column (1)–(3), we drop observations in the canton of Ticino as these were affected by an infrastructure change not recorded in our travel time data. All *coefficients* of the linear probability models are multiplied by 10000, and therefore can be interpreted as basis points. Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

periment that allows us to relax the assumption about time constant matching quality. After the completion of a central tunnel and several new railway connections, the Swiss Federal Railways company (SBB) issued a new timetable on 13 December 2015. It was the most substantial change of the SBB's timetable since 2004, affecting both the frequency of connections and journey times across Switzerland (for additional information see Section A.4 in the appendix). Notably, the planning of Switzerland's public transport schedules is highly centralized; the SBB holds a market share of around 80% in rail traffic so that local providers coordinate their services with the SBB. This centralisation brings about nation-wide changes in public transport connections triggered by newly established connections of the federal railway. Moreover, it facilitates reliable timetable queries from webservices such as *search.ch* (our data source).

We re-estimate equation (10b.), but instead of using movers, we use public transport travel times as a measure for distance, and hence identify η_1 based on changes in the public transport timetable.[19] To reduce noise, we employ data from three-months windows prior and after the change in the public transport timetable, namely June 2015–August 2015 and March 2016–May 2016. Furthermore, we restrict the sample to individuals who keep the same billing address, so that the estimates of η_1 are not affected by the potentially endogenous moving decision. Table 3 shows the results for Switzerland (columns 1–3) and a subsample of individuals living in the canton of Zurich (columns 4–6), where the largest changes were implemented. As in the previous models, distance, now measured by public transport travel times, is negatively correlated with the probability that two

[19]Note that public transportation is frequently used in Switzerland; e.g. public transportation covers about 60 percent of the commutes in the Zurich area.



(a) Level

(b) Ratio, Base Distance = 15 min.

Notes: FE-LPM (fixed-effects, linear probability model): table 2, column (4); FE-LPM (ZH) (fixed-effects, linear probability model for canton Zurich): table 3, column (6); LDV-LPM (lagged dependent variable, linear probability model): table 2, column (6); LDV-Logit (lagged dependent variable, logit model): table B.1, column (4). Models including controls are evaluated at the following values: Same Workplace=0, Common Contacts=0, Degree=mean, Same Gender=1, Same Age=1, Age Diff=0, $g_{ij,t-1}$=0, FE=0.

Figure 5: Predicted Probability to Form a Tie

agents form and maintain a link. Although the estimates' precision drops somewhat, the negative effect of distance remains when introducing pair fixed effects, while the square term again enters positive pointing to a convexly decreasing relation.

We now inspect the magnitude and the functional relation between distance and tie formation in more detail. Figure 5a displays the predicted probability for $g_{ij} = 1$ based on various specifications. Figure 5b plots the relative probability for $g_{ij} = 1$ compared to the base probability at a distance of 15 minutes travel time. Although the models differ regarding the level prediction, they consistently reveal a convexly decreasing relation between link formation and distance. Overall, the graphs illustrate that the effect of distance on social interactions is highly localized; the probability of forming a link is about twice as large for neighbors than for people living 10 minutes apart. This probability continues to fall quickly up to a distance of 30 minutes, beyond which the negative effect of travel time flattens out.

In summary, the evidence shows that distance is highly detrimental to social interactions. This suggests that face-to-face interactions and phone communication are complementary. If distance between two individuals did not impose costs on their social exchange, it would be difficult to argue that regional differences in population density should impact the topography of social networks. In what follows, we examine whether distance costs indeed lead to significant differences in the topography of social networks across urban

and rural areas. First, we examine the consequences regarding network size, and then we turn our attention to network efficiency.

5.2 Cities and Network Size

In order to directly test whether cities are favourable to network size, we estimate a series of pooled OLS models, which are reported in Table 4. We use two sets of key explanatory variables, including the trichotomous classification for urbanisation by EUROSTAT (i.e. urban core, hinterland, periphery) as well as a continuous measure for population density. The latter is defined as the log of the population living within a 15-minute radius of an individual's postcode area. Network size is measured on a monthly basis as degree centrality, i.e. the number of unique contacts an individual calls during one month.

Columns (1) and (2) contain the results for the discretized measure of urbanisation, the former excluding and the latter accounting for individual controls in the regression. Agents who live in the hinterland or periphery have on average a smaller network than city residents. The correlations are statistically highly significant, with an average difference of -1.1 to -1.7 percent when comparing the periphery to the urban core, and -2.4 to -2.5 percent when comparing the hinterland to the urban core.

The continuous population density measure in column (3) is negatively correlated with network size. This unexpected result is due to non-linearities, as the results in columns (1) and (2) already indicate; although the hinterland has a higher population density than peripheral municipalities, the hinterland coefficient is significantly smaller than the periphery coefficient. When a squared-term is included (column (4)), the results indeed reveal a convex relation between population density and network size, with the marginal effect of population density turning positive around its mean value.

Overall, these findings lend support to the hypothesis that dense urban areas facilitate social interactions. So far it is unclear, however, whether the effect has a causal interpretation or is driven by the sorting of high sociability types to urban centres.

In a next step, the regressions include individual fixed effects to back out any person specific characteristics and thereby eliminate the sorting channel. Consequently, inference is now based on customers who changed their billing address during the 12 months period covered. Columns (5) to (7) of Table 4 display results for the baseline fixed effects regression, while columns (8) to (10) show a robustness check based on people who changed their residence by at least 30 minutes driving time. The results stand in stark contrast to the pooled OLS regressions and clearly reject the hypothesis that cities have a causal impact on network size. All coefficients related to regional differences in population density are practically zero and statistically insignificant.

Table 4: Regional Differences in Network Size

	Pooled OLS				FE: Full Sample			FE: Moving Distance > 30min.		
Dependent Variable: $D_{ir,t}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Hinterland (vs. Cities)	-0.024***	-0.025***			0.000			-0.006		
	(0.001)	(0.001)			(0.003)			(0.006)		
Periphery (vs. Cities)	-0.011***	-0.017***			0.000			-0.001		
	(0.001)	(0.001)			(0.004)			(0.007)		
Ln(Pop. Density)			-0.008***	-0.222***		-0.002	-0.001		-0.002	-0.006
			(0.000)	(0.002)		(0.001)	(0.012)		(0.002)	(0.017)
Ln(Pop. Density)^2				0.012***			-0.000			0.000
				(0.000)			(0.001)			(0.001)
R^2	0.011	0.067	0.067	0.068	0.011	0.011	0.011	0.011	0.011	0.011
Further Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Language Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Groups	-	-	-	-	60,514	60,514	60,514	16,874	16,874	16,874
Observations	10,117,645	9,353,794	9,353,679	9,353,679	669,825	669,825	669,825	185,676	185,663	185,663

Notes: We use monthly data for June 2015–May 2016. The *sample* in columns (1)-(4) covers customers who used their phone every month at least once. The *sample* in columns (5)-(10) covers movers who used their phone every month at least once. *Further controls* include commuting distance, language (pooled OLS), dummy for belonging to language minority, gender (pooled OLS), and age (pooled OLS). Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

Figure B.2 in the appendix plots the degree of movers over time. It shows that agents expand their social network in the three months prior to moving, and then revert to their initial level within two months. To test the robustness of our results with respect to this dynamic, we re-estimate the fixed effects models for movers who changed their residence by at least 30 minutes driving time and successively exclude periods around the moving month (see Table B.3 in the appendix). These additional results support the conclusions drawn from benchmark analysis in Table 4.

One further concern may be that urban residents use messenger apps more frequently than people in rural areas, which could lead to a downward bias in population density. Such concerns seem unsubstantiated for three reasons. *First*, at least two thirds of the customers have zero marginal costs for domestic calls as they subscribe to flatrate contracts which price discriminate via data usage. Hence, most customers have little incentive to substitute from phone calls to data intensive messenger apps or voice over IP communication. *Second*, messenger apps and mobile phone calls are most likely complements not substitutes. We decompose messenger usage along gender and language region, based on a survey conducted by *comparis.ch* in 2014. It shows that messenger apps are more often used among men than women and are more widespread in French-speaking than German-speaking regions. The same ranking unfolds for network size in terms of mobile phone calls. This indicates that the two media are complements not substitutes.[20] *Third*, we conduct a series of robustness checks, in which we control for an individual's technology preferences. Table B.8 adds two proxies for technology usage to our benchmark model: the ratio of outgoing SMS versus outgoing calls, as traditional text messages are the most likely technology to be substituted by messenger apps. We further include the ratio of outgoing landline calls versus the total number of calls, because apps may be used to call another mobile phone but not landlines. This robustness check does not alter the results in Table 4, as density remains uncorrelated with network size. Note that we report analogous robustness checks for all other network measures in Appendix B.7.

It remains, then, that the correlation between population density / urbanisation and network size is driven by the sorting of above-average sociable people to the urban core and cannot be attributed to a positive externalities of people living close together. A variance decomposition, which computes the contributions of individual fixed effects, local fixed effects, and time specific factors to the total variance of $D_{i,t}$, also supports the conclusion that regional differences play a small role in explaining differences in network size. Individual components contribute 73.0 percent to the overall variance of degree centrality, while local factors only explain 2.3 percent. The remaining variation can be attributed to time specific factors (0.3%) and to the residual (24.4%), i.e. individual and

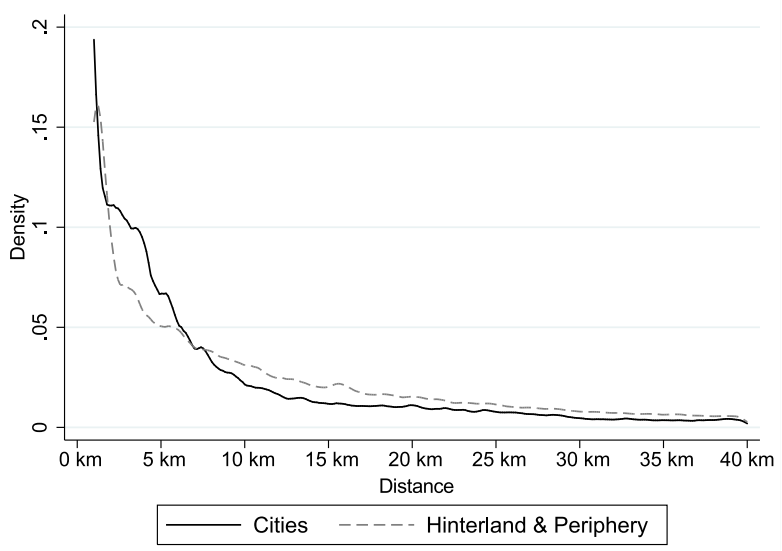
[20] Charlot and Duranton (2006) also show that telephone usage is complementary to all other modes of communication.

time variant components.

This raises the question of why people with an above-average sociable predisposition move to cities. Sociability may thereby refer to the mental capability of maintaining ties, as suggested by the social brain hypothesis, and/or to personality traits, as advocated by Asendorpf and Wilpers (1998). One evident explanation could be that cities provide a favourable environment for social interactions, which does not manifest itself in terms of network size but rather with respect to network efficiency. If this were the case, individuals with a preference for and capability of maintaining large networks would disproportionally benefit from moving to cities, which could explain the sorting pattern uncovered in the above analysis.

