Social Networks and Interactions in Cities*

Robert W. Helsley† Yves Zenou‡

January 8, 2013

Abstract

We examine how interaction choices depend on the interplay of social and physical distance, and show that agents who are more central in the social network, or are located closer to the geographic center of interaction, choose higher levels of interactions in equilibrium. As a result, the level of interactivity in the economy as a whole will rise with the density of links in the social network and with the degree to which agents are clustered in physical space. When agents can choose geographic locations, there is a tendency for those who are more central in the social network to locate closer to the interaction center, leading to a form of endogenous geographic separation based on social distance. We also show that the market equilibrium is not optimal because of social externalities. We determine the value of the subsidy to interactions that could support the first-best allocation as an equilibrium and show that interaction effort and the incentives for clustering are higher under the subsidy program. Finally, we interpret our model in terms of labor-market networks and show that the lack of good job contacts would be here a structural consequence of the social isolation of inner-city neighborhoods.

Keywords: Social networks, urban-land use, spatial mismatch, network centrality.

JEL Classification: D85, R14, Z13.

---

*We thank the participants of the 57th Annual North American Meetings of the Regional Science Association International, in particular, Jens Suedekum and Jacques Thisse, for helpful comments. We are also grateful to Antonio Cabrales, Jens Josephson and Mathias Thoenig for very insightful comments.

†Sauder School of Business, University of British Columbia, Vancouver, Canada. E-mail: robert.helsley@sauder.ubc.ca.

‡Stockholm University and Research Institute of Industrial Economics (IFN), Stockholm, Sweden, and GAINS. E-mail: yves.zenou@ne.su.se.

1 Introduction

Cities exist because proximity facilitates interactions between economic agents. There are few, if any, fundamental issues in urban economics that do not hinge in some way on reciprocal action or influence between or among workers and firms. Thus, the localization of industry arises from intra-industry knowledge spillovers in Marshall (1890), while the transmission of ideas through local inter-industry interaction fosters innovation in Jacobs (1969). In fact, the face-to-face interactions that Jacobs emphasizes are believed to be so critical to cities that Gaspar and Glaeser (1997) (and others) have asked whether advances in communication and information technology might make cities obsolete. As Glaeser and Scheinkman (2001, pp. 90) note: "Cities themselves are networks and the existence, growth, and decline of urban agglomerations depend to a large extent on these interactions."

The interactions that underlie the formation of urban areas are also important in other contexts. Following Romer (1986, 1990), Lucas (1988) views the local interactions that lead to knowledge spillovers as an important component of the process of endogenous economic growth. Non-market interactions also figure prominently in contemporary studies of urban crime (Glaeser et al., 1996; Verdier and Zenou, 2004), earnings and unemployment (Topa, 2001, Calvó-Armengol and Jackson, 2004; Moretti, 2004; Bayer et al., 2008; Zenou, 2009), peer effects in education (de Bartolome, 1990; Benabou, 1993; Epple and Romano, 1998), local human capital externalities and the persistence of inequality (Benabou, 1996, and Durlauf, 1996) and civic engagement and prosperity (Putnam, 1993).

While there is broad agreement that nonmarket interactions are essential to cities and important for economic performance more broadly, the mechanisms through which local interactions generate external effects are not well understood. The dominant paradigm lies in models of spatial interaction, which assume that knowledge, or some other source of increasing returns, arises as a by-product of the production marketable goods. 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. There is a rich literature (whose keystones include Beckmann, 1976; Fujita and Ogawa, 1980; and Lucas and Rossi-Hansberg, 2002) that examines how such spatial externalities influence the location of firms and households, urban density patterns, and productivity. There is also a substantial empirical literature (including Jaffee et al., 1993; Rosenthal and Strange, 2003, 2008; and Argazi and Henderson, 2008) demonstrating that knowledge spillovers do in fact attenuate with distance. Finally, there are more specific models that treat part of the interaction process as endogenous. For example, Glaeser (1999) examines 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.

This paper uses recent results from the theory of social networks to open the black box of local nonmarket interactions. We consider a population of agents who have positions within a social network and locations in a geographic space. As in Goyal (2007), Jackson (2008) and Jackson and Zenou (2013), we use the tools of graph theory to model the social network. In this model the value of interaction effort increases with the efforts of others with whom one has direct links in the social network. As in Helsley and Strange (2007) and Zenou (2013), all interactions take place at a point in space, the interaction center.

To be more precise, we consider a geographical model with two locations, the center, where all interactions occur, and the periphery. All agents are located in either the center or the periphery (geographical space). Each agent is also located in a social network (social space). We first assume that locations are exogenous and agents have to decide how often they want to visit the center, given that there a cost of commuting from the periphery. Each visit results in one interaction, so that the aggregate number of visits is a measure of aggregate interactivity. We examine how interaction choices depend on the interplay of social and physical distance and show that there exists a unique Nash equilibrium in agents' effort (i.e. number of visits or interactions to the center). We also show that agents who are more central in the social network, or are located closer to interaction center, choose higher levels of interactions in equilibrium. As a result, the level of interactivity in the economy as a whole will rise with the density of links in the social network and with the degree to which agents are clustered in physical space.

We then look at a subgame-perfect Nash equilibrium where agents first choose their geographic location and then their social effort. We characterize this equilibrium and give the condition under which there is a unique subgame-perfect Nash equilibrium. We also show that there is a tendency for agents who are more central in the social network to locate closer to the interaction center, leading to a form of endogenous geographic separation based on social distance. Interestingly, the network structure plays an important role in the determination of equilibrium. In a regular network where all agents have the same position in the network (like e.g. the complete or the circle network), there can only be two possible equilibria: either all agents live in the center or in the periphery. On the contrary, in a star network, apart from these two equilibria, there can be a core-periphery equilibrium where the star agent resides in the center while all the peripheral agents live in the periphery. More generally, if we define the type of an agent by her position in the network, we show that the number of equilibria is equal to the number of types plus one and we can give the condition under which each equilibrium exists and is unique.

Furthermore, we show that the market equilibrium is not optimal because of social externalities.

We determine the value of the subsidy to interactions that could support the first-best allocation as an equilibrium, and show that interaction effort and the incentives for clustering are higher under the subsidy program. We also look at a policy that subsidizes location in the geographical space and discuss the possibility of subsidizing both social interactions and geographical locations.

