DISCUSSION PAPER SERIES

No. 10501

URBAN SOCIAL STRUCTURE, SOCIAL CAPITAL AND SPATIAL PROXIMITY

Eleonora Patacchini, Pierre M. Picard and Yves Zenou

PUBLIC ECONOMICS

Centre for Economic Policy Research

ISSN 0265-8003

URBAN SOCIAL STRUCTURE, SOCIAL CAPITAL AND SPATIAL PROXIMITY

Eleonora Patacchini, Pierre M. Picard and Yves Zenou

Discussion Paper No. 10501
March 2015
Submitted 12 March 2015

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

URBAN SOCIAL STRUCTURE, SOCIAL CAPITAL AND SPATIAL PROXIMITY†

Abstract

We develop a theoretical model where the existence and intensity of dyadic contacts depend on location. We show that agents tend to interact more with agents that are highly central in the network of social contacts and that are geographically closer. Using a unique geo-coded dataset of friendship networks in the United States, we find evidence consistent with this model. The main empirical challenge, which is the possible endogenous network formation, is tackled by employing a Bayesian methodology that allows to estimate simultaneously network formation and intensity of network contacts.

JEL Classification: R1, R23 and Z13
Keywords: Bayesian estimation, endogenous network formation, geographical space, social interactions and social space

Eleonora Patacchini ep454@cornell.edu
Cornell University, EIEF and CEPR

Pierre M. Picard pierre.picard@uni.lu
CORE, Université Catholique de Louvain and CREA, University of Luxembourg

Yves Zenou yves.zenou@ne.su.se
Stockholm University and IFN, and CEPR

---

† We are grateful to J. Brueckner and P. Maruhenda for insighful discussions and inspiring comments. Yves Zenou gratefully acknowledges the financial support from the French National Research Agency grant ANR-12-INEG-0002.

1 Introduction

Social interactions and social networks are important components of social and economic life. While they have been the focus of sociological research for many decades, social interactions and networks have attracted the attention of economic research only when they were recognized as a critical engine for the economic growth of nations and regions (e.g. Romer, 1986; Lucas 1988). If individual agents are conceived as socially related, networks can shape their preferences, and hence actions. The economics literature in the last two decades shows a great deal of network analysis to understand economics decision making.[1] Surprisingly, key questions such as how important is proximity and the characteristics of places on the construction and intensity of network exchange are yet unanswered. For Glaeser (2000), the existence of cities critically hinges on how social interactions and networks can be facilitated across the space of urban entities. However, traditional models in urban economics (Fujita, 1989) do not consider the presence of social interactions and social capital in cities. On the other hand, most papers from the network economics literature (implicitly) assume that the existence and intensity of dyadic contacts do not depend on location, i.e. do not consider the geographical location of the agents.

This paper is the first to develop a theoretical model that studies the interplay between the geographical location of agents and the intensity of their social interactions.

Specifically, we consider a population of agents who entertain social interactions and develop social networks in a unidimensional geographical space (the city). In this city, the fraction of individuals at each location is determined by a distribution function of a general form. A key innovation is that the intensity of social interactions depends on the location of each individual in the city. Indeed, each agent decides the frequency of her visits (social interactions) to every other agent in the city where the value of each interaction depends on the social network of the visited agents. We define the value of such interactions as the *social capital* of the agent (Putman, 2000). Social capital is thus defined in a recursive fashion: it is higher the larger the volume of interactions with highly social individuals. When deciding how much to interact with others, agents face the following trade off. Each agent can increase her social capital by entertaining more frequent interactions with agents who also entertain many interactions. However, entertaining

[1]For recent overviews, see Ioannides (2012), Jackson (2008, 2014), Jackson et al. (2015), Jackson and Zenou (2015), Topa and Zenou (2015), and Zenou (2015).

social interactions requires costly travelling to the other agents so that spatial centrality may help in building up each agent's social capital.

We study the interplay of social and spatial dimensions when the location of agents is taken as given. In other words, we discuss the building of social capital through space without considering its potential impact on land rents. This strategy is imposed by the constraint of our empirical analysis where students do not pay land rents.[2] We show that agents who are located in a more central place (i.e. close to the city center) develop more intense social networks. Spatial centrality is thus positively correlated with social capital. Our model has also the following implications: (i) Lower travel costs increase social capital for all agents. This means that a spatial concentration of agents increases social capital for all agents in the city. (ii) An increase in the importance of peers' social links, increases each agent's social capital for small enough travel costs. (iii) Cities that are more "spread" in terms of the geographical distribution of their residents have lower social capital for all agents.

Our analysis suggests that better urban transport facilities are likely to enhance social capital in cities, which is a new implication compared to what urban economics usually predicts.[3] This type of policies may be particularly important under the view that social interactions may promote economic growth (Glaeser, 2000; Ioannides, 2012). Indeed, in our modelling framework, the frequency of social interactions and the level of social capital is inefficiently low. This is because each agent creates a positive externality on others. When she exerts more effort in social interaction, she increases her social capital and therefore enhances the benefit of her interaction partners. We then test the main predictions of the model using a unique geo-coded dataset of friendship relationships among students in the United States. Existing empirical studies on the relevance of geographical location for social interactions in real world networks are almost non existent.[4] In fact, it is extremely difficult to find detailed data on social contacts as a function of

[2]Even if we consider the land rent paid by their parents, we do not have this information in the data.

[3]In the standard monocentric models (Fujita, 1999) and in their multicentric extensions (Fujita and Thisse, 2013), unit travel cost is usually the fundamental parameter that determines the location choices of households within cities, their consumption of housing, land use, and the population size of cities. As a result, transportation policies that reduce commuting costs in the city have a direct impact of these outcomes. Recent research (see, in particular, Duranton and Turner, 2011, 2012; Couture et al., 2013) provide precise estimates for this fundamental parameter and show empirically how it varies with population and road infrastructure. None of the existing papers, however, look at the importance of commuting costs for social interactions.

[4]Recent research on online social networks indicates that spatial proximity appears influencing how people con-

geographical distance between agents together with information on relevant socio-economic characteristics. Some evidence can be found in Marmaros and Sacerdote (2006). Using data on email communication between Dartmouth college students, this paper shows that being in the same freshman dorm increases the volume of interactions by a factor of three.

Our empirical investigation is made possible by the use of data from the National Longitudinal Survey of Adolescent Health (AddHealth) in the United States. While the AddHealth data has been used extensively for its social network information based on friends' nomination (see e.g. Calvó-Armengol et al., 2009), this dataset contains another unique information that has not been exploited before. Indeed, the AddHealth also provides longitude and latitude coordinates of the residential location of each respondent. As a result, it is possible to obtain information not only on the precise geometry of social contacts and the strength of their social interactions, but also on the geographical distance between them. Using this dataset, we then test the main predictions of our theoretical model, namely: (i) how the geographical distance between the students and the social capital of the students affect the intensity of social interactions between them; (ii) how the social capital of each student is affected by the social capital of other students.

We explicitly model a possible endogeneity of the social network of social contacts by allowing for the presence of unobserved factors affecting both the intensity of social interactions and social capital formation. Indeed, the intensity of social interactions can be chosen at the same time as the choice of friends with high social capital. We use a Bayesian approach by simultaneity estimating these two choices. Our results show that, indeed, students residing far away from each other tend to interact less and more central students in their friendship network tend to contribute more to social capital than less central students.

The rest of the paper unfolds as follows. In the next section, we describe the contribution of our paper to the existing literature. Section 3 develops the theoretical model. In Section 4.2, we describe our data and provides some descriptive statistics. Section 4 is devoted to the empirical analysis where we describe our empirical methodology and give the empirical results. Finally,

nect to each other. Using an extensive daset of 100,000 Facebook network users, Goldenberg and Levy (2009) document that both the density of social network contacts and the volume of email traffic decreases with geographical distance. Kaltenbrunner et al. (2012) study the relationship between online user interactions and geographic proximity with a detailed study of a large Spanish online social service. Their results show that while geographic distance strongly affects how social links are created, spatial proximity plays a negligible role on user interactions.

Section 5 concludes the paper. All the proofs of the propositions can be found in Appendix A. In Appendix B, we show that the equilibrium social interactions are not efficient. In Appendix C explains the technical details of the Bayesian estimation.

2 Related literature

Our theoretical framework provides a bridge between two literatures: the traditional urban models and the recent social network models. On the one hand, while there is a common recognition of cities as a major places for social interactions, traditional urban models do not consider the presence of social interactions and social capital. On the other hand, network papers (implicitly) assume that the existence and intensity of dyadic contacts do not depend on location, i.e. do not consider the geographical location of the agents.

Let us review each of these literatures while highlighting the contribution of this paper.

Urban economics and agglomeration: Theory There is an important literature in urban economics looking at how interactions between agents create agglomeration and city centers.[5] It is usually assumed that the level of the externality that is available to a particular firm or worker depends on its location relative to the source of the external effect – the spillover is assumed to attenuate with distance – and on the spatial arrangement of economic activity. This literature (whose keystones include Beckmann, 1976; Fujita and Ogawa, 1980; Lucas and Rossi-Hansberg, 2002; Behrens et al., 2014; Helsley and Strange, 2014) examines how such spatial externalities influence the location of firms and households, urban density patterns, and productivity. For example, Glaeser (1999) develops a model in which random contacts influence skill acquisition, while Helsley and Strange (2004) consider a model in which randomly matched agents choose whether and how to exchange knowledge. Similarly, Berliant et al. (2002) show the emergence of a unique centre in the case of production externalities. These models provide an interesting discussion of spatial issues in terms of use of residential space and formation of neighborhoods and show under which condition different types of city structure emerge.

In this paper, we consider a different view. While the literature cited above aims at explaining different urban configurations (monocentric versus polycentric cities) and to derive under which

[5] See Fujita and Thisse (2013) and Duranton and Puga (2015) for excellent literature reviews.

conditions they emerge,[6] we take the urban configuration as given and explain how the location of each agent in the city affects his/her social interactions with other agents in the city. In other words, we simplify the urban configuration of the city but we open the black box of interactions by examining how and why they form.

Social networks: Theory Most of the papers looking at network formation and economic decision making over network structures assume away agents' geographical location. There is, however, a small strand of the literature that studies the incentives to entertain interactions over exogenous sets of social links. This paper thus parallels recent research efforts by Johnson and Gilles (2000), Brueckner and Largey (2008), Helsley and Strange (2007), Zenou (2013), Mossay and Picard (2011, 2013), Helsley and Zenou (2014), Sato and Zenou (2015) who study the theoretical relationship between social network and urban structure in simple city models that include a few location points or in which agents have similar access costs. However, in this literature the social network is usually not directly related to the geographical location of agents. On the contrary, in the present paper, agents can locate anywhere in the city and we show how their location in the geographical space affects their social interactions with other agents located in the city.

Urban and social networks: Empirics While from a theoretical viewpoint, the literature on the interaction between the social space and the geographical space is relatively small, from the empirical viewpoint, it is quasi-inexistent.[7] There are some implicit tests where the social space is indirectly measured. For example, Bayer et al. (2008) consider spatial clustering of individual work locations for a given residential location, as evidence of local referral effects. In order to separately identify labor market referrals from other spatially correlated effects, the authors estimate the excess propensity to work together (in a given city block) for pairs of workers who co-reside in the same city block (distinct from their work location), relative to the baseline propensity to work together for residents in nearby blocks (within a reference group of blocks). They find that residing in the same block raises the probability of sharing the work location by 33%, consistent with local referral effects. Hellerstein et al. (2011, 2014) build on the identification

[6] For example, Fujita and Ogawa (1980), a prominent paper in this literature, consider a "locational potential function" in which a weighted average of pairwise Euclidean distances between firms has a negative effect on firms' profit. This acts as an agglomeration force for firms because it implies a (strictly) penalty cost for firm dispersion.