5.3 Cities and the Perimeter of Social Networks

We begin the discussion of network efficiency by examining variations in network perimeters across regions. Everything else being equal, an agent is better off the less distant her social contacts live, simply because she will incur lower travel costs. Since people residing in cities have a larger pool of potential contacts within close proximity, one would expect them to recruit their social contacts within a narrower perimeter to minimize travel costs.



Notes: The density plot starts at 1 km; links spanning shorter distances (mostly links within the same postcode) were assigned a value of 1 km.

Figure 6: City versus Hinterland / Periphery – Density Plot for Social Ties by Radius.

Figure 6 plots the density of social ties by radius and location type (i.e. cities versus hinterland & periphery). In comparison to individuals living in the hinterland or periphery,

urban residents evidently have a larger mass of social contacts within a 7 km radius, and fewer contacts beyond. This supports the hypothesis that living in a city can lower the costs incurred from social interactions with distant contacts.

In order to examine this claim further we use the network formation model to test whether urban residents value distance differently than people living in less densely populated areas. To this end we interact the log of distance with either population density or the city dummy. The top panel of Table 5 reports the output of the augmented network formation model, with columns (1) and (2) displaying the pooled OLS results and columns (3) and (4) showing the pair fixed effect estimates. All specifications suggest that urban residents incorporate distance costs more strongly in their valuation than people living in peripheral areas. The interaction terms yield statistically significant negative effects, but are quantitatively relatively small.

Next, we resort to our network topography specification using the within-degree, DW_i^r, as dependent variable. The bottom panel of Table 5 reports the outputs of this approach, with columns (1) and (2) displaying the pooled OLS results and columns (3) and (4) showing the fixed effect estimates that account for the sorting of highly sociable individuals to urban areas. As hypothesized, within-degree is largest in urban areas and positively correlated with population density. This holds true for both the pooled OLS estimates, as well as the fixed effects results. According to our causal estimates from the fixed effects specification, urban residents have on average a 10 percent higher within-degree than individuals residing in the hinterland, and a 23 percent higher within-degree than people living in peripheral areas. The results also show that doubling population density leads, on average, to a 6.8 percent higher within-degree.[21] While population density is hardly relevant for overall network size, it has considerable explanatory power regarding the number of close-range contacts. The variance decomposition also reveals that regional factors explain more than twice as much of the within-degree variance than the variance in network size.

Considering that distant social contacts are costly, these results suggest that urban residents bear fewer costs from social interactions than people living in sparsely populated areas. This could – at least partly – explain why sociable people sort into cities, as they disproportionally benefit from this channel and therefore have a higher willingness to pay for housing in cities than less sociable types. This result may also be interpreted as better matching in cities, because geographical distance is essentially one dimension of matching quality. We further explore matching quality in the following section.

[21] As for the degree, we re-estimate the fixed effects models for movers with a minimum moving distance of 30 minutes and successively exclude periods around the moving month. Table B.4 in the appendix shows that this does not alter the main conclusion. We further include proxies for technology preferences in columns (3) and (4) of Table B.8, which also leaves the results unaffected.

Table 5: Regional Differences in the Perimeter of Social Networks

a. Network Formation	Pooled OLS		Panel FE	
Dependent Variable: $g_{ij,t}$	(1)	(2)	(3)	(4)
Ln(Travel Time$_{ij,t}$)	-0.068***	-0.069***	-0.016***	-0.016***
	(0.003)	(0.003)	(0.001)	(0.001)
Ln(Travel Time$_{ij,t}$) × City$_{i,t}$	-0.001***		-0.001**	
	(0.000)		(0.000)	
Ln(Travel Time$_{ij,t}$) × Ln(Pop. Density$_{i,t}$)		-0.001***		-0.001***
		(0.000)		(0.000)
R^2	0.054	0.054	0.088	0.088
Further Controls	Yes	Yes	Yes	Yes
Pair FE	No	No	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Groups	–	–	2,582,702	2,582,702
Observations	27,238,673	27,238,673	27,238,673	27,238,673
b. Network Topography	Pooled OLS		Panel FE	
Dependent Variable: $DW^r_{i,t}$	(1)	(2)	(3)	(4)
Hinterland (vs. Cities)	-0.111***		-0.123***	
	(0.001)		(0.010)	
Periphery (vs. Cities)	-0.208***		-0.231***	
	(0.001)		(0.012)	
Ln(Population Density)		0.086***		0.143***
		(0.000)		(0.004)
R^2	0.049	0.056	0.018	0.033
Further Controls	Yes	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes
Language Region FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Groups	–	–	60,514	60,514
Observations	9,353,794	9,353679	669,825	669,812

Notes: We use monthly data for June 2015–May 2016. The *sample* covers movers who used their phone every month at least once. $DW^r_{i,t}$ is computed using a 15 min. travel distance. *a. Controls in network formation models:* Dummies for same workplace, same language, common contacts, degree of both agents (pooled OLS), same gender (pooled OLS), same age (pooled OLS), and the absolute age difference between agents i and j (pooled OLS). *b. Controls in network topography models:* Commuting distance, language minority dummy, gender (pooled OLS) and age (pooled OLS). Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

5.4 Cities and Matching

Since matching quality cannot be directly observed, we propose two indirect tests for the hypothesis that matching quality improves with population density. In one test we resort to the network formation model, while the second test is based on the network topography approach.

We begin with the *network formation model*, or more specifically with the fixed effect specification given in equation (10b.): The pair fixed effect absorbs any dyad-specific constant factors that either raise or lower the surplus of interaction for the involved agents. Hence, it primarily captures matching quality, $m_{ij}(\cdot)$, which governs the value obtained from forming a link with another person. If agents living in cities indeed benefit from better

matching quality, we would expect that fixed effects associated with their actually formed links are higher than the equivalent fixed effects calculated for agents living in rural areas. To test this claim, we first estimate equation (10b.), and then regress the predicted pair fixed effects $\widehat{m}_{ij}(\cdot)$ for the subsample of active links (i.e. $g_{ij} = 1$) on population density at agent i's place of residence. Because we focus on movers to back out any distance-related effects, the estimates yield the impact of population density weighted by duration of stay.[22] The results are reported in panel (a.) of Table 6. We obtain strong positive and significant effects for population density in column (2), and negative effects for residents of peripheral municipalities in column (1). Comparing the coefficients of interest with the constant suggests that matching quality in cities is about 6 percent higher than in peripheral areas, and 2 percent higher than in the hinterland.[23] Restricting the sample to customers with a minimum driving distance of 30 minutes between their old and new addresses does not affect the results. This backs the claim that densely populated areas lead to favourable matching outcomes.

We reassess the hypothesis by returning to the *network topography approach*. If people change their residence, we would expect them to keep up with some of their previous contacts and replace others with individuals living in their new neighborhood. Since distance makes social interactions costly, only highly valuable contacts at the old place of residence are worthwhile to maintain. Furthermore, if one encounters very good matches at the new place of residence, the replacement of pre-existing ties with new contacts should advance more quickly. We therefore examine whether this social adjustment process systematically varies with population density at the pre- and post-move residence. Consider an individual i moving in t from place *pre* to place *post*. We are interested in the ratio $DW^{post}_{i,t+1}/DW^{pre}_{i,t+1}$ which reflects the number of $t+1$ contacts at the *post-move* place of residence over the number of $t+1$ contacts at the *pre-move* place of residence. Put differently we estimate the speed of replacement of contacts at the pre-move place by new contacts at the post-move place:

$$\frac{DW^{post}_{i,t+1}}{DW^{pre}_{i,t+1}} = \alpha + \beta_{post} \cdot L^{post}_{i,t+1} + \beta_{pre} \cdot L^{pre}_{i,t+1} + X_i'\gamma + \varrho DW^{post}_{i,t-1} + \epsilon_{i,t+1}, \qquad (13)$$

The main explanatory variables are population density and the trichotomous classification for urbanisation at mover i's new address ($L^{post}_{i,t+1}$) and old address ($L^{pre}_{i,t+1}$), complemented with a measure for the number of pre-move contacts at the new address

[22] As a robustness check, we also restrict the sample to movers who change their residence but stay within the same class of municipalities, i.e. moving from city to city or from hinterland to hinterland. As Table B.6 in the appendix reveals, this does not alter the results.

[23] The mean of our matching measure drops by about $56/2493 \simeq 0.02$ and $145/2493 \simeq 0.06$ when comparing the hinterland and the periphery to cities.

Table 6: Regional Differences in the Matching Quality

a. Network Formation	Full Sample		Moving Distance > 30min.	
Dependent Variable: $\widehat{m}(\xi_i, \xi_j, \delta)$	(1)	(2)	(3)	(4)
Hinterland$_{i,t}$	-55.595***		-59.265***	
	(5.928)		(11.127)	
Periphery$_{i,t}$	-145.315***		-117.816***	
	(6.451)		(11.832)	
Ln(Pop. Density$_{i,t}$)		34.954***		17.834***
		(1.856)		(2.815)
Constant	2492.994***	2085.201***	2466.674***	2232.319***
	(492.036)	(115.281)	(250.231)	(84.219)
R^2	0.001	0.001	0.001	0.001
Observations	11,616,147	11,692,984	3,089,595	3,116,907
b. Network Topography	Full Sample		Moving Distance > 30min.	
Dependent Variable: $DW^{post}_{i,t+1}/DW^{pre}_{i,t+1}$	(1)	(2)	(3)	(4)
City^{post}	0.327***		0.090	
	(0.046)		(0.056)	
City^{pre}	-0.449***		-0.275***	
	(0.033)		(0.044)	
Ln(Pop. Density^{post})		0.227***		0.076***
		(0.011)		(0.013)
Ln(Pop. Density^{pre})		-0.326***		-0.138***
		(0.014)		(0.020)
Constant	1.009***	1.801***	0.685***	1.256***
	(0.097)	(0.176)	(0.037)	(0.194)
R^2	0.047	0.078	0.259	0.263
Further Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Language Region FE	Yes	Yes	Yes	Yes
Observations	28,871	28,871	7,887	7,801

Notes: *Dependent Variable in Panel a.:* Predicted dyad specific fixed effect from network formation model outlined in equation (10b). *Dependent Variable in Panel b.:* The number of *post-move* contacts at the *post-move* place of residence over the number of *post-move* contacts at the *pre-move* place of residence. *Controls in Panel b.:* Number of contacts at new address prior to moving, commuting distance, dummy for belonging to language minority, gender and age . Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

($DW^{post}_{i,t-1}$), and individual level characteristics X_i.[24] The results reported in the bottom panel of Table 6 are based on address changes between October 2015 and January 2016 (period t), a pre-move window covering June 2015 to August 2015 (period $t-1$), and a post-move window covering March 2016 to May 2016 (period $t+1$). As hypothesized, the fastest social adjustment process is observed for people who move from the periphery to the city, while movers who lived in urban areas before changing their address keep comparatively large shares of their pre-move contacts. Quantitatively the difference is substantial: While the ratio of new versus old contacts half a year after changing address is on average

[24]Instead of controlling for the pre-move contacts at the new address, we estimated the model for a subsample of customers that move to a location where they have no prior contacts, i.e. $DW^{post}_{i,t-1} = 0$. This does not alter the conclusion, as Table B.6 (Panel b.) in the appendix shows. We further include proxies for technology preferences in columns (7) and (8) of Table B.8, which also leaves the results unaffected.