Finally, to better understand the policy implications of the model, we interpret the network as a labor-market network so that each visit (or interaction) to the center leads to an exchange of job information. If we further assume that the more central positions in the network are occupied by white workers while black workers are located in less central positions, our model can show that the less central agents in the network (i.e. black workers who do not have an old-boy network) will reside further away from jobs (i.e. the center) than more central agents (whites) and thus will experience adverse labor-market outcomes. This provides a new explanation of the so-called "spatial mismatch hypothesis" where distance to jobs is put forward as the main culprit for the adverse labor-market outcomes of minority workers. We provide here a new mechanism by putting forward the role of the social space (social mismatch) on the geographical space (spatial mismatch).

The paper is organized as follows. The next section highlights our contribution to the literature. Section 3 presents the basic model of interaction with social and physical distance, and solves for equilibrium interaction patterns. Section 4 extends the model to consider location choice and shows that agents who are more central in the social network will tend to locate closer to the center of interactions, ceteris paribus. Section 5 considers efficient interaction patterns and policies that will support the optimum as an equilibrium. Section 6 discusses the implications of our model in terms of black and white workers' outcomes when each visit to the center leads to an exchange of job information. Finally, Section 7 discusses our results and proposes some extensions.

2 Related literature

Our paper lies at the intersection of two different literatures. We would like to expose them in order to highlight our contribution.

Urban economics and economics of agglomeration There is an important literature in urban economics looking at how interactions between agents create agglomeration and city centers.[1] However, as stated in the Introduction, in most of these models, nonmarket interactions are basically a black box. There are recent papers where the nonmarket interactions are modeled in a more satisfactory way. Mossay and Picard (2011, 2012)[2] propose interesting models in which each

[1] See Fujita and Thisse (2002) for a literature review.

[2] See also Picard and Tabuchi (2010).

agent visits other agents so as to benefit from face-to-face communication and, as in our model, each trip involves a cost which is proportional to distance. The 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. Their models are different to ours since the network and its structure are not explicitly modeled. Furthermore, Ghiglino and Nocco (2012) extend the standard economic geography model a la Krugman to incorporate conspicuous consumption. In their model, agents are sensitive to comparisons within their own type group, which depends on the network structure. They show that agglomeration patterns depend on the network structure. Their model is quite different to ours and the networks considered are very specific (complete, segregated or star networks).

Peer effects and social networks There is a growing interest in theoretical models of peer effects and social networks (see e.g. Akerlof, 1997; Glaeser et al., 1996; Ballester et al., 2006; Calvó-Armengol et al., 2009). However, there are very few papers that consider the interaction of social and physical distance. Brueckner, Thisse and Zenou (2002), Helsley and Strange (2007), Brueckner and Largey (2008) and Zenou (2013) are exceptions but, in these models, the social network is not explicitly modeled.[3] Schelling (1971) is clearly a seminal reference when discussing social preferences and location. Shelling's model shows that, even a mild preference for interacting with people from the same community can lead to large differences in terms of location decision. Indeed, his results suggest that total segregation persists even if most of the population is tolerant about heterogeneous neighborhood composition.[4] Our model is conceptually very different from models a la Schelling since there is an explicit network structure and agents decide how much effort to exert in interacting with others. Finally, Johnson and Gilles (2000) extend the Jackson and Wolinsky (1996)'s connection model by introducing a cost of creating a link which is proportional to the geographical distance between two individuals. The model is very different since there is no location choice and no effort decision.[5]

To the best of our knowledge, our paper is the first one to provide a model that interacts the location of an agent in a social network and her geographical location.[6] It is also conceptually very

[3] See Ioannides (2012, Chap. 5) who reviews the literature on social interactions and urban economics.

[4] This framework has been modified and extended in different directions, exploring, in particular, the stability and robustness of this extreme outcome (see, for example, Zhang, 2004 or Grauwin et al., 2012).

[5] Brueckner (2006) proposes a model where individuals in a friendship network decide how much effort to exert in their relationships. The model is quite different since there is no location choice and the network is stochastically formed.

[6] Recent empirical researches have shown that the link between these two spaces is quite strong, especially within community groups (see e.g. Bayer et al., 2008; Hellerstein et al., 2011 and Patacchini and Zenou, 2012).

different from models of social preferences and location a la Schelling. Thus, the paper provides a first stab at a very important question in both social networks and urban economics.

3 Equilibrium interactions with exogenous location

3.1 The model

3.1.1 Locations and the social network

There are n agents in the economy. The geography consists of two locations, a *center*, where all interactions occur, and a *periphery*. All agents are located in either the center or the periphery. The distance between the center and the periphery is normalized to one. Thus, letting x_i represent the location of agent i, defined as her distance from the interaction center, we have $x_i \in \{0, 1\}, \forall i = 1, 2, ..., n$. In this section we assume that locations are exogenous; location choice is considered in Section 4.

The social space is a network. A network g is a set of *ex ante* identical agents $N = \{1, \ldots, n\}$, $n \geq 2$, and a set of *links* or *direct connections* between them. These connections influence the benefit that an agent receives from interactions, in a manner that is made precise below. The adjacency matrix $\mathbf{G} = [g_{ij}]$ keeps track of the direct connections in the network. By definition, agents i and j are directly connected if and only if $g_{ij} = 1$; otherwise, $g_{ij} = 0$. We assume that if $g_{ij} = 1$, then $g_{ji} = 1$, so the network is *undirected*.[7] By convention, $g_{ii} = 0$. $\mathbf{G}$ is thus a square $(0, 1)$ symmetric matrix with zeros on its diagonal. The *neighbors* of an agent i in network g are denoted by $\mathcal{N}_i$. We have: $\mathcal{N}_i = \{\text{all } j | g_{ij} = 1\}$. The *degree* of a node i is the number of neighbors that i has in the network, so that $d_i = |\mathcal{N}_i|$.