[7] For recent surveys, see Ioannides (2012), Topa and Zenou (2015), Overman et al. (2015).

strategy of Bayer et al. (2008) using matched employer-employee data at the establishment level. They also find that residential labor market network effects are stronger within than across races, suggesting racial stratification within residential social networks.[8]

However, none of the papers mentioned above use data on the precise geometry of social contacts or on detailed geographical distance between them and do not focus on the same outcomes. In addition, their focus is on the importance of location and networks for labor-market outcomes. This is the first paper that provides evidence on the importance of location for the intensity of social interactions in real-word networks.

3 The model

3.1 Notations and definitions

Consider a *linear city* on the line segment $x \in [-b, b]$ where b is the city border, and let $\lambda(x) : [-b, b] \to R^+$ measure the number of agents located at x. We focus on a city with unit mass population: $\int_{-b}^{b} \lambda(y)\mathrm{d}y = 1$.

Each agent *visits every other agent* and benefits from social interactions. First, the utility from social interactions is given by

$$S(x) = \int_{-b}^{b} v\left(n(x, y)\right) s(y)\lambda(y)\mathrm{d}y$$

where $n(x, y)$ is the number or, more exactly, the *frequency* of interactions that agent at x initiates with an agent at y who offers an interaction value $s(y)$. For the sake of tractability, we assume that

$$v\left(n\left(x, y\right)\right) = n\left(x, y\right) - \frac{1}{2}\left[n\left(x, y\right)\right]^2 \tag{1}$$

This expression assumes decreasing returns to the frequency of interactions with a given agent; it even assumes negative returns (saturation) above $n = 1$.

Second, each agent located at x incurs a cost of visiting another agent residing at y, $c(x - y)$, which is symmetric and increases with distance $|x - y|$: $c(z) = c(-z)$ and $c'(z) > 0\ \forall z > 0$. For simplicity, we consider the class of travel cost functions $c(x)$ that are differentiable except at

[8] See also Schmutte (2015).

$x = 0$. We define the slope at $x = 0$ as $c'_+(0) \equiv \lim_{x\to 0, x>0} c'(0) \geq 0$, recognizing the possible kink at $x = 0$. The total social interaction cost of an agent located at x is given by

$$C(x) = \int_{-b}^{b} n(x,y)c(x-y)\lambda(y)\mathrm{d}y$$

which increases with the number of social interactions.

Third, the interaction value offered by an agent residing at y is assumed to be equal to

$$s(y) = 1 + \alpha \int_{-b}^{b} n(y,z)s(z)\lambda(z)\mathrm{d}z \tag{2}$$

The first constant term (normalized to 1) represents the idiosyncratic interaction value that the agent located at y provide to her visitors. The second term, $\alpha \int_{-b}^{b} n(y,z)s(z)\lambda(z)\mathrm{d}z$, reflects the value of her social network for her visitors. It increases with $n(y,z)$, the number of interactions, and $s(z)$, the value of her interactions. The parameter $\alpha > 0$ measures the importance of others' social capital in an agent's social capital formation. The higher is α, the higher is the impact of the social network of "friends of friends". We refer to $s(y)$ as the *social capital* of the agent located at y.

The social capital function $s(y)$ defined in (2) can be interpreted in various ways according to the context under discussion. In the context of information transmission (for example, about job opportunities) and/or knowledge (about a product or technique), the first term may represent the information endowed to or produced by the agent located at y while the second term may reflect the information she received during her visits to other agents. The parameter α then measures the imperfection of information transmission and its retention. In the context of a service sector like advertising, law, etc. (Arzaghi and Henderson 2008), the first term represents the idiosyncratic productivity of a firm located at y while the second term reflects the potential and the ability to quickly subcontract parts of a project to other competent firms. In the context of friendship, community or political participation, the first term gives a measure of the pleasure or interest in a specific interaction (e.g. with a college friend, priest or politician) while the second term may reflect the sense of belonging to a community (e.g. alumni, confession or political group).

We now consider the question of how social capital is distributed across space when agents are exogenously located.

3.2 Social capital and space

We assume that λ, the population density at each location, is exogenously fixed. Each agent located at x chooses the profile of interactions $n(x,\cdot)$ that maximizes her utility

$$U(x) = S(x) - C(x) = \int_{-b}^{b} \{v\,(n(x,y))\,s(y) - n(x,y)c(x-y)\}\,\lambda(y)\mathrm{d}y$$

Note that her utility depends on the profile of other agent's social capital $(s(y), y \neq x)$. It also depends on her own social capital $(s(y), y = x)$ but only on a set of measure zero.[9] As a result, the optimal number of interactions of an agent located at x depends only on the social capital $s(y)$ of the other agents located at y at a non-zero distance to her. The optimal number of interactions $n^*(x,y)$ of an agent located at x (that we call agent x) is therefore found by *differentiating pointwise* $U(x)$ with respect to $n(x,y)$, taking $s(y)$ as given. This pointwise differentiation yields:

$$v'\,(n^*(x,y))\,s(y) - c(x-y) = 0.$$

Using (1), this is equivalent to:

$$[1 - n(x,y)]\,s(y) = c(x-y).$$

So, the optimal number of interactions is equal to:

$$n^*(x,y) = 1 - \frac{c(x-y)}{s(y)}. \tag{3}$$

The number of interactions $n^*(x,y)$ of individual x with an agent located at y (i.e. agent y) increases with y's social capital and decreases with the distance between x and y. For simplicity, we assume away corner solutions and assume *global interactions* so that agents interact with every other agent in the city, i.e.

$$n^*(x,y) > 0 \Leftrightarrow s(y) > c(x-y),\ \forall x, y$$

A sufficient condition for this inequality to hold is

$$\min_y s(y) > c(2b) \tag{4}$$

[9]Under the assumption that $\lambda(x) < +\infty$, the agent has no incentive to raise her number of interactions $n(x,\cdot)$ to increase her own social capital $s(x)$. In other words, since one agent's social capital benefits "almost" exclusively other agents, an agent has no incentives to be strategic with respect to increasing her own social capital.

Let us define the *access cost measure* as

$$g(y) \equiv \int_{-b}^{b} c(y-z)\lambda(z)\mathrm{d}z, \tag{5}$$

which is lower than the maximum travel cost $c(2b)$. By plugging (3) into (2) and using (5), we obtain the equilibrium level of social capital $s^*(y)$, which is given by:

$$s^*(y) = 1 + \alpha \int_{-b}^{b} s(z)\lambda(z)\mathrm{d}z - \alpha g(y). \tag{6}$$

Integral equations do not often accept simple analytical solutions, if any. Yet, under the above utility specification, a solution can be obtained. Indeed, integrating $s(z)\lambda(z)$ and simplifying, we obtain:

$$\int_{-b}^{b} s(z)\lambda(z)\mathrm{d}z = \frac{1}{1-\alpha}\left[1 - \alpha \int_{-b}^{b} g(z)\lambda(z)\mathrm{d}z\right]. \tag{7}$$

Inserting this result into (6) yields a closed-form solution for the equilibrium social capital given by:

$$s^*(y) = s_0 - \alpha g(y), \tag{8}$$

where

$$s_0 = \frac{1 - \alpha^2 \int_{-b}^{b} g(z)\lambda(z)\mathrm{d}z}{1-\alpha}, \tag{9}$$

and where $g(y)$ is defined by (5). Under the condition that $0 < \alpha < 1$, the optimal social capital $s^*(y)$ has a finite solution. To guarantee global interactions, we must have $s_0 - \alpha g(y) > c(x-y)$ for all x, y. Using (4), a sufficient condition is

$$s_0 - \alpha \left[\max_y g(y)\right] > c(2b) \tag{10}$$

To summarize,

Proposition 1 *Assume* $0 < \alpha < 1$ *and (10). Then, there exists a unique equilibrium* $(n^*(x,y), s^*(y))$, *defined for all* x, y, *such that*

$$n^*(x,y) = 1 - \frac{c(x-y)}{s^*(y)}$$

and

$$s^*(y) = \frac{1 - \alpha^2 \int_{-b}^{b} g(z)\lambda(z)\mathrm{d}z}{1-\alpha} - \alpha \int_{-b}^{b} c(y-z)\lambda(z)\mathrm{d}z \tag{11}$$

Let us discuss the properties of the equilibrium social capital $s^*(y)$, defined in (11),[10] in a spatial environment.

First, lower travel costs increase social capital for all agents. This conclusion arises simply because social capital increases when the access measure $g(y)$ falls. An upward shift in the travel cost function $c(x)$ raises this access measure and therefore each agent's social capital $s^*(y)$. As a result, travel cost can be seen as a *barrier to social capital formation.* Improvements in urban transportation infrastructure should therefore enhance social capital.

Second, a rise in the importance of peers' social links in the creation of own social capital α, has ambiguous effects. Indeed, differentiating $s(y)$ yields

$$s_\alpha(y) = \int_{-b}^{b} n^*(y,z)s(z)\lambda(z)\mathrm{d}z + \alpha\int_{-b}^{b} n^*(y,z)s_\alpha(z)\lambda(z)\mathrm{d}z + \alpha\int_{-b}^{b} n^*_\alpha(y,z)s(z)\lambda(z)\mathrm{d}z$$

where $s_\alpha(y)$ and $n^*_\alpha(y,z)$ denotes the derivatives of $s(y)$ and $n^*(y,z)$ with respect to α. Thus, an agent's social capital increases with higher α because she values more the social capital of her interaction partners (first term) and because her partners themselves have higher social capital (second term). However, as $n^*_\alpha(y,z) = -c(y-z)s^*_\alpha(z)/\left(s^*(z)\right)^2 \leq 0$, she reduces her frequency of interactions with the partners with higher social capital, which reflects a *substitution effect* between the *frequency* and the *quality* of social interactions (third term). We can get a clearer result by using the optimal frequency of interaction and its associated social capital (6). Differentiating the latter expression with respect to α leads to:

$$s_\alpha(y) = \int_{-b}^{b} s(z)\lambda(z)\mathrm{d}z - g(y) + \alpha\int_{-b}^{b} s_\alpha(z)\lambda(z)\mathrm{d}z. \tag{12}$$

Multiplying this expression by $\lambda(y)$, integrating and simplifying gives:

$$\int_{-b}^{b} s_\alpha(z)\lambda(z)\mathrm{d}z = \frac{1}{(1-\alpha)^2}\left[1 - \int_{-b}^{b} g(z)\lambda(z)\mathrm{d}z\right]$$

Plugging this expression and (7) into (12) yields

$$s_\alpha(y) = \frac{1}{(1-\alpha)^2}\left[1 - \int_{-b}^{b} g(z)\lambda(z)\mathrm{d}z\right] - g(y)$$

As expected, this expression is ambiguous in sign. However, it is positive for small enough access cost measure $g(\cdot)$ and therefore low enough travel costs $c(\cdot)$. We summarize these findings in the following proposition:

[10] Once we know the comparative statics results with respect to $s^*(y)$, then it is straightforward to deduce those of $n^*(x,y)$.