1.3 for people that move into a city, it is only 0.6 for people that move out of the city.[25] Since maintaining spatially distant contacts is costly, this suggests that contacts formed in cities generate on average a higher surplus and are therefore more likely to be maintained. Hence, this test supports the hypothesis that densely populated areas improve matching quality. In Appendix B.6 we document that the superior matching in cities can be further substantiated by a number of sensitivity checks which focus on subgroups of movers and exploit *bilateral* network adjustment.

5.5 Cities and Clustering

The final network property that we examine is clustering. Agents face a trade-off in terms of efficient information exchange (i.e. low clustering) and benefits related to reciprocity (i.e. high clustering). The optimal balance may vary regionally due to factors that alter this trade-off. Additionally, one would expect that more populous neighborhoods display lower average clustering, simply because randomly established links are less likely to form triadic structures when the pool of potential contacts grows larger. To test the first claim, we resort to the network formation model. Even if there is no evidence that urban residents value triadic relations differently than people living in rural areas, the mechanical relation between population density and clustering may lead to measurable regional differences. If this is the case, the network topography approach should uncover them.

In the network formation model we interact the dummy for common contacts with either population density or the city dummy. In order to back out spurious clustering due to the grouping of similar types, we focus on the pair fixed effects specification. The top panel of Table 7 reports the results for these regressions. In both specifications, the interaction terms are negative and statistically significant at the 10 percent level. Hence, this analysis suggests that sharing a common link is valued less by urban residents than by residents of peripheral areas. Magnitude wise the impact is fairly substantial, as it amounts to approximately 20 percent of the effect attributed to the common contact dummy. While sharing a common contact increases the probability of forming and maintaining a link by 17.3 basis points, the effect is only 13.7 basis points among city residents.

Given the results of the network formation analysis, we expect lower clustering in cities than in peripheral areas. The bottom panel of Table 7 displays the results of the network topography analysis with clustering as the dependent variable. Both the pooled OLS regressions in columns (1) and (2), as well as the fixed effects specifications in columns (3) and (4) suggest that cities attenuate network clustering. The effect ranges between -0.010 and -0.014 in the pooled OLS regressions, which is roughly 11 to 15 percent of the sample

[25] According to columns (1) the average ratio is about 1.009 which increases to 1.336 for individuals moving to the city, i.e. $City^{post} = 1$ and reduces to 0.56 for those moving out of the city, i.e. $City^{pre} = 1$.

Table 7: Regional Differences in the Transitivity of Social Networks

a. Network Formation			Panel FE	
Dependent Variable: $g_{ij,t}$			(3)	(4)
>0 Common Contacts$_{ij,t-1}$			17.337***	16.477***
			(1.268)	(1.058)
>0 Common Contacts$_{ij,t-1}$ × City$_{i,t}$			-3.681^{+}	
			(2.009)	
>0 Common Contacts$_{ij,t-1}$ × Ln(Pop. Density$_{i,t}$)				-1.745^{+}
				(0.957)
R^2			0.124	0.124
Further Controls			Yes	Yes
Pair FE			Yes	Yes
Month FE			Yes	Yes
Groups			2,582,702	2,582,702
Observations			27,238,673	27,238,138
b. Network Topography	Pooled OLS		Panel FE	
Dependent Variable: $C_{ir,t}$	(1)	(2)	(3)	(4)
Hinterland (vs. Cities)	0.010***		0.002*	
	(0.001)		(0.001)	
Periphery (vs. Cities)	0.014***		0.002**	
	(0.001)		(0.001)	
Ln(Population Density)		-0.004***		-0.001**
		(0.001)		(0.000)
R^2	0.022	0.022	0.001	0.001
Further Controls	Yes	Yes	Yes	Yes
Individual FE	No	No	Yes	Yes
Language Region FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Groups	–	–	60,507	60,507
Observations	9,252,183	9,252,183	664,343	664,330

Notes: We use monthly data for June 2015–May 2016. The *sample* covers movers who used their phone every month at least once. *a. Controls in network formation models:* Dummies for same workplace, same language, common contacts, degree of both agents (pooled OLS), same gender (pooled OLS), same age (pooled OLS), and the absolute age difference between agents i and j (pooled OLS). *b. Controls in network topography models:* Commuting distance, dummy for belonging to language minority, gender (pooled OLS) and age (pooled OLS). Standard errors in parentheses. $^{+}$ p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

mean. The difference between city and hinterland / periphery drops in the fixed effects specifications, but remains significant at the 5 percent level or higher.[26]

Despite the evidence that population density has no impact on the number of social interactions, cities are shown to facilitate the diffusion of information due to below-average clustering. This can have important consequences for local labor markets, as discussed in Sato and Zenou (2015), for example. In conjunction with the findings on network size, matching quality and distance costs, this suggests that cities may encourage not a larger number but rather more valuable social interactions.

[26]Periods around the moving month are excluded in Table B.5. We also include proxies for technology preferences in columns (5) and (6) of Table B.8, which leaves the results unaffected.

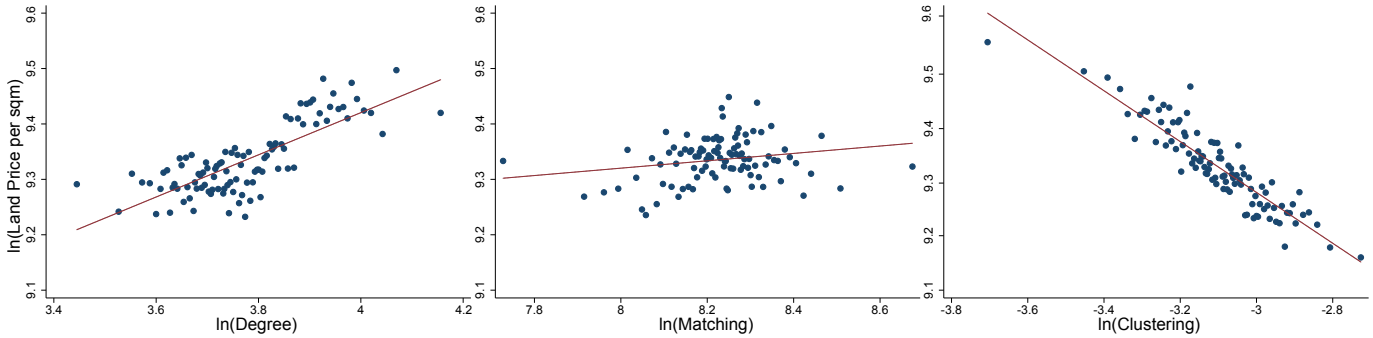
6 Assesing the Value of Social Interactions

Combining land prices with the estimates from above we can perform a simple back-of-the-envelope calculation to asses the monetary value of superior networks in cities. In a standard model of spatial equilibrium, migration of households equalizes utility across space such that higher nominal wages or superior consumption amenities are offset by higher housing costs (see Rosen, 1979; Roback, 1982). In the following, we refrain from disentangling the effect of networks on local productivity – and thus nominal wages – versus utility benefits in form of consumption amenities. Instead, we examine to what extent superior social networks reflect in higher land prices which may be a consequence of either increased productivity, or additional amenities, or both.

In order to quantify the link between network measures introduced above and land prices we resort to micro-level information about rental prices and house characteristics. In particular, we employ a dataset covering all rental offers posted on the most popular search engines in Switzerland.[27] Using rental prices instead of house prices has the advantage that our results are not affected by expected house price changes. Moreover, the rental share amounts to 59 percent in Switzerland. We estimate a standard hedonic regression where we regress the logarithm of the rent per square meter on a comprehensive set of residence specific covariates and absorb all location specific factors by postcode fixed effects. The covariates include information about the number of rooms, living area, building age, floor, the availability of a garden, balcony, attic, elevator, parking and parking garage as well an indicator on whether the residence has a view (to a lake or to the mountains). As postcodes are small regional entities with a median population of about 950 inhabitants (median size 7.3 square kilometers) we capture location specific determinants of rents reasonably well and explain almost 70 percent of the total variation in the log of rents per square meter. Hence, assuming that our detailed covariates capture relevant variation of the quality of structures, the fixed effects obtained from the hedonic regression serve as a suitable measure for the local value of land.

These data are used as a dependent variable to study the elasticity of land values with respect to social network measures. Among others, the local value of land is determined by local amenities and agglomeration economies. Following the argument that efficient social networks represent a channel through which agglomeration economies operate we expect that the quality of social networks exerts a positive effect on land prices even

[27] The information is provided by *Meta-Sys.ch*. The data provide a good coverage of rural and urban areas which we test by comparing the data with the universe of houses and apartments listed in the official building registry. In total we observe information for 546,456 residences across 2,790 postcodes offered for rent in the years 2015–2016. Due to potential measurement error that might arise from the difference between offer and transaction prices, we focus on rental data because rents are almost never negotiated in Switzerland.



Notes: Binned scatterplots between logarithmized network measures (degree, matching, clustering) and logarithmized land prices at the postcode-level conditioned on population density.

Figure 7: Capitalization of Network Measures

when controlling for population density. Figure 7 plots the partial correlation between the logarithmized network measures and logarithmized land prices. Consistent with our hypothesis, we find that higher average network degree and matching quality are associated with higher land prices while less efficient networks characterized by high clustering feature lower land prices.

If efficient social networks represent a channel through which agglomeration economies operate we not only expect that the quality of social networks exerts a positive effect on land prices but also that the positive effect of population density on local land prices is reduced when measures of social networks are included. Table 8 reports the corresponding results. The first column reports a positive correlation between land prices and population density as typically observed in the literature and attributed to agglomeration economies (e.g. Albouy and Ehrlich, 2012). We further included the trichotomous classification for urbanisation with the urban core as the reference category. Overall we find that the periphery and the hinterland display 17 and 7 percent lower land prices than the city. Moreover, all three measures of social networks enter significantly and with the expected signs. Higher average network degree and matching quality are associated with higher land prices while less efficient networks characterized by high clustering feature lower land prices (see columns (2)–(4)).[28] As is evident from columns (1) and (6) in Table 8 the effect of population density is reduced by about 30 percent when controlling for network measures. The same holds true for the indicators of hinterland and periphery which drop by 27–28 percent when accounting for social networks. Hence, these results indicate that a significant part of agglomeration economies observed in land prices can be attributed to superior social networks.

The magnitudes suggest elasticities of land prices with respect to network measures

[28]Note that common friends are valued for individual links while a high aggregate share of triadic closure/clustering lowers the value of the local network. We interpret this discrepancy as a negative externality.