3.1.2 Preferences

Consumers derive utility from a numeraire good z and interactions with others according to the transferrable utility function

$$U_i(v_i, \mathbf{v_{-i}}, g) = z_i + u_i(v_i, \mathbf{v_{-i}}, g), \tag{1}$$

where v_i is the number of visits (effort) that agent i makes to the center, $\mathbf{v_{-i}}$ is the corresponding vector of visits for the other $n-1$ agents, and $u_i(v_i, \mathbf{v_{-i}}, g)$ is the subutility function of interactions. Thus, utility depends on the visit choice of agent i, the visit choices of other agents and on agent i's position in the social network g. We imagine that each visit results in one interaction, so that

[7] Our model can be extended to allow for *directed* networks (i.e. non-symmetric relationships) and *weighted links* in a straightforward way.

the aggregate number of visits is a measure of aggregate interactivity. For tractability, we assume that the subutility function takes the linear quadratic form

$$u_i(v_i, \mathbf{v_{-i}}, g) = \alpha v_i - \frac{1}{2}v_i^2 + \theta \sum_{j=1}^{n} g_{ij} v_i v_j, \tag{2}$$

where $\alpha > 0$ and $\theta > 0$ (the roles of these parameters will become clear shortly). Equation (2) imposes additional structure on the interdependence between agents; under (2) the utility of agent i depends on her own visit choice and on the visit choices of the agents with whom she is directly connected in the network, i.e., those for whom $g_{ij} = 1$.

Agents located in the periphery must travel to the center to interact with others. Letting y represent income and t represent marginal transport cost, budget balance implies that expenditure on the numeraire is

$$z_i = y - t x_i v_i. \tag{3}$$

Using this expression to substitute for z_i in (1), and using (2), gives

$$U_i(v_i, \mathbf{v_{-i}}, g) = y + \alpha_i v_i - \frac{1}{2}v_i^2 + \theta \sum_{j=1}^{n} g_{ij} v_i v_j, \tag{4}$$

where $\alpha_i = \alpha - t x_i$. We assume $\alpha > t$, so that $\alpha_i > 0$, $\forall x_i \in \{0, 1\}$ and hence $\forall i = 1, 2, ...n$. Note from (4) that utility is concave in own visits, $\frac{\partial^2 U_i}{\partial v_i^2} = -1$. Note also that the marginal utility of v_i is increasing in the visits of another with whom i is directly connected, $\frac{\partial^2 U_i}{\partial v_i \partial v_j} = \theta$, for $g_{ij} = 1$. Thus, v_i and v_j are *strategic complements* from i's perspective when $g_{ij} = 1$. Each agent i chooses v_i to maximize (4) taking the structure of the network and the visit choices of other agents as given. Before analyzing this game, we introduce a useful measure of an agent's importance in the social network.

3.1.3 The Katz-Bonacich network centrality measure

There are many ways to measure the importance or centrality of an agent in a social network. For example, degree centrality measures importance by the number of direct connections that an agent has with all others, while closeness centrality measures importance by the average distance (in terms of links in the network) between an agent and all others. See Wasserman and Faust (1994) and Jackson (2008) for discussions of these, and many other, characteristics of social and economic networks. The Katz-Bonacich centrality measure (due to Katz, 1953, and Bonacich, 1987), which has proven to be extremely useful in game theoretic applications (Ballester et al., 2006), "presumes that the power or prestige of a node is simply a weighted sum of the walks that emanate from it" (Jackson, 2008, pp. 41).

To formalize this measure, let $\mathbf{G}^k$ be the kth power of $\mathbf{G}$, with elements $g_{ij}^{[k]}$, where k is an integer. The matrix $\mathbf{G}^k$ keeps track of the indirect connections in the network: $g_{ij}^{[k]} \geq 0$ gives the number of walks or paths of length $k \geq 1$ from i to j in the network g. In particular, $\mathbf{G}^0 = \mathbf{I}$. Consider the matrix $\mathbf{M} = \sum_{k=0}^{+\infty} \theta^k \mathbf{G}^k$. The elements of this matrix, $m_{ij} = \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]}$, count the number of walks of all lengths from i to j in the network g, where walks of length k are weighted by θ^k. These expressions are well-defined for small enough values of θ.[8] The parameter θ is a decay parameter that scales down the relative weight of longer walks. Note that, when $\mathbf{M}$ is well-defined, one can write $\mathbf{M}-\theta\mathbf{G}\mathbf{M} = \mathbf{I}$ and hence $\mathbf{M} = [\mathbf{I}-\theta\mathbf{G}]^{-1}$.[9] The Katz-Bonacich centrality of agent i, denoted, $b_i(g,\theta)$ is equal to the sum of the elements of the ith row of $\mathbf{M}$:

$$b_i(g,\theta) = \sum_{j=1}^{n} m_{ij} = \sum_{j=1}^{n} \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]}. \tag{5}$$

The Katz-Bonacich centrality of any agent is zero when the network is empty. It is also zero for $\theta = 0$, and is increasing and convex in θ for $\theta > 0$. For future reference, it is convenient to note that the $(n \times 1)$ vector of Katz-Bonacich centralities can be written in matrix form as

$$\mathbf{b}(g,\theta) = \mathbf{M}\mathbf{1} = [\mathbf{I} - \theta\mathbf{G}]^{-1}\mathbf{1}, \tag{6}$$

where $\mathbf{1}$ is the $n-$dimensional vector of ones. We can also define the *weighted* Katz-Bonacich centrality of agent i as:

$$b_{\alpha_i}(g,\theta) = \sum_{j=1}^{n} \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]} \alpha_j, \tag{7}$$

[8] The matrix power series $\sum_{k=0}^{+\infty} \theta^k \mathbf{G}^k$ converges if and only if

$$\|\mathbf{G}\| < r = \lim_{k\to\infty} \inf \left|\theta^k\right|^{-1/k} = \frac{1}{\theta}$$

where r is the radius of convergence and $\|\mathbf{G}\|$ is the "norm" of the matrix $\mathbf{G}$. This norm is generally taken to be the "spectral radius" of $\mathbf{G}$, written $\rho(\mathbf{G}) = \max_i |\lambda_i|$, where λ_i is an eigenvalue of $\mathbf{G}$. Thus, the matrix power series converges, and $\mathbf{M}$ is well-defined, for $\theta\rho(\mathbf{G}) < 1$. Convergence of the matrix power series constructively establishes the existence of the inverse $[\mathbf{I} - \theta\mathbf{G}]^{-1}$, where $\mathbf{I}$ is the identity matrix. The condition $\theta\rho(\mathbf{G}) < 1$ relates the payoff function to the network topology. When this condition holds, the local payoff interdependence θ is lower than the inverse of the spectral radius of $\mathbf{G}$, which is a measure of connectivity in the network. When this condition does not hold, existence of equilibrium becomes an issue because the strategy space is unbounded (see Ballester et al., 2006).