Proposition 2 *Lower travel costs increase social capital for all agents. An increase in α, the importance of peers' social links, increases each agent's social capital for small enough travel cost.*

We now look at the impact on social capital of a *wider geographical dispersion of agents.* Consider a mean preserving increase in the spread of the spatial distribution λ; that is, a change in λ that *second-order stochastically dominates* the present distribution. Expanding expression (8), the social capital $s(y) = s_0 - \alpha g(y)$ can be found to be a linear function of

$$-\left((1-\alpha)\, g(y) + \alpha \int_{-b}^{b} g(z)\lambda(z)\mathrm{d}z \right),$$

which can be rewritten as

$$-\int_{-b}^{b} \left[(1-\alpha)\, c(y-z) + \alpha g(z) \right] \lambda(z)\mathrm{d}z.$$

We can then apply standard results from the analysis of uncertainty. Namely, a mean preserving spread of λ will decrease this expression if the square-bracketed expression is a convex function of z for any y. Conversely, it will increase this expression if the square bracket term is a concave function of z for any y. A sufficient condition for a decrease (resp. an increase) of this expression is that both $c(\cdot)$ and $g(\cdot)$ are convex functions (resp. concave functions). For our class of travel cost functions, we find that

$$g''(x) = \int_{-b}^{b} c''(x-y)\lambda(y)\mathrm{d}y + 2c'_{+}(0)\lambda(x)$$

where $c'_{+}(0)$ is the positive slope at the possible kink of the travel cost function. Therefore $g(\cdot)$ is convex for any travel cost function that is piece-wise linear or convex. This includes linear travel cost $c(x) = c_1 |x|$ and quadratic travel cost $c(x) = c_2 x^2$ where c_1 and c_2 are constants. Intuitively, a spread of the spatial distribution of agents increases the trip distances and costs, which decreases the incentives to interact. So, *larger spatial dispersion of agents reduces social capital in cities.*[11]

Finally, agents located at the urban center have a better access to others and have incentive to increase their social interactions and social capital. One therefore expects that the social capital is less spatially dispersed than the agents. To make this argument formally, let us measure the

[11] Note that general results cannot be obtained for travel cost functions that are piece-wise concave (like $c(x) = 1 - \exp(-|x|)$) because these functions are neither convex or concave.

spatial dispersion of a distribution function ϕ by the ratio of "spatial variance" over its mean value, i.e.

$$\mathrm{Disp}(\phi) \equiv \frac{\int_{-b}^{b} z^2 \phi(z) \mathrm{d}z}{\int_{-b}^{b} \phi(z) \mathrm{d}z}$$

A mean preserving spread of the function ϕ around $x = 0$ increases this dispersion measure because it puts higher values to more distant locations. Under this definition, social capital is less spatially dispersed if and only if $\mathrm{Disp}(s\lambda) < \mathrm{Disp}(\lambda)$. Using (8), it is shown in Appendix A that this is equivalent to $\mathrm{Disp}(g\lambda) < \mathrm{Disp}(\lambda)$. That is, the function $g\lambda$ should be more dispersed than the agent's spatial distribution function λ. We further show that this is true irrespective of the travel cost function g when $x^2\lambda(x)/\int z^2\lambda(z)\mathrm{d}z$ is a mean preserving spread of the distribution of $\lambda(x)$ around its mean $x = 0$. This applies for any uniform spatial distribution λ and for most symmetric spatial distribution functions of interest. We summarize these results in the following proposition:

Proposition 3 *Suppose linear or convex travel cost functions. Then,*

(i) A mean preserving increase in the spread of a symmetric distribution λ decreases social capital for all agents;

(ii) Social capital is less spatially dispersed than agents if $x^2\lambda(x)/\int z^2\lambda(z)\mathrm{d}z$ is a mean preserving spread of the distribution of $\lambda(x)$ around its mean $x = 0$.

The main point of Proposition 3 is to show that, provided that travel costs have appropriate regularity properties, the level and the geographical dispersion of social capital are monotone functions of the dispersion of individuals.

Application to linear travel costs Let us now apply the above analysis to *linear travel costs*, which are heavily used in urban economics for their convenient and realistic properties (see, e.g. Fujita, 1989 or Zenou, 2009). In the present paper, they permit closed-form solutions. Suppose, indeed, that $c(x) = c_1 |x|$ where $c_1 > 0$. Then,

$$g(y) \equiv c_1 \int_{-b}^{y} (y - z)\lambda(z)\mathrm{d}z + c_1 \int_{y}^{b} (z - y)\lambda(z)\mathrm{d}z$$

$$g'(y) = c_1 \int_{-b}^{y} \lambda(z)\mathrm{d}z - c_1 \int_{y}^{b} \lambda(z)\mathrm{d}z$$

$$g''(y) = 2c_1\lambda(y) > 0$$

So, the access cost measure g is a convex function of the distance to the center. Social capital is a concave function that is distributed so that $s''(y) = -2\alpha c_1 \lambda(y) < 0$. Assume further that the spatial distribution of agents λ is *symmetric* ($\lambda(x) = \lambda(-x)$). Then, $g(x)$ is also symmetric and therefore equal to

$$g(x) = g_0 + 2c_1 \int_0^x \int_0^y \lambda(z)\mathrm{d}z\mathrm{d}y, \qquad x \geq 0$$

where $g_0 = 2c_1 \int_0^b z\lambda(z)\mathrm{d}z$. So, for $x \geq 0$, and assuming $0 < \alpha < 1$ and (10), then the unique equilibrium $(n^*(x,y), s^*(y))$ is given by

$$n^*(x,y) = 1 - \frac{c_1 |x-y|}{s^*(y)}$$

and

$$s^*(x) = s_0 - 2c_1 \int_0^x \int_0^y \lambda(z)\mathrm{d}z\mathrm{d}y$$

where

$$s_0 = \frac{1 - 2c_1\alpha^2 \left[\int_0^b z\lambda(z)\mathrm{d}z - 2\int_0^b \left(\int_0^x \int_0^y \lambda(z)\mathrm{d}z\mathrm{d}y\right)\lambda(x)\mathrm{d}x\right]}{1-\alpha}$$

It is clear that lower travel costs c_1 increase social capital for all agents. For small enough travel costs c_1, higher α increases s_0 and therefore each agent's social capital.

The above analysis formalizes the link between the *social capital and spatial centrality.* Travel costs and spatial dispersion of agents are barriers to social capital formation. Social capital tends to be more concentrated than agents themselves. Because of the externalities that agents exert on each other (i.e. each agent does not take into account the effect of her choice of $n^*(x,y)$ and $s^*(x)$ on the other agents in the city), *the equilibrium is clearly not efficient.* We investigate in detail this issue in Appendix B and show that the equilibrium frequency of interactions $n^*(x,y)$ and the level of social capital $s^*(x)$ are lower than the efficient ones. In other words, there are two few social interactions both in quantity and quality compared to the social optimum.

We would like now to *partly* test this theoretical model. In fact, we mainly want to test Proposition 1, i.e. the relationship between the intensity of social relationships $n(x,y)$ and the quality of social interactions $s(y)$ as given by (3) and (2).

4 Empirical evidence

4.1 Empirical model

Proposition 1 describes the equilibrium value of the intensity of social interactions between an agent located at x and an agent located at y

$$n(x,y) = 1 - \frac{c(x-y)}{s(y)} \tag{13}$$

where the social capital of the agent located at y, i.e. she/her interaction value, reflects the interaction value of she/her social contacts,

$$s(y) = 1 + \alpha \int_{-b}^{b} n(y,z)s(z)\lambda(z)\mathrm{d}z. \tag{14}$$

Equation (14) thus indicates that the social capital of an individual depends on the social capital of the individuals to which she/he is connected to. In this respect, it can be seen as a *continuous version* of the *eigenvector centrality,* where an agent's network centrality (or status) is *recursively* related to the centrality (or status) of the agents to which she/he is connected. Suppose that there is a finite number of agents $i, j \in \{1, ...N\}$ in the economy, then, using our notation, the eigenvector centrality is defined as:

$$s(y_i) = \frac{1}{\lambda_1} \sum_{j=1}^{N} n(y_i, z_j) s(z_j) \tag{15}$$

where λ_1 is the largest eigenvalue of the $N \times N$ matrix $n(y_i, z_j)$ that keeps tracks of the interactions between any couple of agents (i, j) respectively located at y_i and z_j (see, e.g. Waserman and Faust, 1994, Jackson, 2008). In (15), $1/\lambda_1$ is the counterpart of the parameter α in (14), which is supposed to be low enough to guarantee a finite values for the social capital $s(y)$.

We bring to the data the following equations:

$$n_{ij} = z_{ij} + \beta_1 d_{ij} + \beta_2 s_j + \varepsilon_{ij} \tag{16}$$

$$s_j = s_0 + \phi \sum_k n_{jk} s_k + u_j \tag{17}$$

Equation (16) is the linear version of equation (13), where we denote by n_{ij} the intensity of interactions between agent i and agent j; d_{ij} captures the *geographical distance* between i and j,

$z_{ij} = \sum_{m=1}^{M} \beta_{0m}(z_{im} - z_{jm})$ captures the effects of social distance in terms of *observable characteristics* z_{im} and z_{jm} (different from geographic proximity) in explaining the intensity of interactions n_{ij}, and ε_{ij} is a random error term. Equation (17) corresponds to equation (15) and denotes the social capital of agent j by s_j where s_0 is a constant term and u_j a random error. Our model predicts that individuals living physically closer to each other should interact more (thus β_1 should be negative), and that more central individuals provide more social interactions that less less central ones (thus β_2 should be positive). In addition, we also expect that under *homophily* behavior (i.e. the tendency of individuals to associate and bond with others who share common traits; see Currarini et al., 2009), more individuals with similar characteristics (same race, same gender, etc.) will tend to interact more than less similar individuals (thus the β_{0m}s should be negative).

4.2 Data

Our empirical analysis is made possible by the use of a database on friendship networks from the National Longitudinal Survey of Adolescent Health (AddHealth).[12] Three features of the AddHealth data set are unique and central to our analysis: *(i)* the nomination-based friendship information, which allows us to reconstruct the precise geometry of social contacts, *(ii)* the detailed information about the intensity of social interactions between each of two friends in the network; and *(iii)* the geo-coded information on residential locations, which allows us to measure the geographical distance between individuals.

The AddHealth database has been designed to study the impact of the social environment (i.e. friends, family, neighborhood and school) on adolescents' behavior in the United States by collecting data on students in grades 7-12 from a nationally representative sample of roughly 130 private and public schools in years 1994-95 (Wave I). The AddHealth Spatial Analysis Data places the homes of a random subset of adolescents for each school (in-home students) in geographic

[12] This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

proximity to a central point within the community and other respondents' homes. Latitude and longitude coordinates are calculated for each home address and then translated into $X-$ and $Y-$coordinates in meters from the central point. By doing so, the spatial distance between all respondents can be determined.

The friendship information is based upon actual friend nominations. Pupils were asked to identify their best friends from a school roster (up to five males and five females).[13] For each individual i, the friendship nomination file also contains detailed information on the frequency and nature of interaction with each nominated friend j. The precise questions are: "Did you go to {NAME}'s house during the past seven days?"; "Did you meet {NAME} after school to hang out or go somewhere during the past seven days?"; "Did you spend time with {NAME} during the past weekend?"; "Did you talk to {NAME} about a problem during the past seven days?"; "Did you talk to {NAME} on the telephone during the past seven days?". "Yes" or "No" are the possible answers. These answers are coded by one and zero, respectively. We measure the intensity of social interactions between students i and j, that is $n(x, y)$ in the model, by summing all these items so that the maximum value of $n(x, y)$ is 5 and the minimum is 0.