Table 8: Capitalization of Network Measures

	Dependent Variable: Ln(Land Price per m^2)					
	(1)	(2)	(3)	(4)	(5)	(6)
Pop. Density	0.040***	0.073***	0.040***	0.065***	0.044***	0.028***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Degree		0.257***			0.205***	0.171***
		(0.050)			(0.050)	(0.049)
Clustering			-0.442***		-0.424***	-0.358***
			(0.044)		(0.044)	(0.043)
Matching Quality				0.115***	0.113**	0.072*
				(0.034)	(0.034)	(0.032)
Hinterland	-0.169***					-0.124***
	(0.014)					(0.013)
Periphery	-0.236***					-0.169***
	(0.015)					(0.015)
R^2	0.351	0.286	0.363	0.275	0.378	0.418
Language Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,155	2,194	2,194	2,189	2,189	2,150

Notes: Standard errors are clustered on the postcode level and reported in parentheses. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

between 0.07 to 0.36 (column(6) of Table 8). These findings are robust to the inclusion of various proxies for local geography and (dis-)amenities (e.g. ruggedness, altitude, taxes) as well as proxies for the local composition of structures such as the local share of second-homes and the local share of single-homes. However, the estimates have to be interpreted carefully as the quality of local networks may still be correlated with unobservable factors that determine land prices. A conservative back-of-the envelope calculation suggests that superior network quality leads to a difference of land prices in the periphery relative to the city of about 6 percent.[29] Recalling that land prices are generally 17 percent lower in the periphery than in the city this is well in line with the 30 percent reduction of the elasticity of land prices with respect to population density when accounting for network quality in Tables 8.

7 Conclusion

The results of this study show that cities provide a superior environment for social interactions, which is fundamentally important to the mechanics of agglomeration forces. Contrary to many theoretical models, the advantages of densely populated areas do not translate into larger social networks but rather into improvements in terms of matching

[29] According to Tables 4, 7, and 6 network size is 1.7 percent lower, clustering is 15.2 percent higher, and matching quality is 4.8 percent lower in the periphery than in the cities. This yields a price difference of $1.7\% \times 0.171 + 4.8\% \times 0.072 + 15.2\% \times 0.358 = 6.1\%$. Note that this price difference includes the sorting of high sociability types into the cities which reinforces the network externalities. The corresponding price difference is considerably lower and amounts to between 1.1 and 2.7 percent once we adjust for sorting.

quality, smaller distance costs, and a favourable structure for information diffusion (i.e. lower clustering).

Evidently, modern communication technologies do not render cities obsolete. Our analysis has demonstrated that they remain important as catalyst of valuable social exchange and, consequently, as potential engines of growth. As a significant part of agglomeration economies observed in land prices can be attributed to superior social networks we conclude that these effects are also quantitatively important: About 30 percent of the difference in land prices between the periphery and cities can be explained by measures of quality of networks. From a policy perspective, this result provides micro-level evidence for the positive externalities of densely populated areas, which should be taken into account, for example, in the design of zoning policies, or the pricing of mobility.

There are many potential extensions of the work described in this paper. First, we focused exclusively on private social interactions, thus it would be fruitful to examine whether the same conclusions apply to networks from business communication. Second, it would be interesting to distinguish the role of social networks for productivity versus consumption amenities. Third, we barely scratched the surface of the information available in the mobility data recorded from transmitting antennas. Such data would allow, for instance, to evaluate the costs of commuting in terms of social capital.

References

Abel, J. and R. Deitz (2015). Agglomeration and job matching among college graduates. *Regional Science and Urban Economics 51*, 14–24.

Albouy, D. and G. Ehrlich (2012). Housing productivity and the social cost of land-use restrictions. NBER Working Paper No. 18110.

Alesina, A. and E. La Ferrara (2000). Participation in heterogeneous communities. *Quarterly Journal of Economics 115*(3), 847–904.

Ali, N. and D. Miller (2009). Enforcing cooperation in networked societies. Society for Economic Dynamics, *mimeo*.

Ambrus, A., M. Mobius, and A. Szeidl (2014). Consumption risk-sharing in social networks. *American Economic Review 104*(1), 149–182.

Angrist, J. and J.-S. Pischke (2009). *Mostly Harmless Econometrics. An Empiricist's Companion.* Princeton: Princeton University Press.

Asendorpf, J. and S. Wilpers (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology 74*(6), 1531–1544.

Berliant, M., R. Reed, and P. Wang (2006). Knowledge exchange, matching, and agglomeration. *Journal of Urban Economics 60*, 69–95.

Blondel, V., A. Decuyper, and G. Krings (2015). A survey of results on mobile phone datasets and analysis. *EPJ Data Science 4*, 1–55.

Blum, B. S. and A. Goldfarb (2006). Does the internet defy the law of gravity? *Journal of International Economics 70*, 384–405.

Blumenstock, J., G. Cadamuro, and R. On (2015). Predicting poverty and wealth from mobile phone metadata. *Science 350(6264)*, 1073–1076.

Brueckner, J. and A. Largey (2008). Social interaction and urban sprawl. *Journal of Urban Economics 64*, 18–34.

Burley, J. (2015). The built environment and social interactions: Evidence from panel data. University of Toronto, *mimeo*.

Burt, R. (1987). A note on strangers, friends and happiness. *Social Networks 9*(4), 311–331.

Cairncross, F. (2001). *The Death of Distance: How the Communication Revolution Is Changing Our Lives.* Cambridge: Harvard Business School Press.

Charlot, S. and G. Duranton (2004). Communication externalities in cities. *Journal of Urban Economics 56*, 581–631.

Charlot, S. and G. Duranton (2006). Cities and workplace communication: Some quantitative french evidence. *Urban Studies 43*(8), 1365–1394.

Ciccone, A. and R. Hall (1996). Productivity and the density of economic activity. *American Economic Review 86*(1), 54–70.

Combes, P.-P., G. Duranton, and L. Gobillon (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics 63*, 723–742.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010). Estimating agglomeration effects with history, geology, and worker fixed effects. In E. Glaeser (Ed.), *Agglomeration Economics*, pp. 15–65. Chicago: Chicago University Press.

Combes, P.-P. and L. Gobillon (2015). The empirics of agglomeration economies. In G. Duranton, J. V. Henderson, and W. Strange (Eds.), *Handbook of Regional and Urban Economics*, pp. 247–348. Amsterdam: Elsevier.

ComCom, E. K. (2015). Tätigkeitsbericht der comcom 2015. Published online `http://www.comcom.admin.ch/dokumentation/00564/index.html?lang=de` (01.06.2016).

Cosslett, S. (1981). Maximum likelihood estimator for choice-based samples. *Econometrica 49*, 1289–1316.

Currarini, S., M. Jackson, and P. Pin (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica 77*(4), 1003–1045.

D'Costa, S. and H. Overman (2014). The urban wage growth premium: Sorting or learning. *Regional Science and Urban Economics 48*, 168–179.

De la Roca, J. and D. Puga (2017). Learning by working in big cities. *Review of Economic Studies 84*, 106–142.

Dunbar, R. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution 20*, 469–493.

Dunbar, R. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences 16*, 681–735.

Dunbar, R. (2015). Do online social media cut through the constraints that limit the size of offline social networks? *Royal Society Open Science 3*, 1–9.

Duranton, G. and D. Puga (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson and J. Thisse (Eds.), *Handbook of Regional and Urban Economics*, pp. 2063–2115. Amsterdam: Elsevier.

Fischer, C. (1982). *To Dwell among Friends. Personal Networks in Town and City.* Chicago: Chicago University Press.

Forman, C., A. Goldfarb, and S. Greenstein (2005). How did location affect adoption of the commercial Internet? Global village vs. urban leadership. *Journal of Urban Economics 58*, 389–420.

Forman, C., A. Goldfarb, and S. Greenstein (2012). The Internet and Local Wages: A Puzzle. *American Economic Review 102(1)*, 556–575.

Gaspar, J. and E. Glaeser (1998). Information technology and the future of cities. *Journal of Urban Economics 43*, 136–156.

Glaeser, E. (1999). Learning in cities. *Journal of Urban Economics 46*, 254–277.

Glaeser, E., H. Kallal, J. Scheinkman, and A. Shleifer (1992). Growth in cities. *Journal of Political Economy 100*(6), 1126–1152.

Graham, B. (2014). Methods of identification in social networks. NBER Working Paper No. 20414.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology 78*(6), 1360–1380.

Granovetter, M. (2005). The impact of social structure on economic outcomes. *Journal of Economic Perspectives 19*(1), 33–50.

Gui, B. and R. Sugden (2005). Why interpersonal relations matter for ecconomics. In B. Gui and R. Sugden (Eds.), *Economics and Social Interaction*, pp. 1–23. New York: Cambridge University Press.

Guryan, J. (2001). Desegregation and black dropout rates. NBER Working Paper No. 8345.

Helsley, R. and Y. Zenou (2014). Social networks and interactions in cities. *Journal of Economic Theory 150*, 426–466.

Hilber, C. (2010). New housing supply and the dilution of social capital. *Journal of Urban Economics 67*, 419–437.

Ioannides, Y., H. Overman, E. Rossi-Hansberg, and K. Schmidheiny (2008). The effect of information and communication technologies on urban structure. *Economic Policy 23*(54), 201–242.

Jackson, M. (2008). *Social and Economic Networks*. Princeton: Princeton University Press.

Jackson, M. (2014). Networks in the understanding of economic behavior. *Journal of Economic Perspectives 28*(4), 3–22.

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

Jackson, M., B. Rogers, and Y. Zenou (2017). The economic consequences of social-network structure. *Journal of Economic Literature 55*(3), 1–47.

Jacobs, J. (1969). *The Economy of Cities*. New York: Random House.

Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics 95*(2), 391–413.

Levy, M. and J. Goldenberg (2014). The gravitational law of social interaction. *Physica A 393*, 418–426.

Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics 22*, 3–42.

Manski, C. and S. Lerman (1977). The estimation of choice probabilities from choice based samples. *Econometrica 45*(8), 8.

Marshall, A. (1890). *Principles of Economics*. London: Macmillan.

McPherson, M., L. Smith-Lovin, and J. Cook (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology 27*, 415–444.

Milgram, S. (1967). The small-world problem. *Psychology Today 1*(1), 61–67.

Powell, J., P. Lewis, N. Robers, M. Garcia-Finana, and R. Dunbar (2012). Orbital prefrontal cortex volume predicts social network size: An imaging study of individual differences in humans. *Proceedings of the Royal Statistical Society B 279*, 1–6.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy 90*(6), 1257–1278.

Rosen, H. (1979). Housing decisions and the u.s. income tax: An econometric analysis. *Journal of Public Economics 11*(1), 1–23.

Sato, Y. and Y. Zenou (2015). How urbanization affect employment and social interactions. *European Economic Review 75*, 131–155.

Schläpfer, M., L. Bettencourt, S. Grauwin, M. Raschke, R. Claxton, Z. Smoreda, G. West, and C. Ratti (2014). The scaling of human interactions with city size. *Journal of the Royal Society Interface 11*(98), 1–9.

Stiller, J. and R. Dunbar (2007). Perspective-taking and memory capacity predict social network size. *Social Networks 29*, 93–104.

Travers, J. and S. Milgram (1969). An experimental study of the small world problem. *Sociometry 32*(4), 425–443.