[9] Indeed, expanding the power series gives

$$\mathbf{M} = \mathbf{I} + \theta\mathbf{G} + \theta^2\mathbf{G}^2 + ...,$$

which implies,

$$\theta\mathbf{G}\mathbf{M} = \theta\mathbf{G} + \theta^2\mathbf{G}^2 + \theta^3\mathbf{G}^3 + ...$$

Subtracting the latter from the former gives $\mathbf{M}-\theta\mathbf{G}\mathbf{M} = \mathbf{I}$.

where the weight attached to the walks from i to j is α_j. For any $n-$dimensional vector $\boldsymbol{\alpha}$, the matrix equivalent of (7) is given by:

$$\mathbf{b}_{\boldsymbol{\alpha}}(g,\theta) = \mathbf{M}\boldsymbol{\alpha} = [\mathbf{I} - \theta\mathbf{G}]^{-1}\boldsymbol{\alpha}$$

3.2 Nash equilibrium visits and interactivity

The first-order condition for a maximum of (4) with respect to v_i gives the best-response function

$$v_i^* = \alpha_i + \theta\sum_{j=1}^{n} g_{ij}v_j^* \qquad \forall i = 1, 2, ...n. \tag{8}$$

Thus, due to the linear quadratic form in (2), the optimal visit choice of agent i is a linear function of the visit choices of the agents to whom i is directly connected in the network. In matrix form the system in (8) becomes $\mathbf{v} = \boldsymbol{\alpha} + \theta\mathbf{G}\mathbf{v}$, where $\boldsymbol{\alpha}$ is the $(n \times 1)$ vector of the α_i's. Solving for $\mathbf{v}$ and using (6) gives the Nash equilibrium visit vector $\mathbf{v}^*$:

$$\mathbf{v}^* = [\mathbf{I}-\theta\mathbf{G}]^{-1}\boldsymbol{\alpha} = \mathbf{M}\boldsymbol{\alpha}. \tag{9}$$

The Nash equilibrium visit choice of agent i is

$$v_i^*(x_i, \mathbf{x_{-i}}, g) = \sum_{j=1}^{n} m_{ij}\alpha_j = \sum_{j=1}^{n}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}\alpha_j, \tag{10}$$

where $\mathbf{x_{-i}}$ is the vector of locations for the other $n-1$ agents. The expression on the right in (10) is the weighted Katz-Bonacich centrality of agent i defined in (7) above. This analysis is summarized by the following proposition where $\rho(\mathbf{G})$ is the spectral radius of the adjacency matrix $\mathbf{G}$:[10]

Proposition 1 (Equilibrium visits) *For any network g and for sufficiently small θ, i.e. $\theta\rho(\mathbf{G}) < 1$, there exists a unique, interior Nash equilibrium in visit choices in which the number of visits by any agent i equals her weighted Katz-Bonacich centrality,*

$$v_i^*(x_i, \mathbf{x_{-i}}, g) = b_{\alpha_i}(g,\theta). \tag{11}$$

The Nash equilibrium number of visits $v_i^*(x_i, \mathbf{x_{-i}}, g)$ depends on position in the social network and geographic location. Proposition 1 implies that an agent who is more central in the social network, as measured by her Katz-Bonacich centrality, will make more visits to the interaction center in equilibrium. Intuitively, agents who are better connected have more to gain from interacting with others and so exert higher interaction effort for any vector of geographic locations.

[10] All proofs can be found in the Appendix.

We would like to see how the equilibrium number of visits $v_i^*(x_i, \mathbf{x_{-i}}, g)$ varies with the different parameters of the model. It is straightforward to verify that $v_i^*(x_i, \mathbf{x_{-i}}, g)$ increases with α and decreases with commuting costs t. It is also straighforward to analyze the relationship between $v_i^*(x_i, \mathbf{x_{-i}}, g)$ and the intensity of social interactions θ, which is also a measure of complementarity in the network.[11] We have the following the result.

Proposition 2 (Intensity of social interactions) *Assume $\theta\rho(\mathbf{G}) < 1$. Then, for any network, an increase in the intensity of social interactions θ raises the equilibrium number of visits $v_i^*(x_i, \mathbf{x_{-i}}, g)$ by any agent i.*

When there are a lot of synergies from social interactions, each agent finds it desirable to visit the center more because the benefits are higher. The same intuition prevails for α. On the contrary, when commuting costs increase, then the number of visits to the center decreases.

Let us now analyze aggregate effects. From (10), $v_i^*(x_i, \mathbf{x_{-i}}, g)$ is non-increasing in x_i,

$$v_i^*(1, \mathbf{x_{-i}}, g) - v_i^*(0, \mathbf{x_{-i}}, g) = -t\, m_{ii} \leq 0 \tag{12}$$

since $\mathbf{M}$ is a non-negative matrix. Any agent for whom $m_{ii} > 0$ will make more interaction visits, or exert higher interaction effort, when located in the center rather than the periphery. In fact, reflecting the complementarity in visit choices, the equilibrium visit choice of agent i is non-increasing in the distance of *any* agent from the interaction center. Letting $\mathbf{x_{-ik}}$ be the vector of locations for all agents except i and k, so $\mathbf{x_{-i}} = (x_k, \mathbf{x_{-ik}})$, we have

$$v_i^*(x_i, (1, \mathbf{x_{-ik}}), g) - v_i^*(x_i, (0, \mathbf{x_{-ik}}), g) = -t\, m_{ik} \leq 0, \qquad \forall k \neq i. \tag{13}$$

Let $V^*(g)$ represent the equilibrium aggregate level of visits, or, for simplicity, the equilibrium aggregate level of interactions. From (10) and (7), we have

$$V^*(g) = \sum_{i=1}^{i=n} v_i^*(x_i, \mathbf{x_{-i}}, g) = \sum_{i=1}^{i=n} b_{\alpha_i}(g, \theta) \tag{14}$$

Consider an alternative social network g', $g' \neq g$ such that for all i, j, $g'_{ij} = 1$ if $g_{ij} = 1$. It is conventional to refer to g and g' as *nested* networks, and to denote their relationship as $g \subset g'$. As discussed in Ballester et al. (2006), the network g' has a denser structure of network links: some agents who are not directly connected in g are directly connected in g'. Then, given the

[11] Recall that
$$\frac{\partial^2 U_i}{\partial v_i \partial v_j} = \theta \text{ for } g_{ij} = 1.$$

complementarities in the network, it must be the case that equilibrium visits are weakly larger for all agents, which implies $V^*(g') > V^*(g)$. Similarly, (12) and (13) imply that $V^*(g)$ is non-increasing in the distance of any agent from the interaction center. Thus, the more compact is the spatial arrangement of agents, the greater is the level of aggregate interactions for any network g. Furthemore, because of local complementarities, denser networks also increase each bilateral interaction between two individuals. This analysis is summarized in the following proposition:

Proposition 3 (Aggregate interactions) *For sufficiently small θ, aggregate interactions as well as the entire vector of individual interactions increase with the density of network links and decrease with the distance of any agent from the interaction center.*

This is an interesting result since it analyzes the relationship between network structure and aggregate interactions as well as individual interactions. It says, for example, that a star-shaped network will have fewer social interactions than a complete network because agents enjoy fewer *local complementarities* in the former than in the latter.