Our final sample of in-home students (and their friends) with non-missing information on our target variables consists of 1,567 individuals distributed over 219 networks. This large reduction in sample size with respect to the original sample (about 20,000 individuals) is mainly due to the network construction procedure - roughly 20 percent of the students do not nominate any friends and another 20 percent cannot be correctly linked.[14] In addition, we focus on network sizes between 4 and 50 members as the strength of peer effects may be too different in too small or too large networks (see Calvó-Armengol et al., 2009). The mean and standard deviation of network size are both equal to roughly 7 pupils, thus showing a notable dispersion about this mean value. Table 1 provides a description of the variables used in our study, as well as the summary statistics. Among the adolescents selected in our sample of students, 52% are female and 19% are blacks. Slightly less than 70% live in a household with two married parents, although about 30% come from a single parent family. The average parental education is high school graduate. The performance at school, as measured by the grade point average or GPA, exhibits a mean of 2.9, meaning slightly

[13] The limit in the number of nominations is not binding (even by gender). Less than 1% of the students in our sample show a list of ten best friends.

[14] The representativeness of the sample is preserved. Summary statistics are available upon request.

less than a grade of "B". The average family income is about 50,000 in 1994 dollars, although 11% of parents chose not to report such information. The maximum geographical distance between two students, which is calculated for each school separately, is about 70 kilometers. The average distance is about 7 kilometers, which roughly coincides with the median.

[*Insert Table 1 here*]

4.3 Identification and estimation strategy

From an econometric viewpoint, if the network formation process is exogenous, then equations (16) and (17) can be estimated separately using traditional estimation econometric methods. Indeed, equation (16) could be estimated by OLS, whereas equation (17) is a spatial lag model and requires maximum likelihood estimation to deal with simultaneity issues (see, e.g. Anselin, 1988). If the associations between individuals is exogenous (exogenous network), OLS and ML estimates are consistent.

However, under homophily, linked individuals are likely to be similar not only in terms of *observed characteristics* but also in terms of *unobserved characteristics* that could affect *network fixed effects* and also influence their behavior. To address this issue, most of the existing papers in the social network literature (see, in particular, Bramoullé et al., 2009; Calvó-Amengol et al., 2009; Lin, 2010; Lee et al., 2010; Liu et al., 2012) use the architecture of the networks by introducing *network fixed effects* in the empirical model. The underlying assumption is that the unobservable factors that drive friendship formation are common to all individuals belonging to the same network. However, if there are individual-level unobservables that drive both network formation and outcome choices, this strategy will not work.[15] There is a growing consensus in the social network literature that this is most likely to be the case (Jackson, 2014; Jackson et al., 2015; Graham, 2014, 2015). Recent papers (Goldsmith-Pinkham and Imbens 2013, Hsieh and Lee 2013, Mele 2013, Sheng, 2015) tackle those issues using a Bayesian estimation approach where both network formation and behavior over networks are jointly estimated. We follow this approach here. Network fixed effects are nevertheless still included.

Let us define a vector ξ of individual-level *unobserved characteristics* that is relevant in shaping

[15] For a general discussion and overview on these issues, see Blume et al. (2011), Jackson et al. (2015), Graham (2015).

the intensity of interactions and social capital equations in (16)–(17). In that case, (16) can be written as:

$$n_{ij} = z_{ij} + \beta_1 d_{ij} + \beta_2 s_j + \underbrace{\beta_3(\xi_i - \xi_j) + \epsilon_{ij}}_{\varepsilon_{ij}} \tag{18}$$

where we explicitly model the impact of both observable and unobservable characteristics on the intensity of interactions n_{ij}. Let us also assume that ξ_j is correlated with u_j in Model (17) according to a bivariate normal distribution

$$(\xi_j, u) \sim N\left(\left(\begin{array}{c} 0 \\ 0 \end{array}\right), \left[\begin{array}{cc} \sigma^2_{u_j} & \sigma_{u_j\xi_j} \\ \sigma_{u_j\xi_j} & \sigma^2_{\xi_j} \end{array}\right]\right). \tag{19}$$

Joint normality implies that the error term u_j in equation (17) can be replaced by its expected value $\sigma_{u_j\xi_j}\xi_j$, yielding:

$$s_j = s_0 + \phi \sum_{k \neq j} n_{jk} s_k + \underbrace{\sigma_{u_j\xi_j}\xi_j + v_j}_{u_j}, \tag{20}$$

where v_j is now an i.i.d. error term uncorrelated with the z_js and the unobservable ξ_j. Observe that if $\sigma_{u_j\xi_j} \neq 0$, then s_j is endogenous and the parameter estimates that are obtained when estimating (16)–(17) separately are biased. A solution to this problem is to consider (20) as a linear version of the Selection-Corrected SAR model proposed by Hsieh and Lee (2013), where instead of binary links we have a continuous measure of intensity of interactions. To estimate the system, following Hsieh and Lee (2013), we need to impose[16] that $(u, \xi)|x_i \sim \mathcal{N}(0, \Sigma)$, with $\mathbf{\Sigma} = \left[\begin{array}{cc} \sigma^2_u & \sigma_{u\xi} \\ \sigma_{u\xi} & \sigma^2_\xi \end{array}\right]$. We then rely on Bayesian techniques that allow us to draw samples of ξ from an appropriate conditional distribution.

Bayesian estimation

We define the parameter vector $\mathbf{\Theta} = (\beta, \alpha, \phi, \mathbf{\Sigma}, \sigma^2_\epsilon)$, and an $n \times n$ matrix of intensities of interaction $\mathbf{N} = \{n_{ij}\}$, a vector of social capital $\mathbf{s}$ and a matrix of observables $\mathbf{Z}$. We can then specify our Bayesian probabilistic model as

$$P(\xi, \mathbf{\Theta}|\mathbf{N}, \mathbf{s}) \propto \pi(\xi, \mathbf{\Theta}) \times \mathcal{L}(\mathbf{N}, \mathbf{s}|\xi, \mathbf{\Theta}), \tag{21}$$

[16] As Hsieh and Lee (2013) note, linearity of conditional mean function is enough to obtain $E(u|\xi) = \sigma_{u\xi}\xi$, but normality assumption will be required for estimation anyway.

where $\mathcal{L}(\mathbf{N}, \mathbf{s}|\xi, \mathbf{\Theta})$ is a joint likelihood function and $\pi(\xi, \theta)$ a joint prior distribution.[17] The usual independency assumption allows us to write $\pi(\xi, \mathbf{\Theta}) = \pi_1(\xi) \times \pi_2(\mathbf{\Theta})$ and $\mathcal{L}(\mathbf{N}, \mathbf{s}|\xi, \mathbf{\Theta}) = \mathcal{L}(\mathbf{N}|\xi, \mathbf{\Theta}) \times \mathcal{L}(\mathbf{s}|\mathbf{N}, \xi, \mathbf{\Theta})$. The independence also applies for all elements of $\mathbf{\Theta}$.

Bayesian inference requires the computation of marginal distribution for all parameters. However, since this requires integration of complicated distributions over several variables, a closed form solution is not readily available and Markov Chain Monte Carlo (MCMC) techniques are usually employed to obtain random draws from posterior distributions. The Gibbs sampling algorithm allows us to draw random sample values for each parameter from their posterior distribution, given previous samples of other parameters. Once stationarity of the Markov Chain has been achieved, the random draws can be used to study empirical distributions of the posterior.[18]

We start our algorithm by picking $(\beta^{(1)}, \alpha^{(1)}, \phi^{(1)}, \mathbf{\Sigma}^{(1)}, \sigma_\epsilon^{2(1)})$ as starting values. For $\beta^{(1)}, \alpha^{(1)}, \phi^{(1)}$ we use OLS and ML estimates, while we set the variances $\sigma^{2(1)_u}, \sigma^{(1)_{u\xi}}, \sigma^{2(1)_\epsilon}$ at 2, 1.5 and 1 respectively.[19] In the first step we ought to draw samples of ξ_i from

$$P(\xi_i|\mathbf{N}, \mathbf{s}, \mathbf{\Theta}, \xi_{-i}) \propto \pi(\xi_i) \times \mathcal{L}(\mathbf{N}|\xi_{-i}, \mathbf{\Theta}) \times \mathcal{L}(\mathbf{s}|\mathbf{N}, \xi_{-i}, \mathbf{\Theta}), \quad i = 1, \cdots, n \times n$$

To do this we first draw a candidate ξ_i^c from a normal distribution with mean ξ_i^{t-1}, then we rely on a Metropolis-Hastings (M-H) decision rule: if ξ_i^c is accepted we set $\xi_i^t = \xi_i^c$, otherwise $\xi_i^t = \xi_i^{t-1}$. Once all ξ_i are sampled, we move to the sampling of β in the second step. By specifying a normal prior and a normal likelihood we can now easily sample β^t it from a multivariate normal distribution. A diffuse prior for σ_ϵ^2 allows us to sample it from an inverse chi-squared distribution in the third step.

In the fourth step, we follow the Bayesian spatial econometric literature by sampling ϕ^c from a uniform distribution with support $[-1/\tau, +1/\tau]$, where τ is the highest eigenvalue of $\mathbf{N}$. A M-H step is then performed over a normal likelihood: if accepted, then $\phi^t = \phi^c$, or $\phi^t = \phi^{t-1}$ otherwise. In the fifth step we are again dealing with normal prior and normal likelihood, so α is easily sampled from a multivariate normal. The sixth and final step deals with the sampling of $\sigma_u^{2(c)}, \sigma_{\xi u}^c$ from a truncated bivariate normal over an admissible region $\mathcal{O}$ such that $\mathbf{\Sigma}^c$ is positive definite.

[17] Exogenous variables in $\mathbf{Z}$ have been dropped to simplify the exposition, since the conditioning always applies.

[18] For a brief introduction to Monte Carlo methods in Bayesian econometrics, the reader can refer to Chib (2001) or Robert and Casella (2004).

[19] The algorithm is robust to different starting values, however speed of convergence may increase significantly.

Acceptation or rejection is determined by the usual M-H decision rule. A detailed step-by-step description of the algorithm can be found in Appendix C.

4.4 Empirical results

Table 2 reports the estimation results. Columns (1) to (4) report the results when network exogeneity is assumed, with an increasing set of controls. In particular, columns (1) and (3) estimate equation (16) alone using OLS while columns (2) and (4) estimate equation (17) alone using MLE. Finally, columns (5) and (6) display the Bayesian estimation results where we estimate simultaneously equations (18) and (20). They account for a possible network endogeneity.[20]

Columns (1) and (2) use a basic set of demographic characteristics (age, race, sex, parental education and family income), while Columns (3) and (4) extends the number of control variables to include other specific characteristics possibly related to the intensity of social interaction and social capital, such as family structure (household size and a dummy if both parents are living in the same home), school performance, religion practice and physical development.

Table 2 shows that the Bayesian estimates are remarkably similar to the ones that are obtained using traditional estimation methods, with coefficients that are only slightly lower. This indicates that conditional on the observed characteristics, there are no troubling unobserved factors, i.e. unobserved variables that are correlated with both social capital (network) formation and intensity of social interactions between the formed social contacts. Indeed, the estimate of $\sigma_{u_j \xi_j}$ in model (16)–(17) is not statistically significant. It appears that , for example, being of a different gender or different race would decrease the level of social interactions, whereas the same level of parental education or same family size would increase social contacts (results in columns 1, 3 and 5). Turning our attention to our target parameters, the evidence shows that β_1, β_2 and ϕ are statistically significant and in line with the theoretical predictions: $\beta_1 < 0$, $\beta_2 > 0$ and $\phi > 0$. $\beta_1 < 0$ means that, the higher is the geographical distance between two students, the lower are their social interactions (measured as the time spent together in terms of telephone calls or going to the friend's house, etc.). In terms of magnitude, however, the reduction of social interactions with distance is not extremely large. Our estimates shows that in order to mimic the effect of

[20]We show the results that are obtained when using networks with size between 4 and 50 individuals to ease the Bayesian computational burden. When slightly changing the network size window, the results remain qualitatively unchanged.

gender differences on interactions, two students need to be located almost 80 kilometers apart from each other. $\beta_2 > 0$ means that the more central (in terms of eigenvector centrality) is a student, the more intense are his/her social interactions with other students. Furthermore, $\phi > 0$ indicates that more central students tend to interact more with more central students. Finally, most of the β_{0m}s are negative, which is in line with homophily behavior so that the closer students are in terms of observable characteristics (such as gender, race, grade, etc.), the higher are their social interactions.