Watts, D. (1999). Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology 105*(2), 493–527.

A Appendix: Data

A.1 Descriptive Statistics – Municipalities and Postcode Areas

Table A.1: Main Descriptives for Municipalities and Postcode Areas

	Mean	SD	Min	Max
		Municipal Level (N=2322)		
Area in km^2	17.412	31.434	0.327	438.562
Population (from 2010 Census)	2396	3397.175	50	384786
Market Share of Swisscom	0.577	0.096	0.090	0.997
Degree of Urbanization				
Core	0.035	–	0	1
Periphery	0.336	–	0	1
Hinterland	0.629	–	0	1
Main Language				
German	0.628	–	0	1
French	0.295	–	0	1
Italian	0.065	–	0	1
Rhaeto-Romanic	0.012	–	0	1
Distance: Municipality i to j				
Euclidean Distance (km)	107.611	58.955	0.581	348.644
Travel Time by Car (min.)	134.004	66.897	0.692	433.696
Travel Time by Public Trans. (min.)	253.447	93.206	1.061	712.070
		Postcode Level (N=3201)		
Area in km^2	12.927	19.215	0.014	242.904
# Customers within 15 min. Radius	14683	16818.31	50	107549
Distance: Postcode i to j				
Euclidean Distance (km)	111.931	59.501	0.336	353.852
Travel Time by Car (min.)	142.804	69.033	0.283	453.508
Travel Time by Public Trans. (min.)	269.904	100.181	1.008	713.000

Sources: Municipal and postcode areas from Swisstopo; municipal population, language shares, and degree of urbanisation from Federal Statistical Office; car travel times from *search.ch*; number of private mobile phone customers from *Swisscom*. Data from postcodes and municipalities with less than 50 customers were deleted due to privacy concerns.

Table A.1 summarizes the variables used on municipality and postcode level. In total there are $2,322$ municipalities and $3,201$ postcodes in Switzerland. Population by language stems from the most recent census carried out in 2010. Municipality and postcode shapefiles were provided by Swisstopo and used to assign residences of individuals to postcodes and municipalities. Travel time between postcodes and municipalities is computed in terms of public transport and car travels on the basis of data from *search.ch*. Degree of urbanization is provided by the Federal Statistical Office and refers to the OECD classification of functional areas (also used by EUROSTAT).

A.2 Phone Usage Statstics

Table A.2 displays monthly phone activity and call duration statistics of private customers subdivided into device and message type, i.e. mobile phone calls, text messages sent from mobile phones and landline calls. We restrict our analysis to mobile phones and then filter the data as motivated in Section 3. In particular we restrict the analysis to the first 28 days of a month, *keeping* (*i.a.*) calls between mobile phones, (*i.b.*) customers that registered only one mobile phone, (*ii.*) outgoing calls, (*iii.*) calls with a duration of more than 10 seconds, (*iv.*) mobile phones with a monthly call duration between 1 minute and 56 hours. The filtered data comprises about 40% of private mobile phones calls in the data representing 60% of the total call duration; the filtering skews the sample towards relatively long-lasting calls as very short calls are deleted from the data set in step (*iii.*).

Table A.2: Call Duration (in Mio. Minutes) between June 2015 to May 2016

	Phone Activity (in Mio.)					Call Duration (in Mio. Minutes)			
	MP-Calls	SMS	Landline	**Total**	*Filtered*	MP-Calls	Landline	**Total**	*Filtered*
Jun. 2015	166.3	90.9	64.3	**321.6**	*66.0*	351.2	296.2	**647.4**	*222.4*
Jul. 2015	157.3	91.9	57.8	**307.0**	*62.0*	324.8	271.1	**595.9**	*202.2*
Aug. 2015	153.6	89.0	59.7	**302.3**	*60.3*	337.0	283.6	**620.6**	*211.3*
Sep. 2015	153.8	85.2	61.9	**300.9**	*61.6*	343.0	294.2	**637.2**	*216.9*
Oct. 2015	133.6	76.3	59.9	**269.8**	*53.7*	307.5	284.8	**592.3**	*192.6*
Nov. 2015	138.1	77.7	62.1	**277.9**	*56.5*	333.1	298.5	**631.6**	*208.7*
Dec. 2015	154.1	79.1	61.6	**294.8**	*62.0*	347.4	298.1	**645.5**	*218.5*
Jan. 2016	155.7	78.5	62.0	**296.2**	*61.0*	376.0	312.4	**688.4**	*235.5*
Feb. 2016	167.6	77.5	60.6	**305.7**	*66.3*	393.3	299.6	**692.9**	*246.7*
Mar. 2016	163.3	74.9	58.6	**296.8**	*65.4*	378.1	286.8	**664.9**	*240.3*
Apr. 2016	164.2	70.7	59.9	**294.8**	*65.7*	378.8	286.1	**664.9**	*241.1*
Mai 2016	161.1	68.6	55.9	**285.7**	*64.9*	353.5	264.6	**618.1**	*228.3*

Notes: These figures base on phone usage statistics of private customers.

A.3 Descriptive Statistics – Individual Level

Table A.3 displays the correlation coefficients of population figures from the census data and the customers numbers from our data by age group, language group, and gender. It is evident that our data is highly representative for the Swiss population at the local level. This holds even true when we study specific subgroups of the population as the correlation coefficients are always well above 0.9 except for Italian speaking part of Switzerland (Ticino). In Ticino, which represents only about 5 percent of Swiss municipalities, we still observe a correlation coefficient of about 0.9 but other phone providers seem to be relatively strong for the age group 30 where the correlation coefficients is only 0.765.

Table A.4 lists the monthly number of movers as identified from the billing address

Table A.3: Correlation between Census Population and Number of Customers at the Municipality Level

	All	Male	Female	German	French	Italian
Age All	0.987	0.984	0.988	0.992	0.990	0.893
Age 20	0.945	0.946	0.944	0.960	0.946	0.916
Age 30	0.953	0.955	0.951	0.953	0.973	0.765
Age 40	0.968	0.963	0.971	0.983	0.993	0.875
Age 50	0.985	0.982	0.984	0.993	0.988	0.914
Age 60	0.990	0.988	0.987	0.994	0.984	0.922

Notes: These figures base on customer information of active phones during June 2015 and the most recent census conducted by the Federal Statistical Office in 2010.

Table A.4: Number of Private Mobile Phone Customers with a Change in Residence

			DEGURBA Classification of Movers			
Month	All	Distance > 30min	City to Hint./Peri.	Hint./Peri. to City	Within Hint./Peri.	No Change
July	13880	4461	1468	1858	2864	7690
August	14212	4572	1431	1930	2923	7928
September	15636	4842	1584	2044	3160	8848
October	15673	4795	1572	2052	3229	8820
November	14820	4612	1537	1977	3070	8236
December	14053	4202	1396	1836	3229	7592
January	13292	4432	1194	2207	2708	7183
February	13705	4333	1275	2033	2807	7590
March	15171	4671	1501	2060	3181	8429
April	15838	4873	1529	2111	3234	8964

Notes: Movers are identified based on address changes in the customer database. Columns 3 to 6 document the moving pattern along the DEGURBA classification.

recorded in the provider's customer database. Between 13,292 to 15,838 mobile phone users reported a change of their billing address each month; this amounts to about 6.5 percent of all customers within the 12 months period covered. For one-third of movers the moving distance is larger than 30 minutes driving time; we use this subsample to check the robustness of our benchmark estimations, which are based on all movers irrespective of distance between the old and new place of residence. Table A.4 further documents the moving pattern along the DEGURBA classification. In particular, it shows that about 10% of movers change their residence from an urban municipality to the hinterland/periphery while about 15% move from the hinterland/periphery to the city. About 20% of residence changes cover moves from the periphery to the hinterland, or vice versa. For the majority, the post- and pre-move residence do not change in terms of DEGURBA-classification.

One concern maybe that movers are systematically different from non-movers. Table A.5 compares phone usage statistics, network characteristics, and sociodemographics between movers and non-movers. While movers are considerably younger than non-movers

Table A.5: Comparing Non-movers to Movers, Main Descriptive Statistics

	Non-Movers			Movers			
	Mean	SD	N	Mean	SD	N	*Difference*
Monthly Phone Usage Statistics, June 2015 – May 2016 (pooled)							
Number of Calls	110.525	109.039	9 564 636	126.170	114.840	834 913	*-15.646*
Duration (Minutes)	250.840	293.322	9 564 636	302.285	316.835	834 913	*-51.445*
Monthly Network Characteristics, June 2015 – May 2016 (pooled)							
Degree Centrality	9.164	7.912	9 564 636	9.633	7.875	834 913	*-0.468*
Within-Degree	7.163	7.266	9 564 636	5.971	6.721	834 913	*1.192*
Clustering Coefficient	0.092	0.134	9 423 136	0.081	0.114	825 787	*0.011*
Sociodemographics - Private Mobile Phones							
Age	35.307	13.734	797 053	31.038	10.624	69 593	*4.269*
Female	0.522	–	797 053	0.527	–	69 593	*-0.005*
Language: German	0.680	–	797 053	0.703	–	69 593	*-0.023*
Language: French	0.271	–	797 053	0.251	–	69 593	*0.020*
Language: Italian	0.043	–	797 053	0.039	–	69 593	*0.004*
Language: English	0.006	–	797 053	0.007	–	69 593	*-0.001*

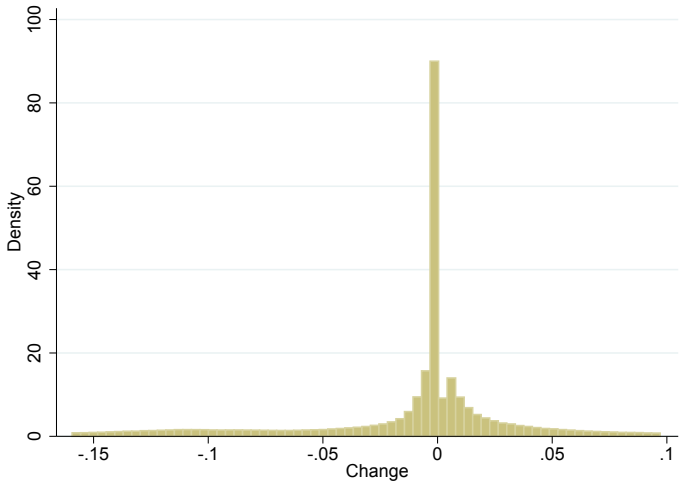
Notes: The table is based on the subsample of customers with phone activity in all 12 months, which we also use in the main analysis. Further filters as described in Section 3. Phone usage statistics include in- and outgoing calls. The *within-degree* measures network size within a radius of 15 minutes around an agent's residence.

(∼4.2 years, ∼1/3 SD), they only marginally differ along the other dimensions: Movers call slightly more (∼15 calls, ∼1/7 SD) and longer (∼50 min., ∼1/6 SD), have a marginally higher degree (∼0.5, ∼1/17 SD), lower within degree (∼1.2, ∼1/6 SD), and lower clustering (∼0.01, ∼1/11 SD). In terms of gender and language, movers are practically identical to non-movers.