3.3 Example

To illustrate the previous results, consider the following star-shaped social network g with three agents (i.e. $n = 3$), where agent 1 holds a central position whereas agents 2 and 3 are peripherals:

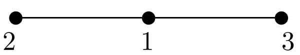

Figure 1: A star network with 3 individuals

The adjacency matrix for this social network is given by:

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Its is a straightforward algebra exercise to compute the powers of this matrix, which are:

$$\mathbf{G}^{2k} = \begin{bmatrix} 2^k & 0 & 0 \\ 0 & 2^{k-1} & 2^{k-1} \\ 0 & 2^{k-1} & 2^{k-1} \end{bmatrix} \quad \text{and} \quad \mathbf{G}^{2k+1} = \begin{bmatrix} 0 & 2^k & 2^k \\ 2^k & 0 & 0 \\ 2^k & 0 & 0 \end{bmatrix}, k \geq 1.$$

For instance, we deduce from $\mathbf{G}^3$ that there are exactly two walks of length three between agents 1 and 2, namely, $12 \to 21 \to 12$ and $12 \to 23 \to 32$. Obviously, there is no walk of this length (and, in general, of odd length) from any agent to herself. It is easily verified that:

$$\mathbf{M} = [\mathbf{I} - \theta \mathbf{G}]^{-1} = \frac{1}{1-2\theta^2} \begin{bmatrix} 1 & \theta & \theta \\ \theta & 1-\theta^2 & \theta^2 \\ \theta & \theta^2 & 1-\theta^2 \end{bmatrix}$$

We can now compute the agents' centrality measures using (11). We obtain:[12]

$$\begin{bmatrix} v_1^* \\ v_2^* \\ v_3^* \end{bmatrix} = \begin{bmatrix} b_{\alpha_1}(\theta, g) \\ b_{\alpha_2}(\theta, g) \\ b_{\alpha_3}(\theta, g) \end{bmatrix} = \frac{1}{1-2\theta^2} \begin{bmatrix} \alpha_1 + \theta(\alpha_2 + \alpha_3) \\ \theta\alpha_1 + \left(1-\theta^2\right)\alpha_2 + \theta^2\alpha_3 \\ \theta\alpha_1 + \theta^2\alpha_2 + \left(1-\theta^2\right)\alpha_3 \end{bmatrix}$$

Suppose now that, for exogenous reasons, individual 1 resides in the center, i.e., $x_1 = 0$ while individuals 2 and 3 live at the periphery, i.e., $x_2 = x_3 = 1$. This implies that $\alpha_1 = \alpha$ and $\alpha_2 = \alpha_3 = \alpha - t > 0$. Thus, we now have:

$$\begin{bmatrix} v_1^* \\ v_2^* \\ v_3^* \end{bmatrix} = \frac{1}{1-2\theta^2} \begin{bmatrix} \alpha + 2\theta(\alpha - t) \\ \alpha(1+\theta) - t \\ \alpha(1+\theta) - t \end{bmatrix} \tag{15}$$

It is easily verified that:[13]

$$v_1^* > v_2^* = v_3^*$$

In that case, the effort exerted by agent 1, the most central player, is the highest one. As a result, agents located closer to the center have higher centrality $b_{\alpha_i}(g, \theta)$ and thus higher effort (i.e. they visit more often the center to interact with other people). Note that, in equilibrium, each agent i's effort is affected by the location of all other agents in the network but distant neighbors have less impact due to the decay factor θ in the Katz-Bonacich centrality.

The equilibrium aggregate level of interactions in a network is then given by:

$$V^*(g) = \sum_{i=1}^{i=n} v_i^* = \frac{(3+4\theta)\alpha - 2(1+\theta)t}{\left(1-2\theta^2\right)}$$

Let us now illustrate Proposition 3. Consider the network described in Figure 1 and add one link between individuals 2 and 3 so that we switch from a star-shaped network to a complete one.

[12] Note that this centrality measures are only well-defined when $\theta < 1/\sqrt{2}$ or $\theta^2 < 1/2$ (condition on the largest eigenvalue).

[13] Observe that this inequality is true because we have assumed that $\theta < 1/\sqrt{2}$ (this guarantees that the Katz-Bonacich centrality is well-defined) and $\alpha > t$.

Suppose that we have the same geographical configuration, i.e. individual 1 resides in the center while individuals 2 and 3 lives at the periphery, i.e., $\alpha_1 = \alpha$ and $\alpha_2 = \alpha_3 = \alpha - t > 0$.[14] We easily obtain:[15]

$$\begin{bmatrix} v_1^* \\ v_2^* \\ v_3^* \end{bmatrix} = \frac{1}{\left(1-\theta-2\theta^2\right)} \begin{bmatrix} \alpha\left(1+\theta\right)-2t\theta \\ \alpha\left(1+\theta\right)-t \\ \alpha\left(1+\theta\right)-t \end{bmatrix}$$

Not surprisingly, given that $\theta < 0.5$, $v_1^* > v_2^* = v_3^*$ since all individuals have the same position in the social network but individual 1 has an "advantage" in the geographical space by locating in the center. Total activity in this network, denoted by $g^{[+23]}$, is then equal to:

$$V^*(g^{[+23]}) = \frac{(3\alpha - 2t)\left(1+\theta\right)}{1-2\theta^2-\theta} > V^*(g)$$

This confirms the fact that denser networks (complete networks) generate more aggregate and bilateral activities than less dense networks (star networks).

4 Location choice

4.1 Model and subgame-perfect equilibrium

This section extends our model of social networks and interaction to allow agents to choose between locating in the center and the periphery. We suppose that there is an exogenous cost differential $c > 0$ associated with the central location. Assuming that the center has more economic activity generally, this cost differential might arise from congestion effects or reflect a difference in location land rent from competition among other activities for center locations. Agents choose locations to maximize net utility, that is, utility from interactions minus the exogenous location cost, taking the visits of other agents as given.