[*Insert Table 2 here*]

For completeness, in the upper panels of Figures 1, 2 and 3, we show kernel density estimates of the posterior distributions of our target parameters: β_1, β_2 and ϕ. The lower panels display a time-series sample of draws, which shows that stationarity of the chain has been achieved.

[*Insert Figures 1*, 2 and 3 *here*]

5 Concluding remarks

Our theoretical model proposes a behavioral microfoundation for the relationship between geographical distance and social interactions. We characterize the equilibrium in terms of optimal level of social interactions and social capital for a general distribution of individuals in the geographical space. An important prediction of the model is that the level of social interactions is inversely related to the geographical distance. Travel costs and spatial dispersion of agents are barriers to the development of social capital formation. Social capital tends to be more concentrated than agents themselves. We also show that a spread of the spatial distribution of agents in the city increases the trip distances and their costs, which decrease the incentives to interact socially. As a result, a larger spatial dispersion of agents reduces social capital in cities. Because of the externalities that agents exert on each other, we demonstrate that the equilibrium levels of social interactions and social capital are lower than the efficient ones. In other words, there are two few social interactions both in quantity and quality compared to the social optimum. Our empirical analysis shows that, indeed, geographical distance is an hinder to social interactions, although similarity in gender or race are greater determinants.

References

[1] Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Publishers.

[2] Argazi, M. and V. Henderson (2008), "Networking off Madison avenue," *Review of Economic Studies* 75, 1011-1038.

[3] Bayer, P., Ross, S.L. and G. Topa (2008), "Place of work and place of residence: Informal hiring networks and labor market outcomes," *Journal of Political Economy* 116 ,1150-1196.

[4] Beckmann, M.J. (1976), "Spatial equilibrium in the dispersed city," In: G.J. Papageorgiou (Ed.), *Mathematical Land Use Theory*, Lexington Books, pp. 117-125.

[5] Behrens, K., Duranton, G. and F. Robert-Nicoud (2014), "Productive cities: Sorting, selection, and agglomeration," *Journal of Political Economy* 122, 507-553.

[6] Berliant, M., Peng, S.-K. and P. Wang (2002), "Production externalities and urban configuration," *Journal of Economic Theory* 104, 275–303.

[7] Blume, L.E., Brock, W.A., Durlauf, S.N. and Y.M. Ioannides (2011), "Identification of social interactions," In: J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), *Handbook of Social Economics*, Vol. 1B, Amsterdam: Elsevier Publisher, pp. 853-964.

[8] Bramoullé, Y., Djebbari, H. and B. Fortin (2009), "Identification of peer effects through social networks," *Journal of Econometrics* 150, 41-55.

[9] Brueckner, J.K. and A.G. Largey (2008), "Social interactions and urban sprawl," *Journal of Urban Economics* 64, 18-34.

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

[11] Chib, S. (2001), "Markov chain Monte Carlo methods: Computation and inference," In: J.J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics, Vol. 5*, Amsterdam: Elsevier Science, pp. 3569-3649.

[12] Chib, S. and E. Greenberg (1995), "Understanding the Metropolis-Hastings algorithm," *American Statistician* 49, 327-335.

[13] Chib, S. and E. Greenberg (1996), "Markov chain Monte Carlo simulation methods in econometrics," *Econometric Theory* 12, 409-431.

[14] Couture, V., Duranton, G. and M.A. Turner (2013), "Speed," Unpublished manuscript, University of Toronto.

[15] Currarini, S., Jackson, M.O., and P. Pin (2009), "An economic model of friendship: Homophily, minorities, and segregation," *Econometrica* 77, 1003-1045.

[16] Duranton, G. and D. Puga (2015), "Urban land use," In: G. Duranton, V. Henderson and W. Strange (Eds.), *Handbook of Regional and Urban Economics, Vol. 5*, Amsterdam: Elsevier Publisher, forthcoming.

[17] Duranton, G. and M.A. Turner (2011), "The fundamental law of road congestion: Evidence from US cities," *American Economic Review* 101, 2616-2652.

[18] Duranton, G. and M.A. Turner (2012), "Urban growth and transportation," *Review of Economic Studies* 79, 1407-1440.

[19] Fujita, M. (1989), *Urban Economic Theory*, Cambridge: Cambridge University Press.

[20] Fujita, M. and J.-F. Thisse (2013), *Economics of Agglomeration. Cities, Industrial Location, and Regional Growth*, Second Edition, Cambridge University Press, Cambridge.

[21] Fujita, M. and H. Ogawa (1980), "Equilibrium land use patterns in a nonmonocentric city," *Journal of Regional Science* 20, 455-475.

[22] Glaeser, E.L. (1999), "Learning in cities," *Journal of Urban Economics* 46, 254-277.

[23] Glaeser, E.L. (2000), "The future of urban economics: Non-market interactions," *Brookings-Wharton Papers on Urban Affairs* 1, 101-150.

[24] Goldenberg, J. and M. Levy (2009), "Distance Is not dead: Social interaction and geographical distance in the Internet era," *Computers and Society* 2, 1-22.

[25] Goldsmith-Pinkham, P. and G. Imbens (2013), "Social networks and the identification of peer effects," *Journal of Business & Economic Statistics* 31, 253-264.

[26] Graham, B.S. (2014), "An empirical model of network formation: Detecting when agents are heterogenous," Unpublished manuscript, U.C. Berkeley.

[27] Graham, B.S. (2015), "Methods of identification in social networks," *Annual Review of Economics*, forthcoming.

[28] Hellerstein, J.K., M.P. McInerney and D. Neumark (2011), "Neighbors and co-workers: The importance of residential labor market networks," *Journal of Labor Economics* 29, 659-695.

[29] Hellerstein, J.K., Kutzbach, M.J. and D. Neumark (2014), "Do labor market networks have an important spatial dimension?" *Journal of Urban Economics* 79, 39-58.

[30] Helsley, R.W. and W.C. Strange (2004), "Knowledge barter in cities," *Journal of Urban Economics* 56, 327-345.

[31] Helsley, R.W. and W.C. Strange (2007), "Urban interactions and spatial structure," *Journal of Economic Geography* 7, 119-138.

[32] Helsley, R.W. and W.C. Strange (2014), "Coagglomeration, clusters, and the scale and composition of cities," *Journal of Political Economy* 122, 1064-1093.

[33] Helsley, R.W. and Y. Zenou (2014), "Social networks and interactions in cities," *Journal of Economic Theory* 150, 426-466.

[34] Hsieh, C.S. and L.F. Lee (2013), "A social interactions model with endogenous friendship formation and selectivity," Unpublished manuscript, Ohio State University.

[35] Ioannides, Y.M. (2012), *From Neighborhoods to Nations: The Economics of Social Interactions*, Princeton: Princeton University Press.

[36] Jackson, M.O. (2008), *Social and Economic Networks*, Princeton University Press, Princeton.

[37] Jackson, M.O. (2014), "Networks in the understanding of economic behaviors," *Journal of Economic Perspectives* 28, 3-22.

[38] Jackson, M.O., Rogers, B.W. and Y. Zenou (2015), "The economic consequences of social network structure", CEPR Discussion Paper 10406.

[39] Jackson, M.O. and Y. Zenou (2015), "Games on networks," in: P. Young, S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier Science, pp. 91-157.

[40] Jacobs, J. (1969), *The Economy of Cities*, New York: Vintage.

[41] Johnson, C. and R.P. Gilles (2000), "Spatial social networks," *Review of Economic Design* 5, 273-299.

[42] Kaltenbrunner, A., Scellato, S., Volkovich, Y., Laniado, D., Currie, D., Jutemar, E.J. and C. Mascolo (2012), "Far from the eyes, close on the web: Impact of geographic distance on online social interactions," Proceedings of the 2012 ACM workshop on Workshop on online social networks, pp. 19-24.

[43] Lee, L.F., Liu, X. and X. Lin (2010), "Specification and estimation of social interaction models with network structures," *Econometrics Journal* 13, 145-176.

[44] Lin, X. (2010), "Identifying peer effects in student academic achievement by a spatial autoregressive model with group unobservables," *Journal of Labor Economics* 28, 825-860.

[45] Liu, X., Patacchini, E., Zenou, Y. and L-F. Lee (2012), "Criminal networks: Who is the key player?" CEPR Discussion Paper No. 8772.

[46] Lucas, R.E. (1988), "On the mechanics of economic development," *Journal of Monetary Economics* 22, 3-42.

[47] Lucas, R.E. and E. Rossi-Hansberg (2002), "On the internal structure of cities," *Econometrica* 70, 1445-1476.

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

[49] Overman, H.G., Gibbons, S. and E. Patacchini (2015), "Spatial methods," In: G. Duranton, V. Henderson and W. Strange (Eds.), *Handbook of Regional and Urban Economics, Vol. 5*, Amsterdam: Elsevier Publisher, forthcoming.

[50] Marmaros, D. and B.I. Sacerdote (2006), “How do friendships form?” *Quarterly Journal of Economics* 121, 79-119.

[51] Mele, A. (2013), “A structural model of segregation in social networks,” Unpublished manuscript, Johns Hopkins Carey Business School.

[52] Mossay P. and P.M. Picard (2011), “On spatial equilibria in a social interaction model,” *Journal of Economic Theory* 146, 2455-2477.

[53] Mossay P. and P.M. Picard (2013), “Spatial segregation and urban structure,” CORE Discussion Paper No. 2013/19.

[54] Putnam, R.D. (2000), *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon & Schuster.

[55] Robert, C.P. and G. Casella (2004), *Monte Carlo Statistical Methods*, New York: Springer-Verlag.

[56] Romer, P.M. (1986), “Increasing returns and long-run growth,” *Journal of Political Economy* 94, 1002-1037.

[57] Rosenthal, S.S. and W. C. Strange (2008), “The attenuation of human capital externalities,” *Journal of Urban Economics* 64, 373-389.

[58] Sato, Y. and Y. Zenou (2015), “How urbanization affect employment and social interactions,” *European Economic Review* 75, 131-155.

[59] Schmutte, I.M. (2015), “Job referral networks and the determination of earnings in local labor markets,” *Journal of Labor Economics* 33, 1-32.

[60] Sheng S. (2015), “A structural econometric analysis of network formation games,” Unpublished manuscript, University of California at Los Angeles.

[61] Topa, G. and Y. Zenou (2015), “Neighborhood and network effects,” In: G. Duranton, V. Henderson and W. Strange (Eds.), *Handbook of Regional and Urban Economics, Vol. 5*, Amsterdam: Elsevier Publisher, forthcoming.

[62] Wasserman, S. and K. Faust (1994), *Social Network Analysis. Methods and Applications*, Cambridge: Cambridge University Press.

[63] Zenou, Y. (2009), *Urban Labor Economics*, Cambridge: Cambridge University Press.