A.4 Public Transport: Change of the Federal Railway Timetable

In Section 5.1 we exploit changes in the federal railway timetable to infer the causal impact of distance on link formation. The new timetable was put into effect on 13 December 2015, splitting our sample of phone data – that spans from June 2015 to May 2016 – into 6.5 and 5.5 months periods. Notably, the planning of Switzerland's public transport schedules is considerably centralised; the Swiss Federal Railways company (SBB) holds a market share of around 80% in rail traffic so that local providers typically coordinate their services with the SBB timetable. For instance, Switzerland's largest city transport network in Zürich – the Zürcher Verkehrsverbund (ZVV) – also revised its timetables on 13 December 2015 in order to synchronize their connections with SBB. This centralisation facilitates reliable timetable queries from websites such as *search.ch*, and also brings about nationwide changes in public transport connections triggered by revisions in SBB's scheduling.

The change of timetable in December 2015 was the largest of its kind since 2004. It



Notes: Illustrates the LN-differences for travel times before and after the change in the federal railway timetable on 13 December 2015; Mean: -0.012, Std. Dev.: 0.084, Share of Zeros: 0.256.

Figure A.1: The Impact of the Revised Timetable on Travel Times between Postcode Pairs

aimed at incorporating new regional and interregional connections affecting travel times both through longer/shorter journey times and through longer/shorter waiting times.[30] To calculate the changes in travel time between two places, we proceed as follows: *search.ch* kindly provided data on the quickest connections between all pairs of serviced public transport stops, i.e. about 26,000×26,000 different routes, including the frequency of available connections for a two hour window between 6am and 8am. The data covers four randomly chosen weekdays in 2015 (before the change of the timetable) and four randomly chosen weekdays in 2017 (after the change of the timetable). We build a cleaned and integrated file for 2015 and 2017, where we select the day with the shortest journey time; typically journey times do not vary across different days of the week unless construction or maintenance work causes temporary delays. As the data includes x-y-coordinates of each public transport stop, we can reliably assign them to a postcode/municipality; we then extract the quickest transport link for every postcode/municipality pair in 2015, including the stop-ids, journey time in minutes, and the number of available connections between 6am and 8am. Our final measure of public transport travel times incorporates both journey and waiting time, and is defined as

$$Public\ Transport\ Travel\ Time = \frac{120\ min.}{\#Available\ Connections} + Journey\ Time. \quad \text{(A.1)}$$

[30]Detailed summaries of all changes made in December 2015 can be found on the SBB's website, e.g. `https://stories.sbb.ch/fahrplanwechsel-dezember-2015/2015/11/10/` or `https://company.sbb.ch/de/medien/medienstelle/medienmitteilungen/detail.html/2015/11/1111-1`.

To obtain a comparable travel time matrix for 2017, we use the same selection of stops as in the 2015 matrix. Any changes between travel times in 2015 and 2017 can then be attributed to the change of timetable on 13 December 2015. Figure A.1 plots the distribution of percentage changes in the travel time between 2015 and 2017, while summary statistics for public transport travel times in 2015 are shown in Table A.1. The largest changes occured around Zurich, which is why we estimate the models for Switzerland as well as a subsample consisting of Zurich and its neighboring cantons, namely Schaffhausen, Thurgau, St. Gallen, Schwyz, Zug, and Aargau.

The change in the federal railway timetable affected travel times for three quarters of postcode pairs; on average the modifications lowered travel times by 1.2% ranging from reductions of 15% up to increases of 10%.

A.5 Data about Housing Rents

Our data on houses offered for rent is provided by *MetaSys.ch* and includes the location as well as detailed information about characteristics of the residences. Comparing the dataset to official statistics about monthly rents (Volkszaehlung und Strukturerhebung, 2014 published by the Federal Statistical Office) shows that our data is highly representative even at the local level. We merge the data with a shapefile about postcodes to obtain the postcode fixed effects. Table A.6 shows the benchmark hedonic regression which we use to predict the postcode fixed effects.

Table A.6: Housing Rents.

Dependent Variable: Ln(Rent per m^2)	Point Estimate	Std. Error
No. Rooms	0.039***	(0.001)
Ln(Living Area)	-0.388***	(0.001)
Age	-0.001***	(0.000)
Age^2	0.000***	(0.000)
Single House	-0.001	(0.003)
Garden	0.037***	(0.001)
Balcony	0.061***	(0.001)
Parking or Garage	0.033***	(0.009)
Elevator	0.049***	(0.001)
Elevator × Floor	-0.002***	(0.000)
View	0.066	(0.048)
Chimney	0.018***	(0.001)
Conservatory	0.032***	(0.003)
Low Energy	0.097***	(0.001)
Observations	293,777	
Adj. R^2	.681	
Time FE	Yes	
Floor FE	Yes	
Postcode FE	Yes	

Notes: The *sample* covers all rental offers published in the years 2015–2016. Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

B Appendix: Robustness

B.1 Robustness: Nonlinear Model of Network Formation

We also estimated *Logit models* of network formation to accommodate for the binary dependent variable and check the robustness of these results. Since the incidental parameter problem can induce severe bias in the logit fixed effects estimates (e.g Lancaster, 2000), we only show results for the pooled logit model and the lagged dependent variable model. Table B.1 summarized the results of the logit model. We compare the logit estimates to the alternative specifications in Figures 5a and 5b in the main text.

Table B.1: Logit Models of Network Formation

	Pooled Logit		Lagged Dependent Var.	
	(1)	(2)	(3)	(4)
Ln(Travel Time$_{ij,t}$)	-1.410***	-0.877***	-1.131***	-0.976***
	(0.002)	(0.049)	(0.001)	(0.005)
Same Workplace$_{ij,t}$		0.893***		1.085***
		(0.161)		(0.013)
Same Language$_{ij,t}$		1.813***		1.685***
		(0.057)		(0.005)
>0 Common Contacts$_{ij,t-1}$		7.363***		4.786***
		(0.122)		(0.070)
>1 Common Contacts$_{ij,t-1}$		2.323***		-0.018
		(0.352)		(0.071)
g$_{ij,t-1}$			12.218***	9.868***
			(0.003)	(0.029)
Const.	-7.357***	-12.951***	-8.958***	-11.170***
	(0.010)	(0.249)	(0.007)	(0.026)
Pseudo R^2	0.138	0.376	0.492	0.527
Further Controls	No	Yes	No	Yes
Pair FE	No	No	No	No
Month FE	Yes	Yes	Yes	Yes
Observations	30,996,082	27,238,673	28,411,817	28,411,817

Notes: We use monthly data for June 2015–May 2016. The *sample* covers movers who used their phone every month at least once. *Further controls* include the degree of both agents (log), dummies for same gender and same age, as well as the absolute age difference between agents i and j. Standard errors in parentheses.
+ p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

B.2 Robustness: Network Formation with Nonlinear Distance Function

In Table B.2 we report the results for specifications with distance bin dummies instead of parametric distance control functions. We assign units to distance bins in the range of $5min$ to $60min$ with the reference group being at a distance of above $60min$ travel time. The coefficients confirm that tie formation is a convex function in distance. The base probability for forming a tie at a distance above 60min travel time amounts to 0.005 basis points in the linear probability model (without lagged dependent variable and fixed effects, column (1)). This probability increases to 0.26 basis points for agents residing in

a distance of $20min$ and reaches 0.035 percent for individuals living within a $5min$ radius. We illustrates these results in the figure below.

Table B.2: Network Formation with Nonlinear Distance

	LPM	LPM-LDV	LPM-FE
	(1)	(2)	(3)
$< 5min$	3.474 ***	1.628 ***	0.241 ***
	(0.023)	(0.011)	(0.006)
$5 - 10min$	0.683 ***	0.322 ***	0.055 ***
	(0.003)	(0.002)	(0.002)
$10 - 15min$	0.420 ***	0.198 ***	0.029 ***
	(0.001)	(0.008)	(0.001)
$15 - 20min$	0.255 ***	0.120 ***	0.019 ***
	(0.001)	(0.001)	(0.001)
$20 - 30min$	0.127 ***	0.060 ***	0.010 ***
	(0.000)	(0.000)	(0.000)
$30 - 40min$	0.057 ***	0.027 ***	0.005 ***
	(0.000)	(0.000)	(0.000)
$40 - 50min$	0.027 ***	0.013 ***	0.002 ***
	(0.000)	(0.000)	(0.000)
$50 - 60min$	0.013 ***	0.006 ***	0.001 ***
	(0.000)	(0.000)	(0.000)
Const.	0.005 ***	0.002 ***	0.024 ***
	(0.000)	(0.000)	(0.000)
Adj. R^2	0.001	0.275	0.071
Pair FE	No	No	Yes
Month FE	Yes	Yes	Yes
Observations	30'996'082	28'411'817	30'996'082

Notes: *LPM* refers to the linear probability model, *LDV* refers to the lagged-dependent variable model, and *FE* refers to pair fixed effects. All coefficients of the are multiplied by 10000, and therefore can be interpreted as basis points. We use monthly data for June 2015–May 2016. The *sample* covers movers who used their phone every month at least once. Standard errors in parentheses.
+ p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

B.3 Robustness: Moving Dynamic and Degree Centrality

Figure B.2 plots the degree dynamic for movers (minimum distance of 30 min. driving time) around the moving month. The y-axis depicts the deviation in degree relative to the first month at the new address, while the x-axis reflects the timeline in terms of relocation. This event-study type of graph reveals that the average degree of movers gradually increases three months prior to relocation, and then converges back to the pre-moving period within two months. To test the robustness of our benchmark results with respect to this particular adjustment process, we step-by-step exclude periods around the moving date ($t = 0$), which we define as the first month at the new residence. In particular, the following tables contrast the benchmark results for movers with a minimum moving distance of 30 minutes (All Months, Column 0) to estimates obtained from regressions excluding the moving month ($t \neq 0$, Column 1), excluding the moving month plus the two adjacent months ($t \leq -2 \vee t \geq 2$, Column 2), and so forth.