The timing is now as follows. In the first stage, agents decide where to locate ($x = 0$ or $x = 1$) while, in the second stage they decide their optimal effort in the network. Thus, we look at subgame-perfect equilibria. As usual, we solve the model backward. The second stage has already been solved and Proposition 1 showed that, if $\theta\rho(\mathbf{G}) < 1$, the exists a unique effort level for each individual i given by: $v_i^*(x_i, \mathbf{x_{-i}}, g) = b_{\alpha_i}(g, \theta)$. Using the best-response function (8), we can write the equilibrium utility level of agent i as:

[14] This is just for the sake of illustrating Proposition 3. We will see below that such an equilibrium cannot exist in a complete network.

[15] It is easily verified that the condition on the largest eigenvalue is now given by: $\theta < 1/2$.

$$U_i^*(v_i^*, \mathbf{v}_{-\mathbf{i}}^*, g) = y + \frac{1}{2}\left[v_i^*(x_i, \mathbf{x}_{-\mathbf{i}}, g)\right]^2 = y + \frac{1}{2}\left[b_{\alpha_i}(g,\theta)\right]^2 \tag{16}$$

where $v_i^*(0, \mathbf{x}_{-\mathbf{i}}, g)$ and $v_i^*(1, \mathbf{x}_{-\mathbf{i}}, g)$ are the equilibrium effort of individual i if she lives in the center and in the periphery, respectively. As a result, the equilibrium utility of each agent i is equal to her income plus half of her equilibrium effort squared. We need now to solve the first stage of the game, i.e. the location choice. What is complicated here is that the weighted Katz-Bonacich centralities are endogenous equilibrium objects and thus one needs to know the equilibrium location configuration in order to build the equilibrium.

Let us now characterize the equilibrium.

Define $\mathcal{C}$ as the set of *central* agents (i.e. all individuals who live in the center) and $\mathcal{P}$ as the set of *peripheral* agents (i.e. all individuals who live in the periphery). If individual i resides in the center ($x = 0$), her equilibrium utility is equal to:[16]

$$U_i^*(v_i^*(0, \mathbf{x}_{-\mathbf{i}}, g), \mathbf{v}_{-\mathbf{i}}^*, g) = y + \frac{1}{2}\left[\sum_{j\in\mathcal{C}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}\alpha + \sum_{j\in\mathcal{P}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}(\alpha - t) + \sum_{k=0}^{+\infty}\theta^k g_{ii}^{[k]}\alpha\right]^2 - c$$

We have here decomposed the Katz-Bonacich centrality $b_{\alpha_i}(g,\theta)$ into self-loops ($m_{ii} = \sum_{k=0}^{+\infty}\theta^k g_{ii}^{[k]}$) and non self-loops ($m_{ij} = \sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}$) and give different weights to these paths depending if agents live in the center (weight α) or in the periphery (weight $\alpha - t$). Similarly, if individual i resides in the periphery ($x = 1$), her equilibrium utility is equal to:

$$U_i^*(v_i^*(1, \mathbf{x}_{-\mathbf{i}}, g), \mathbf{v}_{-\mathbf{i}}^*, g) = y + \frac{1}{2}\left[\sum_{j\in\mathcal{C}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}\alpha + \sum_{j\in\mathcal{P}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}(\alpha - t) + \sum_{k=0}^{+\infty}\theta^k g_{ii}^{[k]}(\alpha - t)\right]^2$$

As a result, individual i will live at $x = 0$ if and only if $U_i^*(v_i^*(0, \mathbf{x}_{-\mathbf{i}}, g), \mathbf{v}_{-\mathbf{i}}^*, g) > U_i^*(v_i^*(1, \mathbf{x}_{-\mathbf{i}}, g), \mathbf{v}_{-\mathbf{i}}^*, g)$.

Denote by

$$b_\alpha^{[-ii]}(g,\theta) \equiv \alpha\sum_{j=1, j\neq i}^{n} m_{ij} = \alpha\sum_{j\in\mathcal{C}-\{i\}} m_{ij} + \alpha\sum_{j\in\mathcal{P}-\{i\}} m_{ij}$$

the weighted Katz-Bonacich centrality without self-loops and

$$m_{ii}^{(2)} \equiv \frac{-\left(b_\alpha^{[-ii]}(g,\theta) - t\sum_{j\in\mathcal{P}-\{i\}} m_{ij}\right) + \sqrt{\left(b_\alpha^{[-ii]}(g,\theta) - t\sum_{j\in\mathcal{P}-\{i\}} m_{ij}\right)^2 + \frac{2c}{t}(2\alpha - t)}}{2\alpha - t}$$

We have the following result:

[16] Observe that $\mathcal{C}-\{i\}$ and $\mathcal{P}-\{i\}$ denotes respectively the set of all *central* agents but i and the set of all *peripheral* agents but i.

Proposition 4 (Characterization of equilibrium locations) *Assume $\theta\rho(\mathbf{G}) < 1$. Then all individuals i with a $m_{ii} > m_{ii}^{(2)}$ will reside in the center of the city (i.e. $x = 0$) while all individuals i with a $m_{ii} < m_{ii}^{(2)}$ will live at the periphery of the city (i.e. $x = 1$). In other words, in equilibrium,*

$$\mathcal{C} = \{\textit{all is for which } m_{ii} > m_{ii}^{(2)}\}$$

and

$$\mathcal{P} = \{\textit{all is for which } m_{ii} < m_{ii}^{(2)}\}$$

Figure 2 displays the equilibrium characterization of Proposition 4.