[64] Zenou, Y. (2013), "Spatial versus social mismatch," *Journal of Urban Economics* 74, 113-132.

[65] Zenou, Y. (2015), "Networks in economics," In: J.D. Wright (Ed.), *International Encyclopedia of Social and Behavioral Sciences*, 2nd Edition, Oxford: Elsevier Publisher.

APPENDIX

A Proof of Proposition 3

We need to show that (i) $\mathrm{Disp}(s\lambda) < \mathrm{Disp}(\lambda)$ is equivalent to $\mathrm{Disp}(g\lambda) < \mathrm{Disp}(\lambda)$ and (ii) this is true when $x^2\lambda(x)/\int z^2\lambda(z)\mathrm{d}z$ is a mean preserving spread of a symmetric distribution of $\lambda(x)$.

First, note that s_0 is a constant and s and λ are functions of z. One successively gets the following equivalences:

$$
\begin{aligned}
&\mathrm{Disp}(s\lambda) < \mathrm{Disp}(\lambda)\\
&\quad\Leftrightarrow \frac{\int z^2 s\lambda \mathrm{d}z}{\int s\lambda \mathrm{d}z} < \frac{\int z^2\lambda \mathrm{d}z}{\int \lambda \mathrm{d}z}\\
&\quad\Leftrightarrow \frac{\int z^2\left(s_0-\alpha g\right)\lambda \mathrm{d}z}{\int z^2\lambda \mathrm{d}z} < \frac{\int \left(s_0-\alpha g\right)\lambda \mathrm{d}z}{\int \lambda \mathrm{d}z}\\
&\quad\Leftrightarrow s_0-\alpha\frac{\int z^2 g\lambda \mathrm{d}z}{\int z^2\lambda \mathrm{d}z} < s_0-\alpha\frac{\int g\lambda \mathrm{d}z}{\int \lambda \mathrm{d}z}\\
&\quad\Leftrightarrow \frac{\int z^2 g\lambda \mathrm{d}z}{\int z^2\lambda \mathrm{d}z} > \frac{\int g\lambda \mathrm{d}z}{\int \lambda \mathrm{d}z}\\
&\quad\Leftrightarrow \frac{\int z^2 g\lambda \mathrm{d}z}{\int g\lambda \mathrm{d}z} > \frac{\int z^2\lambda \mathrm{d}z}{\int \lambda \mathrm{d}z}\\
&\quad\Leftrightarrow \mathrm{Disp}(g\lambda) > \mathrm{Disp}(\lambda)
\end{aligned}
$$

where, for notation convenience, we have dropped the integrals the boundaries $-b$ and b.

Second, by denoting by

$$\mu(z) \equiv \frac{z^2\lambda(z)}{\int_{-b}^{b} w^2\lambda(w)\mathrm{d}w},$$

we can write the last condition $\mathrm{Disp}(g\lambda) > \mathrm{Disp}(\lambda)$ as $\int_{-b}^{b} g\mu \mathrm{d}z - \int_{-b}^{b} g\lambda \mathrm{d}z = \int_{-b}^{b} g\left(\mu-\lambda\right)\mathrm{d}z > 0$. Integrating by part, we obtain the following condition:

$$-\int_{-b}^{b}\left\{\int_{-b}^{z}\left[\mu(x)-\lambda(x)\right]\mathrm{d}x\right\} g'(z)\mathrm{d}z > 0 \tag{22}$$

Finally, consider the symmetric spatial distribution $\lambda(x)$ around $x=0$. Because $\lambda(x)$ is symmetric around $x=0$, then $g(x)=\int_{-b}^{b} c(x-z)\lambda(z)\mathrm{d}z$ is symmetric around $x=0$. Furthermore we know that $g(x)$ is also convex, which implies that $g'(z)>0$ if and only if $z>0$. A sufficient condition for inequality (22) to be true is that $\int_{-b}^{z}[\mu(x)-\lambda(x)]\,\mathrm{d}x$ is negative for $z>0$ and positive for $z<0$. That is, if

$$\int_{-b}^{z}\mu(x)\mathrm{d}x \leq \int_{-b}^{z}\lambda(x)\mathrm{d}x, \text{ for } z>0$$

and the opposite condition for $z<0$. This condition is satisfied if $\mu(x)$ is a mean preserving spread of the distribution of $\lambda(x)$ around its mean $x=0$. For example, for a uniform distribution $\lambda(x)=1/(2b)$, we get

$$\begin{aligned}\int_{-b}^{z}\mu(x)\mathrm{d}x-\int_{-b}^{z}\lambda(x)\mathrm{d}x &= \int_{-b}^{z}\left(\frac{x^2}{\int_{-b}^{b}w^2\frac{1}{2b}\mathrm{d}w}-1\right)\frac{1}{2b}\mathrm{d}x \\ &= -\frac{1}{2}z\left(b^2-z^2\right)/b^3<0\end{aligned}$$

so that $\mu(x)$ is a mean preserving spread of the distribution of $\lambda(x)$. ▮

B Efficient social interactions

We now study the planner's allocation of interaction frequency for a given location pattern λ. The planner chooses the profiles of social interactions $n(\cdot,\cdot)$ and social capital $s(\cdot)$ that maximize the aggregate utility

$$W=\int_{-b}^{b}U(x)\lambda(x)\mathrm{d}x=\int_{-b}^{b}\left[S(x)-C(x)\right]\lambda(x)\mathrm{d}x$$

subject to the social capital constraint

$$s(x)\leq 1+\alpha\int_{-b}^{b}n(x,z)s(z)\lambda(z)\mathrm{d}z \tag{23}$$

where we put an inequality to express that the agent can always reduce her social capital at no cost (e.g. she erases a part of her address book).

The government chooses the profiles $n(\cdot,\cdot)$ and $s(\cdot)$ that maximize the Lagrangian function

$$\begin{aligned}\mathcal{L}=&\int_{-b}^{b}\int_{-b}^{b}\left\{v\left[n(x,y)\right]s(y)-n(x,y)c(x-y)\right\}\lambda(y)\lambda(x)\mathrm{d}x\mathrm{d}y \\ &-\int_{-b}^{b}\upsilon(x)\left[s(x)-1-\alpha\int_{-b}^{b}n(x,y)s(y)\lambda(y)\mathrm{d}y\right]\lambda(x)\mathrm{d}x\end{aligned}$$

where $\upsilon(x)\geq 0$ (or more precisely $\upsilon(x)\lambda(x)$) is the Kuhn-Tucker multiplier of the social capital constraint. So, $\upsilon(x)$ measures the welfare value of a marginal increase of the social capital of an agent located at x.

Lemma 4 *The efficient interaction frequency and social capital satisfy the following necessary conditions:*

$$v'\left[n(x,y)\right]s(y) - c(x-y) + \alpha\upsilon(x)s(y) = 0 \tag{24}$$

$$\int_{-b}^{b}\left\{v\left[n(x,y)\right] + \alpha\upsilon(x)n(x,y)\right\}\lambda(x)\mathrm{d}x - \upsilon(y) = 0 \tag{25}$$

Equations (24), (25) together with binding (23) solve for the functions $n(x,y)$, $s(y)$ and $\upsilon(x)$.

Proof. Substituting z for y we can write the Lagrangian function as

$$\mathcal{L} = \int_{-b}^{b}\int_{-b}^{b}\left\{v\left[n(x,y)\right]s(y) - n(x,y)c(x-y) + \alpha\upsilon(x)n(x,y)s(y)\right\}\lambda(y)\lambda(x)\mathrm{d}x\mathrm{d}y$$
$$-\int_{-b}^{b}\upsilon(x)\left[s(x)-1\right]\lambda(x)\mathrm{d}x$$

Finally, we note that $\int_{-b}^{b}\upsilon(x)\left[s(x)-1\right]\lambda(x)\mathrm{d}x = \int_{-b}^{b}\upsilon(y)\left[s(y)-1\right]\lambda(y)\mathrm{d}y$. Substituting the latter expression in the last term in the above expression and multiplying it by $\int_{-b}^{b}\lambda(x)\mathrm{d}x$ $(=1)$ we get the following Lagrangian function:

$$\mathcal{L} = \int_{-b}^{b}\int_{-b}^{b}\left\{\begin{array}{c} v\left[n(x,y)\right]s(y) - n(x,y)c(x-y) \\ +\alpha\upsilon(x)n(x,y)s(y) - \upsilon(y)\left[s(y)-1\right]\end{array}\right\}\lambda(y)\lambda(x)\mathrm{d}x\mathrm{d}y \tag{26}$$

We now use variation calculus on the Lagragian function $\mathcal{L} = \int_{-b}^{b}\int_{-b}^{b}F[n(x,y),s(y),x,y]\lambda(y)\lambda(x)\mathrm{d}x\mathrm{d}y$ where $F\left(n,s,x,y\right)$ denotes the integrand in the above curly bracket. It is a differentiable function with partial derivatives F_n' and F_s'. Defining the infinitely small perturbations $\widetilde{n}(x,y)$ and $\widetilde{s}(y)$ on the optimal profiles $n^o(x,y)$ and $s^o(y)$ respectively, we get the variation of the objective function $\mathcal{L}$

$$\Delta\mathcal{L} = \int_{-b}^{b}\int_{-b}^{b}F_n'[n^o(x,y),s^o(y),x,y]\widetilde{n}(x,y)\lambda(y)\lambda(x)\mathrm{d}x\mathrm{d}y$$
$$+\int_{-b}^{b}\int_{-b}^{b}F_s'[n^o(x,y),s^o(y),x,y]\widetilde{s}(y)\lambda(y)\lambda(x)\mathrm{d}x\mathrm{d}y$$

This must be zero for any small perturbations $\widetilde{n}(x,y)$ and $\widetilde{s}(y)$. So, we get $F_n'[n^o(x,y),s^o(y),x,y] = 0$ and $\int_{-b}^{b}\left\{F_s'[n^o(x,y),s^o(y),x,y]\right\}\lambda(x)\mathrm{d}x = 0$. This gives (24) and (25). ▮ ■

Condition (24) captures the main externality at work in the process of social interaction. When the planner chooses the interaction frequency $n(x,y)$, he considers the benefit and cost to agent x

but also the welfare benefit $\upsilon(x)$ of raising the social capital $s(y)$ of the agents that she visits at y. This latter effect is not considered by agent x at the equilibrium. The weight that the planner puts on raising agent y's social capital increases with the importance of interactions, α, and with the social benefit of relaxing the social capital constraint, $\upsilon(x)$.

The second condition (25) is interpreted as follows: when the planner increases the social capital of an agent located at y, he directly raises the utility of all agents who interact with this agent (first term in curly brackets) and indirectly increases the social capital for all those other agents (second term in the curly brackets). In the efficient allocation, this combined effect should be equal to $\upsilon(y)$, the welfare value of a marginal increase of the social capital of an agent located at y.

Condition (24) yields

$$v'\left[n(x,y)\right] = \frac{c(x-y)}{s(y)} - \alpha\upsilon(x) \tag{27}$$

which gives

$$n^o(x,y) = 1 - \frac{c(x-y)}{s^o(y)} + \alpha\upsilon^o(x)$$

under our specification of v. With social capital at y held fixed at the equilibrium level ($s^*(y) = s^o(y)$), this expression is larger than the equilibrium number of visits $n^*(x,y)$ because $\upsilon^o(x) \geq 0$. The question thus becomes how social capital changes in this efficient allocation.