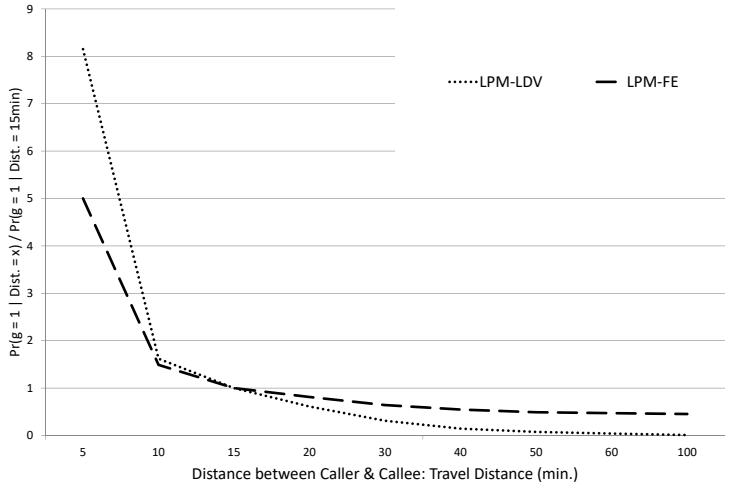


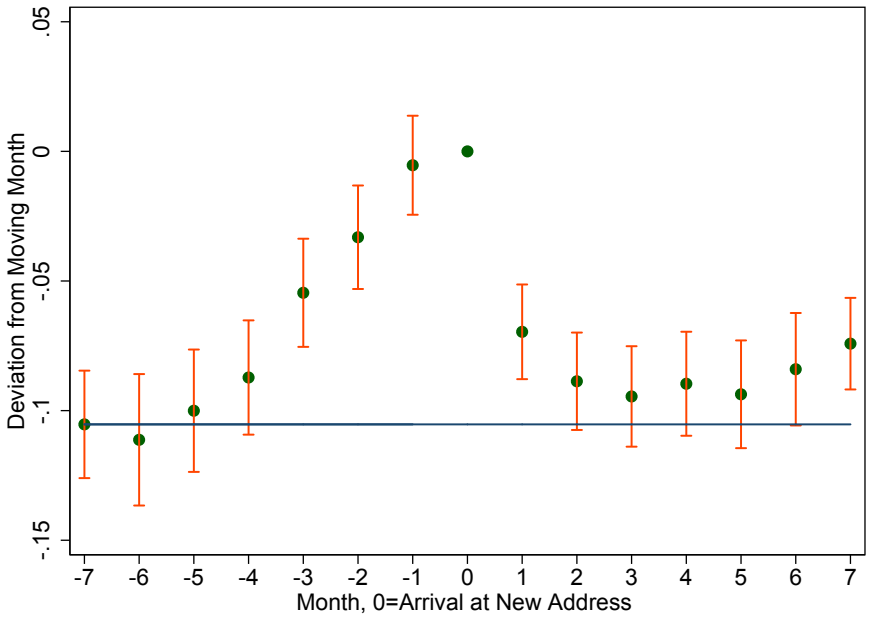
Figure B.1: Ratio, Base Distance = 15 min.

Notes: FE-LPM (fixed-effects, linear probability model); LDV-LPM (lagged dependent variable, linear probability model). Models including controls are evaluated at the following values: Same Workplace=0, Common Contacts=0, Degree=mean, Same Gender=1, Same Age=1, Age Diff=0, $g_{ij,t-1}$=0, FE=0.

Table B.3 compares the benchmark model (Column 0) relating population density and *degree* to estimates obtained from panel FE regressions that exclude periods around the moving months (Column 1 to 5). The main insight from our benchmark analysis, i.e. population density does not affect the size of a person's network, is unaltered by this robustness exercise. Although there is a tendency towards larger point-estimates once observations around the moving month are excluded, 16 out of 18 coefficients are statistically insignificant as in the benchmark model, and the remaining two coefficients (in Column 3) are only significant at the 10% level and enter negatively. Hence, there is very little evidence corroborating the hypothesis that population density affects the size of a person's network.

B.4 Robustness: Moving Dynamic and Within-Degree

Table B.4 compares the benchmark model (Column 0) relating population density and *within-degree* to estimates obtained from panel FE regressions that exclude periods around the moving months (Column 1 to 5). For instance, Column (1) shows estimates obtained from regressions excluding the moving month ($t \neq 0$), while in Column (2) the moving month plus the two adjacent months ($t \leq -2 \vee t \geq 2$) are excluded. The results of our benchmark analysis in Table 5 remain valid: population density enlarges a person's local network. 14 out of 15 coefficients are statistically significant as in the benchmark model. Magnitude-wise the estimates from the benchmark model in Column (0) are very similar



Notes: We regress degree centrality of movers on dummies for the months leading and following the moving date. The reference category is the moving date $t = 0$ which we define as the first month at the new location. The lines illustrates the 95 percent confidence bounds around the point estimates.

Figure B.2: The Degree prior and after Moving

to the one obtained in these additional regressions.

B.5 Robustness: Moving Dynamic and Clustering

Table B.5 compares the benchmark model (Column 0) relating population density and *clustering* to estimates obtained from panel FE regressions that exclude periods around the moving months (Column 1 to 5). It turns out that the negative effect of population density on a person's clustering coefficient is very robust. 12 out of 15 coefficients are statistically significant as in the benchmark model (see Table 7). Magnitude-wise the estimates from the benchmark model are not significantly different from most of the estimates obtained in these additional regressions. However, we observe a tendency towards larger effects once observations around the moving month are excluded.

B.6 Robustness: Matching

The robustness exercise in Table B.6 concerns the inferred impact of population density on matching quality. We performed additional analyses both in the network formation framework (panel a.) and the network topography framework (panel b.).

Table B.3: Robustness – Cities and Network Size

Dependent Variable: $D_{ir,t}$	All Months	Excluding Months around Change of Residence				
		$t \neq 0$	$t \leq -2 \vee t \geq 2$	$t \leq -3 \vee t \geq 3$	$t \leq -4 \vee t \geq 4$	$t \leq -5 \vee t \geq 5$
	(0)	(1)	(2)	(3)	(4)	(5)
Hinterland (vs. Cities)	-0.006	-0.007	-0.008	-0.017^{+}	-0.012	-0.015
	(0.006)	(0.006)	(0.007)	(0.009)	(0.012)	(0.020)
Periphery (vs. Cities)	-0.001	0.001	0.002	-0.001	-0.005	-0.027
	(0.007)	(0.007)	(0.009)	(0.011)	(0.015)	(0.024)
R^2	0.011	0.011	0.011	0.010	0.009	0.011
Groups	16,874	16,868	16,808	16,743	16,681	16,535
Observations	185,644	167,761	138,883	113,106	90,018	69,675
Ln(Population Density)	-0.006	-0.008	-0.011	-0.048^{+}	-0.030	-0.094
	(0.017)	(0.019)	(0.023)	(0.027)	(0.038)	(0.061)
Ln(Population Density)^2	0.000	0.000	0.000	0.002	0.002	0.005
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
R^2	0.011	0.011	0.011	0.010	0.009	0.008
Groups	16,874	16,868	16,808	16,743	16,680	16,534
Observations	185,644	167,749	138,872	113,097	90,011	69,670
Further Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Language Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The *sample* covers movers (minimum moving distance 30min) who used their phone every month at least once. Column (1) excludes the moving month; column (2) excludes the moving month and the first month prior and after moving; and so on. *Further controls* include commuting distance and a dummy for belonging to language minority. Standard errors in parentheses. $^{+}$ p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

Table B.4: Robustness – Cities and the Within-Degree

Dependent Variable: $DW^{\tau}_{i,t}$	All Months	Excluding Months around Change of Residence				
		$t \neq 0$	$t \leq -2 \vee t \geq 2$	$t \leq -3 \vee t \geq 3$	$t \leq -4 \vee t \geq 4$	$t \leq -5 \vee t \geq 5$
	(0)	(1)	(2)	(3)	(4)	(5)
Hinterland (vs. Cities)	-0.038*	-0.039*	-0.049*	-0.058*	-0.049	-0.084^{+}
	(0.018)	(0.018)	(0.021)	(0.025)	(0.032)	(0.046)
Periphery (vs. Cities)	-0.132***	-0.133***	-0.149***	-0.148***	-0.148***	-0.194***
	(0.020)	(0.021)	(0.024)	(0.028)	(0.036)	(0.053)
R^2	0.012	0.015	0.017	0.016	0.015	0.012
Groups	16,874	16,868	16,808	16,743	16,681	16,535
Observations	185,676	167,761	138,883	113,106	90,018	69,675
Ln(Population Density)	0.076***	0.077***	0.082***	0.085***	0.087***	0.087***
	(0.006)	(0.006)	(0.007)	(0.008)	(0.010)	(0.015)
R^2	0.016	0.019	0.021	0.020	0.018	0.013
Groups	16,874	16,868	16,808	16,743	16,680	16,534
Observations	185,663	167,749	138,872	113,097	90,011	69,670
Further Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Language Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The *sample* covers movers (minimum moving distance 30min) who used their phone every month at least once. Column (1) excludes the moving month; column (2) excludes the moving month and the first month prior and after moving; and so on. *Further controls* include commuting distance and a dummy for belonging to language minority. Standard errors in parentheses. $^{+}$ p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

The top panel of Table B.6 re-estimates these benchmark model according to (10b.) but for a *sample of movers that stay within the same DEGURBA-class*, i.e. people who move from city to another city, from hinterland to hinterland, or from peripheral munic-

Table B.5: Robustness – Cities and Clustering

Dependent Variable: $C_{i,t}$	All Months	Excluding Months around Change of Residence				
		$t \neq 0$	$t \leq -2 \vee t \geq 2$	$t \leq -3 \vee t \geq 3$	$t \leq -4 \vee t \geq 4$	$t \leq -5 \vee t \geq 5$
	(0)	(1)	(2)	(3)	(4)	(5)
Hinterland (vs. Cities)	0.002^{+}	0.002^{+}	0.002	0.004*	0.003	0.002
	(0.001)	(0.001)	(0.001)	(0.002)	(0.003)	(0.004)
Periphery (vs. Cities)	0.002^{+}	0.003*	0.003*	0.006**	0.006*	0.008^{+}
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.004)
R^2	0.001	0.001	0.001	0.001	0.001	0.01
Groups	16,870	16,863	16,802	16,735	16,670	16,518
Observations	183,896	166,130	137,489	111,965	89,099	68,953
Ln(Population Density)	-0.001^{+}	-0.001^{+}	-0.001^{+}	-0.001^{+}	-0.001^{+}	-0.003*
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
R^2	0.001	0.001	0.001	0.001	0.001	0.001
Groups	16,870	16,863	16,802	16,735	16,669	16,517
Observations	183,896	166,118	137,478	111,956	89,092	68,948
Further Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Language Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The *sample* covers movers (minimum moving distance 30min) who used their phone every month at least once. Column (1) excludes the moving month; column (2) excludes the moving month and the first month prior and after moving; and so on. *Further controls* include commuting distance and a dummy for belonging to language minority. Standard errors in parentheses. $+$ p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

ipality to another peripheral municipality. Then, we use the predicted matching quality $\widehat{m}(\xi_i, \xi_j, \delta)$ again and regress it on measures of population density. As in the main analysis, we estimate the model for both the full sample of movers and those with a minimum moving distance of 30min. Qualitatively, this does not affect the results: The fixed-effect component is larger among urban residents, which suggests that matching quality increases with population density. Magnitude-wise the coefficients of interest and the constants get substantially larger than in the benchmark specifications reported in Table 6, while the ratio remains very similar: Comparing urban residents to people living in the periphery, column (1) of Table B.6 suggests that the matching quality is 7% higher in the city than in the periphery, while column (3) suggests that the difference amounts to 4.7%.[31] In the benchmark model in Table 6 this coefficient-constant ratio adds up to about 5.8%.