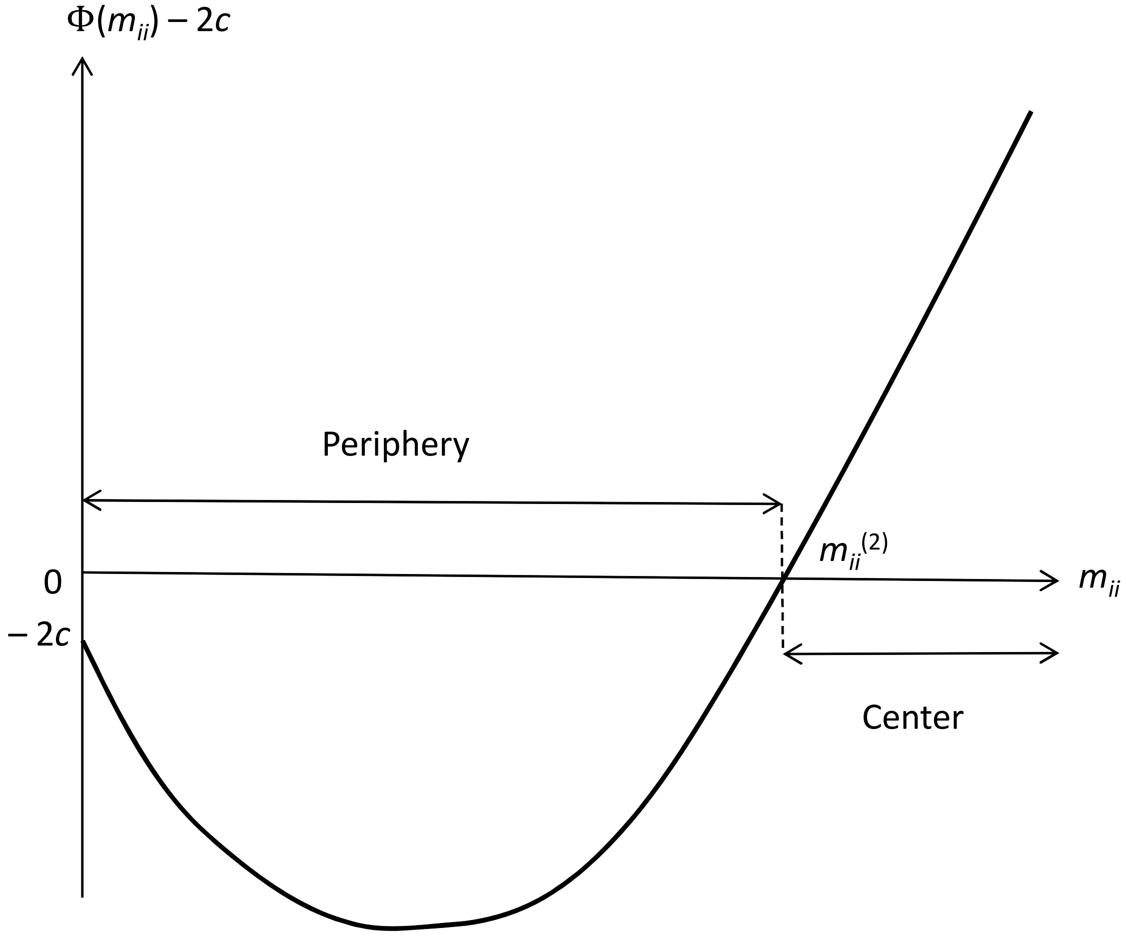

Figure 2: Characterization of equilibrium locations

Proposition 4 expresses the salient relationship between position in the social network and geographic location. Remember that $m_{ii} = \sum_{k=0}^{+\infty} \theta^k g_{ii}^{[k]}$ counts the number of paths in g starting from i and ending at i (self-loops) where paths of length k are weighted by θ^k, and $m_{ij,j\neq i} = \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]}$ counts the number of paths in g starting from i and ending at $j \neq i$ (non self-loops) where paths of length k are weighted by θ^k. Remember also that the Katz-Bonacich centrality is: $b_i(g, \theta) = m_{ii} + \sum_{j=1,j\neq i}^{n} m_{ij}$ while the *weighted* Katz-Bonacich centrality is given by:

$$b_{\alpha_i}(g, \theta) = \alpha_i m_{ii} + \alpha \sum_{j \in \mathcal{C}} m_{ij} + (\alpha - t) \sum_{j \in \mathcal{P}} m_{ij}$$

where $\alpha_i = \alpha$ if i lives in the center and $\alpha_i = \alpha - t$ if i lives in the periphery. As a result, m_{ii} captures the *centrality* in the network of each individual i. If participation in a social network involves costly transportation, then agents who occupy more central positions in the social network will have the most to gain from locating at the interaction center. In our model with two locations, in equilibrium *agents who are most central in the social network* (higher m_{ii}) *will locate at the interaction center*, while *agents who are less central in the social network* (lower m_{ii}) *locate in the periphery*. There is, in effect, endogenous geographic separation by position in the social network.

We would now like to deal with the issues of existence and uniqueness of the subgame-perfect equilibrium location-effort. For that, consider any network with n agents. Rank agents in the network such that we start with agent 1 who has the highest centrality in the network, i.e. $m_{11} = \max_i m_{ii}$, then we have agent 2 who has the next highest centrality, etc. until we reach agent n who has the lowest centrality in the network, i.e. $m_{nn} = \min_i m_{ii}$. Define each agent by her type, where the type of an agent is her Katz-Bonacich centrality (or her m_{ii}). Since two agents can have the same centrality, there are $\omega \leq n$ types in each network of n agents. Denote by

$$\Phi^C(m_{ii}) \equiv t\,(2\alpha - t)\,(m_{ii})^2 + 2t \left(\alpha \sum_{j=N-\{i\}} m_{ij} \right) m_{ii} \qquad (17)$$

where all the m_{ii}s and m_{ij}s are defined by the cells of the matrix $\mathbf{M} = [\mathbf{I} - \theta \mathbf{G}]^{-1}$. We have the following result where "equilibrium" means "Subgame-Perfect Nash equilibrium":

Proposition 5 (Existence and uniqueness of equilibrium locations) *Assume $\theta\rho(\mathbf{G}) < 1$ and consider any network of n agents with $\omega \leq n$ types. In any equilibrium, two agents with the same Katz-Bonacich centrality have to reside in the same part of the city and agents with higher Katz-Bonacich centrality cannot reside further away from the center than agents with lower Katz-Bonacich centrality. Moreover, the number of equilibria is equal to the number of types of agents plus one, i.e. $\omega + 1$.*

If the number of types is the same as the number of agents, we can characterize the locational (subgame-perfect) equilibria as follows:

(i) If

$$2c < \Phi^{\mathcal{C}}(m_{nn})$$

there exists a unique Central equilibrium where all agents live in the center, i.e. $\mathcal{C} = N$ *and* $\mathcal{P} = \varnothing$.

(ii) If

$$\Phi^{\mathcal{C}}(m_{nn}) < 2c < \Phi^{\mathcal{C}}(m_{n-1n-1}) - 2t^2 m_{n-1n} m_{n-1n-1}$$

there exists a unique Core-Periphery equilibrium such that $\mathcal{C} = N - \{n\}$ *and* $\mathcal{P} = \{n\}$.