Inserting (9) in the binding condition (23), we get

$$s^o(x) = 1 + \alpha \int_{-b}^{b} s^o(z)\lambda(z)\mathrm{d}z - \alpha g(x) + \alpha^2\upsilon^o(x) \int_{-b}^{b} s^o(z)\lambda(z)\mathrm{d}z$$

We use the same algebraic manipulation leading to expression (7), multiplying both members of the last expression by $\lambda(x)$, integrating them and simplifying to get the value of $\int_{-b}^{b} s^o(x)\lambda(x)\mathrm{d}x$. We then insert this expression in the previous equality and simplify, getting the following closed-form solution for the efficient level of social capital:

$$s^o(x) = 1 + \alpha \frac{\left[1 + \alpha\upsilon^o(x)\right]\left[1 - \alpha \int_{-b}^{b} g(z)\lambda(z)\mathrm{d}z\right]}{1 - \alpha - \alpha^2 \int_{-b}^{b} \upsilon^o(z)\lambda(z)\mathrm{d}z} - \alpha g(x)$$

If $\upsilon^o(x) = 0$, this yields the equilibrium $s^*(x)$. However, since $\upsilon^o(x) \geq 0$, the numerator is larger and the denominator is smaller than in the equilibrium. It thus must be that $s^o(x) > s^*(x)$. In turn, this implies that $n^o(x,y) \geq n^*(x,y)$.

Proposition 5 *The equilibrium frequency of interactions and level of social capital are smaller than the efficient ones.*

Intuitively, the planner internalizes the effect that each agent has on others' social capital when she entertains more intense social interactions. As a result, the planner asks agents to increase their frequency of social interactions above the equilibrium level. This welfare conclusion confirms Brueckner's (2007) and extends his analysis to the case where agents are distributed across the space.

Can the efficient allocation of social interactions can be decentralized by setting subsidies $\eta(x,y)$ and $\theta(x,y)$ for interactions and travel cost? In this case, the utility becomes

$$\begin{aligned} U(x) &= S(x) - C(x) \\ &= \int_{-b}^{b} \{v\,[n(x,y)]\,[s(y) + \eta(x,y)] - n(x,y)\,[c(x-y) - \theta(x,y)]\}\lambda(y)\mathrm{d}y \end{aligned}$$

while the equilibrium number of interactions becomes

$$n^*(x,y) = 1 - \frac{c(x-y) - \theta(x,y)}{s(y) + \eta(x,y)}$$

Efficient social interactions can therefore be decentralized by setting $\eta(x,y) = 0$ and $\theta(x,y) = \alpha v^o(x) s^o(y)$. Indeed, in this case, we find again $n^*(x,y) = 1 - c(x-y)/s(y) + \alpha v^o(x)$. This means first that social interactions should not be subsidized, and second that trips should subsidized as a function of the locations of the destination and origin partners. The subsidy should be higher for trips to partners who have higher social capital and for trips from partners whose social capital increases more with additional interactions. The optimal subsidy to travel costs is therefore not a uniform one. This suggests that decentralization would be difficult to implement because subsidies depend on both the origins and destinations of social interactions (it is very unlikely that $\theta(x,y)$ reduces to a simple function of x, or y or $x-y$). This result contrasts with Helsley and Zenou (2014), who advocate that the planner should subsidize the most central agents. Their model with a two location points, however, imperfectly captures the full picture of spatial interactions. In the present model, we observe that the planner does not subsidize those agents with high social capital but only subsidizes the trips to those agents.

C Bayesian estimation: technical details

Given the parameter vector $\mathbf{\Theta} = (\beta, \alpha, \phi, \mathbf{\Sigma}, \sigma^2_\epsilon)$, an $n \times n$ matrix of intensities of interaction $\mathbf{N} = \{n_{ij}\}$, a vector of social capital $\mathbf{s}$ and a matrix of observables $\mathbf{Z}$, we can specify our bayesian probabilistic model as

$$P(\xi, \mathbf{\Theta}|\mathbf{N}, \mathbf{s}) \propto \pi(\xi, \mathbf{\Theta}) \times \mathcal{L}(\mathbf{N}, \mathbf{s}|\xi, \Theta),$$

where $\mathcal{L}(\mathbf{N}, \mathbf{s}|\xi, \mathbf{\Theta})$ is the joint likelihood function and $\pi(\xi, \theta)$ the joint prior distribution[21]. The usual independency assumption allows us to write $\pi(\xi, \mathbf{\Theta}) = \pi_1(\xi) \times \pi_2(\mathbf{\Theta})$ and $\mathcal{L}(\mathbf{N}, \mathbf{s}|\xi, \mathbf{\Theta}) = \mathcal{L}(\mathbf{N}|\xi, \mathbf{\Theta}) \times \mathcal{L}(\mathbf{s}|\mathbf{N}, \xi, \mathbf{\Theta})$.

Bayesian inference requires the computation of marginal distribution for all parameters. However, since this requires integration of complicated distributions over several variables, a closed form solution is not readily available and MCMC techniques are usually employed to obtain random draws from posterior distributions.

By using Bayes' theorem we can write conditional marginal densities in order to set up our Gibbs sampling scheme:

i) $P(\xi|\mathbf{N}, \mathbf{s}, \mathbf{\Theta}) \propto \pi(\xi) \times \mathcal{L}(\mathbf{N}, \mathbf{s}|\xi, \mathbf{\Theta})$.

ii) $P(\beta|\mathbf{N}, \xi) \propto \pi(\beta) \times \mathcal{L}(\mathbf{N}|\xi, \beta)$

iii) $P(\phi|\mathbf{s}, \mathbf{N}, \xi, \alpha, \mathbf{\Sigma}) \propto \mathcal{L}(\mathbf{s}|\mathbf{N}, \xi, \alpha, \mathbf{\Sigma})$, with ϕ uniform in the interval $(-1/\tau, 1/\tau)$, and τ equal to the largest eigenvalue of $\mathbf{N}$. This is required for $(\mathbf{I} - \phi\mathbf{N})$ to be invertible according to Gershgorin's theorem.

iv) $P(\alpha|\mathbf{N}, \mathbf{s}, \phi, \mathbf{\Sigma}) \propto \pi(\alpha) \times \mathcal{L}(\mathbf{s}|\mathbf{N}; \xi, \alpha, \phi, \mathbf{\Sigma})$

v) $P(\Sigma|\mathbf{N}, \mathbf{s}, \xi, \alpha, \phi, \mathbf{\Sigma}) \propto \pi(\mathbf{\Sigma}) \times \mathcal{L}(\mathbf{s}|\mathbf{N}, \xi, \alpha, \phi)$.

C.1 MCMC algorithms: a review

Modern bayesian estimation relies on Monte Carlo methods to draw random samples from known distribution. The reader can refer to Chib and Greenberg (1995, 1996) or Robert and Casella (2004). Let us suppose we have a distribution $q^*(x)$ of which we don't know its normalizing

[21] Exogenous variables in X have been dropped to simplify the exposition.

constant or its moments are very difficult or impossible to compute. The goal of MCMC methods is to draw samples from such target distribution in order to explore its characteristics empirically instead of analytically: sample moments and sample quantiles can be used instead of theoretical ones when the simulated distribution converges to the target one.

Gibbs sampling While Gibbs sampling can be shown to be a special case of the M-H algorithm, it's quicker and simpler although it needs the full conditional distribution to be known. Clearly this makes the Gibbs sampler alone less applicable. Let us suppose we need to draw a sample from $q(x_1, x_2, x_3)$. In the bayesian case this could be the full conditional distribution (21). The algorithm works as follows.

1. Chose x_1^0, x_2^0, x_3^0 as starting values of the chain and set an index $t = 0$
2. Draw x_1^{t+1} from $q(x_1|x_2^t, x_3^t)$
3. Draw x_2^{t+1} from $q(x_2|x_1^{t+1}, x_3^t)$
4. Draw x_3^{t+1} from $q(x_3|x_1^{t+1}, x_2^{t+1})$
5. Set $t = t + 1$ and get back to step 2.

When the chain $\{x_1\}$ converges to its stationary distribution $q(\cdot)$, its elements can be considered as a sample from the marginal conditional densities $q(x_i|x_{-i})$.

Metropolis-Hastings within Gibbs If any of the conditional distribution needed for the Gibbs sampling are not suitable for direct sampling (in our example it could be $q(x_3|x_1, x_2)$), Metropolis-Hastings (or Metropolis) can be employed in that step.

Metropolis-Hastings Let us suppose we cannot draw a random observation from $q(x_3|x_1, x_2)$ in the example above. Given a starting draw x_3^t, the M-H algorithm uses a *candidate point* y from a suitable *proposal distribution* $f(y|x_3^t)$. The candidate y is accepted as the 'new' x_3^{t+1} with probability given by

$$\alpha(y, x_3^t) = \min\left[\frac{f(y|x_3^t)\ q(x_3^t|x_1^t, x_2^t)}{f(x_3^t|y)\ q(y|x_1^t, x_2^t)}, 1\right].$$

Note that since the target distribution is on both the numerator and the denominator, normalizing constant does not matter at all for the sampling.

C.2 Choice of priors and estimation

Let $\tilde{\beta}$ be the OLS estimate for β. Our chosen priors are the following:

- $\pi(\xi) \sim \mathcal{N}(0, I_{nn})$
- $\pi(\beta) \sim \mathcal{N}(\tilde{\beta}, \tilde{S}_\beta)$
- $\pi(\phi) \sim \mathcal{U}(-1/\tau, 1/\tau)$, with τ equal to the largest eigenvalue of N
- $\pi(\alpha) \sim \mathcal{N}(\tilde{\alpha}, \tilde{S}_\alpha)$
- $\pi(\Sigma) \sim \mathcal{TN}\left(\begin{pmatrix} \sigma^2_{\epsilon 0} \\ \sigma^2_{\xi\epsilon 0} \end{pmatrix}, \begin{bmatrix} \sigma^2_v & \sigma_{v\xi} \\ \sigma_{v\xi} & \sigma^2_\xi \end{bmatrix}\right)$ as a truncated bivariate normal on the support $\mathcal{S}$ such that Σ is positive definite.
- For $\sigma^2_u \propto (1/\sigma)$ we choose an uninformative diffuse prior.

Given the state of the chain at $t-1$, the t^{th} iteration our M-H within Gibbs MCMC algorithm can be summarized in the following steps.