In the bottom panel of Table B.6 we show robustness results on the social adjustment process of movers. In the benchmark models we controlled for the number of existing links at the new address, while the results in Table B.6 are based on a sample of movers *without any contacts* at the post-move residence prior to moving. Put differently, in this robustness check we exclude all movers that had any pre-move contact with people living within a 15 minutes radius of their new place of residence. Qualitatively, we obtain the same pattern as in the benchmark estimations: Agents that move from the periphery/hinterland to the

[31] The avg. measure for matching quality drops by about $382/5447 \simeq 0.07$ and $252/5347 \simeq 0.047$ when comparing the periphery to cities in columns (1) and (3).

city replace pre-move contacts substantially faster than people moving from the city to the periphery/hinterland. Since maintaining spatially distant contacts is costly, this suggests that contacts formed in cities generate on average a higher surplus and are therefore more likely to be maintained. Hence, this test also supports the hypothesis that densely populated areas improve matching quality

Table B.6: Robustness – Regional Differences in the Matching Quality

a. Network Formation	Full Sample		Moving Distance > 30min.	
Dependent Variable: $\widehat{m}(\xi_i, \xi_j, \delta)$	(1)	(2)	(3)	(4)
$\text{Hinterland}_{i,t}$	-101.581***		-44.090	
	(16.083)		(48.718)	
$\text{Periphery}_{i,t}$	-381.461***		-252.426***	
	(18.874)		(50.975)	
$\text{Ln(Pop. Density}_{i,t})$		88.509***		56.085***
		(5.083)		(9.318)
Constant	5447.207***	4439.930***	5346.683***	4718.905***
	(13.150)	(50.016)	(44.029)	(86.166)
R^2	0.002	0.001	0.001	0.001
Observations	1,631,708	1,646,566	306,798	313,072
b. Network Topography	Full Sample		Moving Distance > 30min.	
Dependent Variable: $DW^{post}_{i,t+1}/DW^{pre}_{i,t+1}$	(1)	(2)	(3)	(4)
City^{post}	0.308***		0.202***	
	(0.046)		(0.056)	
City^{pre}	-0.231***		-0.275***	
	(0.033)		(0.045)	
$\text{Ln(Pop. Density}^{post})$		0.184***		0.099***
		(0.013)		(0.013)
$\text{Ln(Pop. Density}^{pre})$		-0.145***		-0.103**
		(0.016)		(0.031)
Constant	0.738***	0.304+	0.754**	0.670***
	(0.124)	(0.176)	(0.282)	(0.033)
R^2	0.017	0.041	0.018	0.005
Further Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Language Region FE	Yes	Yes	Yes	Yes
Observations	5,718	5,718	3,108	3,194

Notes: *Dependent Variable in Panel a.:* Predicted dyad specific fixed effect from network formation model outlined in equation (10b). *Dependent Variable in Panel b.:* The number of *post-move* contacts at the *post-move* place of residence over the number of *post-move* contacts at the *pre-move* place of residence. *Controls in Panel b.:* Number of contacts at new address prior to moving, commuting distance, dummy for belonging to language minority, and Romansh region), gender and age . Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

An alternative way to identify asymmetries across cities and urban areas in the network adjustment following a relocation builds on dyad specific information $g_{ij,t}$. We expect that high quality matches remain after a relocation while low quality matches drop out relatively quickly. In the following we study whether the network of *stayers* adjusts differently to the relocation of a contact who moves from the periphery to the city compared to one that moves from the city to the periphery. In the former case the set of potential high-quality links in close neighborhood expands for the mover such that we expect the stayer

to drop out of the network with a relatively high likelihood. In the latter case the mover has established the link among a large set of alternatives such that $g_{ij,t-1} = 1$ should represent a relatively high-quality link and should remain relatively stable. Hence, we estimate a network formation model limited to the sample of movers i and stayers j:

$$g_{ij,t+1} = \beta g_{ij,t-1} + \beta_2 g_{ij,t-1} L^{pre}_{i,t-1} T_{ij,t+1} + U_{ij,t+1} \; \forall \; i : movers, j : stayers, \tag{B.1}$$

where we define again $t-1$ as the pre-move period, t as the moving period, and $t+1$ as the post-move period. $L^{pre}_{i,t-1}$ captures information about the movers place of origin which is either a dummy variable on whether the place is a city or the places' population density. Note that all unobservable information about individual characteristics determining link formation in the pre-move period, i.e. $\nu_{i,t-1}$ and $\nu_{j,t-1}$ are absorbed by the lagged dependent variable. Table B.7 shows that links of movers relocating from a city are in fact significantly more stable. The persistence of links formed in cities is about 5.3 percent higher than those formed in the hinterland or the periphery supporting the hypothesis that links established at high density places are characterized by a higher quality of matching and thus persistence.[32] This finding is robust to using population density instead of a city dummy and to focusing on movers who changed their residence by at least 30 minutes driving time.

Table B.7: Robustness – Bilateral Network Adjustment

Network Formation	Full Sample		Moving Distance > 30min.	
Dependent Variable: $g_{ij,t+n} \; \forall \; i : movers, j : stayers$	(1)	(2)	(3)	(4)
Ln(Travel Time$_{ij,t+1}$)	-0.047***	-0.047***	-0.041***	-0.041***
	(0.000)	(0.000)	(0.000)	(0.000)
$g_{ij,t-1}$	2851.853***	2323.578***	2825.895***	2511.691***
	(2.709)	(18.475)	(5.079)	(28.578)
$g_{ij,t-1} \times City^{pre}_j$	149.639***		101.198***	
	(5.750)		(12.394)	
$g_{ij,t-1} \times$Ln(Pop. Density$^{pre}_j$)		58.037***		35.932***
		(1.895)		(3.650)
Constant	0.237***	0.237***	0.208***	0.208***
	(0.000)	(0.000)	(0.001)	(0.001)
R^2	0.184	0.184	0.178	0.178
Observations	16,016,369	16,012,515	4,243,460	4,239,751

Notes: We use monthly data for June 2015–May 2016. The *sample* covers movers and stayers who used their phone every month at least once. All *coefficients* of the linear probability models are multiplied by 10000, and therefore can be interpreted as basis points. Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

[32] According to column 1 the lagged dependent variable indicates an increase of about $150/2852 \simeq 0.053$.

B.7 Robustness: Technology Preferences

One concern may be that urban residents use technology differently than people living in rural areas, which may confound our population density / city dummy estimates. For instance, it may be that messenger apps such as WhatsApp are used more frequently in cities. In this subsection we test the sensitivity of our network topography results with respect to technology usage by including two measures that account for changes in technology usage of movers. First, we use the ratio of outgoing SMS versus outgoing calls, because traditional text messages are the most likely technology to be substituted by messenger apps. In contrast, the pricing schemes of Swiss phone contracts do not invite to substitute calls, as phone companies primarily price discriminate based on data volume and speed rather than cost per call. Second, we include the ratio of outgoing landline calls versus the total number of calls, because apps may be used to call another mobile phone but not landlines. Hence, if moving from rural areas to the city comes along with an increase in messenger app usage, this should reflect in a decrease in the SMS-call ratio and an increase in the share of landline calls.

Table B.8 adds the two technology proxies to our benchmark models shown in Tables 4 (columns 5 and 6), 5b (columns 3 and 4), 7b (columns 3 and 4) and 6a (columns 1 and 2). The main results are not affected by this exercise: density remains to have not significant effect on network size whereas it has a positive and significant effect on within-degree and matching quality. The effect of population density on clustering remains negative and significant.

B.8 Robustness: Endogeneity of Population Distribution

A final robustness concerns the role of unobserved location factors that may affect network characteristics as well as population density. We follow the literature on urban wage premia and instrument population density using historical population counts and measures of soil quality (see Ciccone and Hall (1996) and Combes et al. (2010) for details). Hence, we exploit only variation in population density which is determined by exogenous factors. Neither historical population counts (year 1850) nor soil quality are likely to have a direct effect on social network characteristics measured today. Yet, both instruments are a strong predictor of population density today as is evident from the first-stage regressions. Soil quality used to be an important location advantage in former times when agriculture played an important role. Historical population differences persisted while soil quality ceased to be an important factor for location choice in Switzerland. The corresponding two-stage least square results for the network topography approach and the outcomes Degree, Within-Degree, Clustering and Matching Quality are presented in Table B.8. The

estimates confirm an insignificant effect of population density on network size; positive and significant effects on within degree and matching quality; negative and significant effects on clustering. The magnitude of the coefficients is in line with the corresponding estimates in Tables 4 (column 6), 5b (column 4), 7b (column 4) and 6a (column 2).

Table B.8: Robustness – Controlling for Technology Preferences

	Degree		Within Degree		Clustering		Matching	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hinterland (vs. Cities)	-0.001		-0.124***		0.002*		-64.217***	
	(0.003)		(0.010)		(0.001)		(5.997)	
Periphery (vs. Cities)	-0.002		-0.232***		0.002**		-148.797***	
	(0.004)		(0.011)		(0.001)		(6.523)	
Ln(Pop. Density)		-0.001		0.144***		-0.001**		35.148***
		(0.001)		(0.004)		(0.000)		(1.875)
SMS/Calls	-0.385***	-0.385***	-0.234***	-0.234***	0.001	0.001	-282.707***	-279.417***
	(0.005)	(0.005)	(0.008)	(0.008)	(0.001)	(0.001)	(10.648)	(10.620)
Calls to MP/Total Calls	0.499***	0.499***	0.266***	0.267***	-0.010***	-0.010***	511.954***	512.618***
	(0.004)	(0.004)	(0.006)	(0.006)	(0.001)	(0.001)	(9.671)	(9.647)
R^2	0.067	0.067	0.017	0.025	0.001	0.001	0.003	0.003
Further Controls	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Language Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Groups	60,514	60,514	60,514	60,514	60,507	60,507	-	-
Observations	669,825	669,812	669,825	669,812	664,343	664,330	11,404,385	11,479,868

Notes: We use monthly data for June 2015–May 2016. The *sample* covers movers who used their phone every month at least once. *Further controls* include commuting distance, language (pooled OLS), dummy for belonging to language minority, gender (pooled OLS), and age (pooled OLS). Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.

Table B.9: Robustness – Endogenous Population Density

	Degree		Within Degree		Clustering		Matching	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Pop. Density)	-0.003	-0.001	0.139***	0.132***	-0.002**	-0.001*	44.705***	38.368***
	(0.002)	(0.002)	(0.004)	(0.003)	(0.001)	(0.000)	(2.829)	(2.561)
R^2	0.067	0.067	0.017	0.025	0.001	0.001	0.003	0.003
Instrument	Pop. 1850	Soil	Pop. 1850	Soil	Pop. 1850	Soil	Pop. 1850	Soil
Further Controls	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Language Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Groups	59,356	60,477	59,356	60,477	59,348	60,468	-	-
Observations	632,613	665,443	632,613	665,443	627,360	660,006	10,911,638	11,610,575

Notes: We use monthly data for June 2015–May 2016. The *sample* covers movers who used their phone every month at least once. *Further controls* include commuting distance, language (pooled OLS), dummy for belonging to language minority, gender (pooled OLS), and age (pooled OLS). Instruments are population density in 1850 (Pop. 1850) and soil quality (Soil) which turn out highly relevant in all first-stage regressions. Standard errors in parentheses. + p<0.10, * p<0.05, ** p<0.01 *** p<0.001.