1. Sample ξ_i from $P(\xi|N, s, \Theta)$ by a M-H step for all $i = 1, \cdots, nn$:
   - i) Propose $\xi_i \sim \mathcal{N}(\xi_i^{t-1}, \kappa^t)$ where κ^t is a user-specified time-varying variance such that the acceptance rate is between 40% and 60%.
   - ii) Accept or reject ξ_i according to the M-H rule, using a standard normal as proposal distribution.
2. Sample ϕ from $P(\phi|s, N, \xi, \alpha, \Sigma)$

Table 1: Description of Data and summary statistics

	Variable definition	n.obs.	mean	std.dev	min	max
WAVE I						
Individual socio-demografic variables						
Female	Dummy variable taking value one if the respondent is female.	1537	0.53	0.50	0	1
Grade	Grade of student in the current year, range 7 to 12	1537	9.00	1.66	7	12
Black	Dummy variable taking value one if the respondent is Black or African American. “White” is the reference category.	1537	0.18	0.39	0	1
Other race	Dummy for other races	1537	0.07	0.08	0	1
Religion practice	Answer to the question “In the past 12 months, how often did you attend religious services?”. Coded as 1= “once a week or more”, 2=”once a month or more, but less than once a week”, 3=”once a month”, 4=”never”, 5=”has no religion”	1537	3.84	1.80	1	5
Grade Point Average	Grades defined from “A”=4 to “D and lower”=0. Average of grades in english, math, science and history is taken.	1537	2.90	0.73	0	4
Physical developement	Answer to the question “How advanced is your physical development compared to other boys your age?”. Coded as 1=”I look younger than most”, 2=”I look younger than some”, 3=”I look average”, 4=”I look older than some”, 5=”I look older than most”.	1537	3.23	1.12	1	5
Family background variables						
Family size	Number of people living in the household, range 2 to 11	1537	3.48	1.65	2	11
Married parents	Dummy variable taking value one if the respondent lives in a household with two parents (both biological and non biological) that are married.	1537	0.73	0.48	0	1
Parent education	Schooling level of the (biological or non-biological) parent who is living with the child, coded as 1 = "never went to school", 2 = “some school” and "less than high school", 3= “high school graduate”, “GED”, “went to a business, trade or vocational school”, “some college” 4 = "graduated from college or a university", 5 = "professional training beyond a four-year college". If both parents are in the household the maximum level of schooling is considered.	1537	3.18	2.63	9	19
Family income	Family income in thousands of dollars	1537	50.82	78.42	0	999
Family income refused dummy	Dummy that takes value 1 if parents refused to answer.	1537	0.12	0.32	0	1
Dyadic variables						
Spatial distance	Add-Health data reports X- and Y-coordinates in Kilometers from a central point translated from the Universal Transverse Mercator coordinates. By plotting the coordinates, the spatial relationship between interviewed respondents can be calculated as an euclidean distance.	20406	6.95	7.87	0	68.77
Intensity of interactions	The intensity of interaction measure is calculated as the sum of answers to questions regardin activities performed with another individual previosly identified as friends. Activities are: “going to friend’s house”, “hang out or go somewhere with the friend after school”, “spend time with the friend during the weekend”, “talk to the friend about a problem during the last week” and “talk to the phone during the last seven days”.	20406	0.17	0.76	0	5
Social capital	As proxy for social capital we use the eigenvector centrality (the normalized leading eigenvector of the directed adjacency matrix).	1537	0.16	0.33	-0.68	0.71

Table 2: Empirical results

	Interactions	**Social capital**	**Interactions**	**Social capital**	**Interactions**	**Social capital**
	(1)	(2)	(3)	(4)	(5)	(6)
Social capital (β_2)	0.1961***		0.1900***		0.1903***	
	(0.0246)		(0.0249)		(0.0250)	
Geographical distance (β_1)	-0.0018***		-0.0019***		-0.0012**	
	(0.0007)		(0.0007)		(0.0006)	
Peers' social capital (φ)		0.0830***		0.0840***		0.0762***
		(0.0001)		(0.0001)		(0.0006)
Constant	0.2601***	-0.0997	0.2479***	-0.0961	0.2459***	-0.0203
	(0.0119)	(0.0448)	(0.0143)	(0.0592)	(0.0152)	(0.0193)
Female	-0.0984***	0.0063	-0.0952***	-0.0198	-0.0952***	-0.0279
	(0.0105)	(0.0137)	(0.0167)	(0.0190)	(0.0167)	(0.0191)
Black african	-0.0809***	-0.0212	-0.0828***	-0.0275	-0.0829***	-0.0703
	(0.0181)	(0.018)	(0.0184)	(0.0187)	(0.0186)	(0.0841)
Other races	-0.098***	-0.135	-0.1121***	-0.0648	-0.1097***	0.0269***
	(0.0355)	(0.0822)	(0.0357)	(0.0815)	(0.0359)	(0.0043)
Student grade	-0.0664***	0.0267***	-0.0693***	0.0257***	-0.0697***	-0.0004
	(0.0048)	(0.0041)	(0.0049)	(0.0041)	(0.0049)	(0.0078)
Parental education	0.0162***	-0.0051	0.0161***	-0.0008	0.0164***	0.0000
	(0.0064)	(0.0074)	(0.0066)	(0.0077)	(0.0066)	(0.0001)
Family income	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0112
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0226)
Family income refused	-0.0398***	-0.0174	-0.0418***	-0.0114	-0.0422***	-0.0046
	(0.0132)	(0.0219)	(0.0134)	(0.0221)	(0.0133)	(0.0052)
Family size			0.0128***	-0.0042	0.0126***	0.0097
			(0.0043)	(0.0051)	(0.0043)	(0.0173)
Two parent			-0.0217**	0.0095	-0.0214***	0.0079
			(0.0115)	(0.0169)	(0.0116)	(0.0054)
Religion practice			-0.0024	0.0079	-0.0025	-0.0037
			(0.0048)	(0.0052)	(0.0049)	(0.0063)
Physical devel			0.0211***	-0.0037	0.0209***	-0.0006
			(0.0053)	(0.0061)	(0.0053)	(0.0103)
GPA			-0.0193**	-0.001	-0.0196*	-0.1039
			(0.0098)	(0.0100)	(0.0099)	(0.0606)
Unobservables (σ)					-0.0005	-0.0000
					(0.0046)	(0.0071)
Network fixed effects	yes	yes	yes	yes	yes	Yes
Estimation method	OLS	MLE	OLS	MLE	Bayesian	Bayesian
Obs	20406	1567	20406	1537	20406	1537
Number of networks	219	219	213	213	213	213

Notes: Columns (5) and (6) report the means and the standard deviations of the posterior distributions of the parameters. We draw random samples from each parameter's marginal conditional distribution using Markov Chain Monte Carlo (MCMC) techniques. We let our chain run for 22,000 iterations, discarding the first 2,000 iterations. Ergodicity of the Markov Chain is achieved quite fast (see Figures 1, 2, and 3). In columns (1), (3) and (5) the regressors are differences in terms of the listed variables between friends.

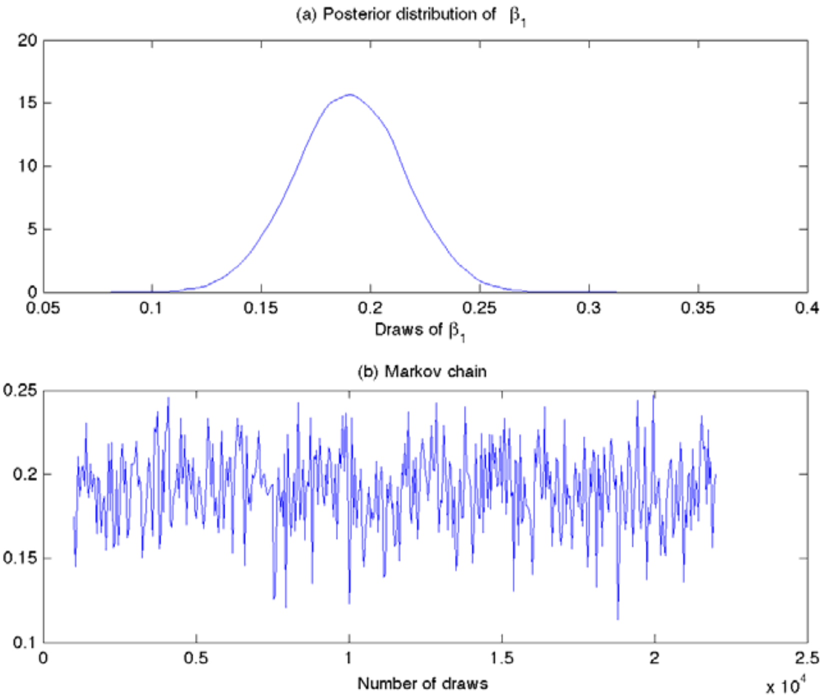

Figure 1: panel (a) shows the kernel density estimate of the posterior distribution of the parameter for geographic location. Panel (b) depicts the Markov chains draws of the parameter.

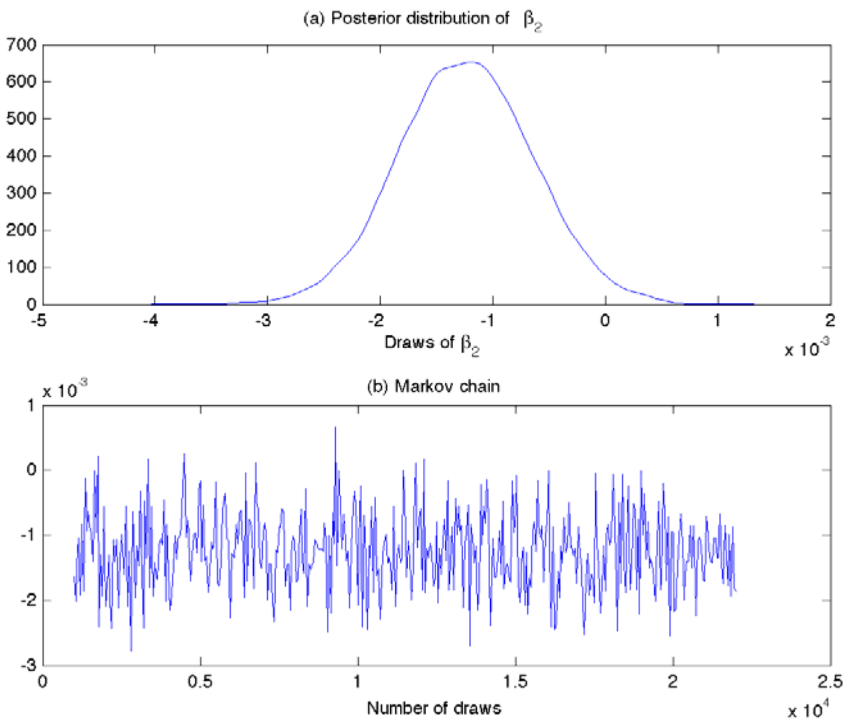

Figure 2: panel (a) shows the kernel density estimate of the posterior distribution of the parameter for social capital. Panel (b) depicts the Markov chains draws of the parameter.

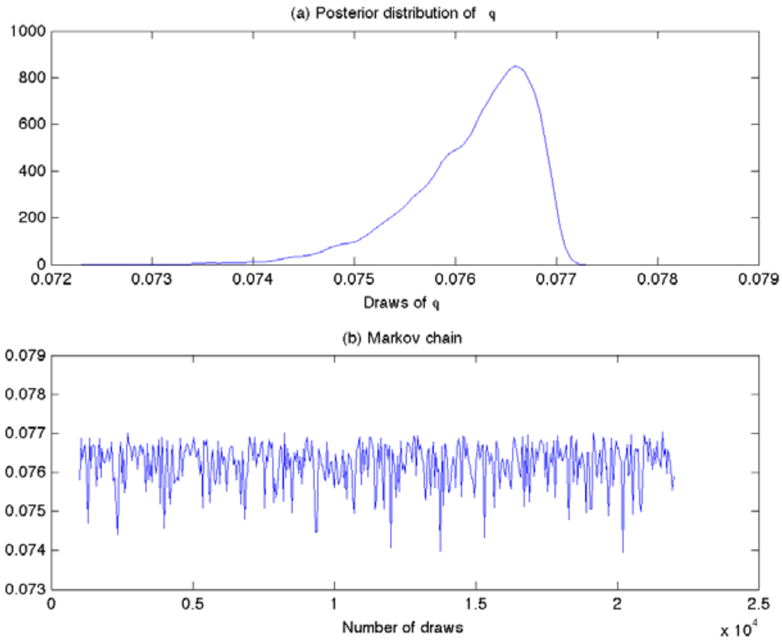

Figure 3: panel (a) shows the kernel density estimate of the posterior distribution of the parameter for social capital formation. Panel (b) depicts the Markov chains draws of the parameter.