(iii) If

$$\Phi^{\mathcal{C}}(m_{n-1n-1}) - 2t^2 m_{n-1n} m_{n-1n-1} < 2c < \Phi^{\mathcal{C}}(m_{n-2n-2}) - 2t^2 \left(m_{n-2n-1} + m_{n-2n}\right) m_{n-2n-2}$$

there exists a unique Core-Periphery equilibrium such that $\mathcal{C} = N - \{n-1, n\}$ *and* $\mathcal{P} = \{n-1, n\}$.

(iv) If

$$\Phi^{\mathcal{C}}(m_{n-2n-2}) - 2t^2 \left(m_{n-2n-1} + m_{n-2n}\right) m_{n-2n-2} < 2c < \Phi^{\mathcal{C}}(m_{n-3n-3}) - 2t^2 \left(\sum_{j \in \mathcal{P}}^{n} m_{n-3j}\right) m_{n-3n-3}$$

there exists a unique Core-Periphery equilibrium such that $\mathcal{C} = N - \{n-2, n-1, n\}$ *and* $\mathcal{P} = N - \{n-2, n-1, n\}$.

(v) etc. until we arrive at agent 1 *who has the highest centrality. Then,*

(vi) If

$$\Phi^{\mathcal{C}}(m_{11}) - 2t^2 \left(\sum_{j \in \mathcal{P} - \{1\}}^{n} m_{1j}\right) m_{11} < 2c$$

there exists a unique Peripheral equilibrium where all agents live in the periphery, i.e. $\mathcal{C} = \varnothing$ *and* $\mathcal{P} = N$.

If the number of types is less than the number of agents, then each step described above has to be made by type and not by individual so that each subscript refers to types and not to individuals.

This proposition totally characterizes the (subgame-perfect Nash) equilibrium locations and shows that there always exists a unique equilibrium within each interval. Interestingly, we could characterize everything in terms of $\Phi^{\mathcal{C}}(m_{ii})$, which is the "incentive function" when there is a

Central equilibrium, i.e. when all agents reside in the center of the city. Indeed, when, for all i, $\Phi^{\mathcal{C}}(m_{ii}) > 2c$, all individuals live in the center and we have a unique Central equilibrium. Then, when we start to move people from the center to the periphery, we need to change the weight in the Katz-Bonacich centrality from α (when living in the center) to $\alpha - t$ (when living in the periphery). This corresponds to the terms of both the right-hand side and left-hand side of each inequality since this is what is needed to be compensated for the agents living at the periphery of the city compared to the Central equilibrium where these agents lived in the center. Interestingly, there cannot be multiple equilibria within the same set of parameters.

Let us now perform a comparative statics exercise of the key parameters of the model.

Proposition 6 (Spatial concentration in the center) *Assume $\theta\rho(\mathbf{G}) < 1$. A decrease in the cost c of locating in the center, a decrease in marginal transport cost t, or an increase in the intensity of social interactions θ, will lead to more spatial concentration of agents in the center.*

Proposition 6 states that a decrease in c will increase the number of agents living in the center, which leads to more spatial concentration at the interaction center. An increase in marginal transport cost t, will have a similar impact. Finally, an increase in θ, the intensity of social interactions, will also lead to more spatial concentration in the center. Indeed, when θ increases, social interactions become more valuable and, because it is costly to commute to the center from the periphery, the spatial concentration at the interaction center increases. Therefore, this proposition allows us to analyze how endogenous spatial location affects the contribution equilibrium efforts. From Proposition 3, we know that aggregate interactions decrease with the distance of any agent from the interaction center. As a result, when, for example, c decreases, more agent choose to live in the center, which, in turn, increases social interactions in the network and thus equilibrium efforts. It is thus interesting here to see how *the geographical space affects the social space.*

4.2 Examples

4.2.1 Star-shaped networks: two types of agents

Let us return to the network described in Figure 1. Remember from Section 3.3, that, if $\theta < 1/\sqrt{2}$, then

$$\mathbf{M} = [\mathbf{I} - \theta\mathbf{G}]^{-1} = \frac{1}{1-2\theta^2}\begin{bmatrix} 1 & \theta & \theta \\ \theta & 1-\theta^2 & \theta^2 \\ \theta & \theta^2 & 1-\theta^2 \end{bmatrix} \tag{18}$$

In particular, this means that,

$$m_{11} = \frac{1}{1-2\theta^2} \text{ and } m_{22} = m_{33} = \frac{1-\theta^2}{1-2\theta^2}$$

We have the following result.

Proposition 7 (Locational equilibrium for a star-shaped network) *Consider the star-shaped network depicted in Figure 1 and assume that $\theta < 1/\sqrt{2} = 0.707$.*

(i) If

$$c < \frac{t(1-\theta)(1+\theta)^2[2\alpha-(1-\theta)t]}{2(1-2\theta^2)^2} \tag{19}$$

there exists a unique Central equilibrium where all agents live in the center, i.e. $\mathcal{C} = \{1,2,3\}$ and $\mathcal{P} = \varnothing$.

(ii) If

$$\frac{t(1-\theta)(1+\theta)^2[2\alpha-(1-\theta)t]}{2(1-2\theta^2)^2} < c < \frac{t[2\alpha(1+2\theta)-t(1+4\theta)]}{2(1-2\theta^2)^2} \tag{20}$$

there exists a unique Core-Periphery equilibrium where the star agent lives in the center while the peripheral agents reside in the periphery, i.e. $\mathcal{C} = \{1\}$ and $\mathcal{P} = \{2,3\}$.

(iii) If

$$c > \frac{t[2\alpha(1+2\theta)-t(1+4\theta)]}{2(1-2\theta^2)^2} \tag{21}$$

there exists a unique Peripheral equilibrium where all agents live in the periphery, i.e. $\mathcal{C} = \varnothing$ and $\mathcal{P} = \{1,2,3\}$.

This proposition shows that, for the star-shaped network, there are only three types of equilibria (i.e. number of types plus 1). It also shows the role of c and of t in the location decision process. For fixed values of α, t and θ, when we increase c, we switch from a central equilibrium to a core-periphery equilibrium and then to peripheral equilibrium. Interestingly, for fixed values of α, t and c, when we decrease θ we obtain the same types of result because an increase in θ means that social interactions are more valuable and thus tend to induce people to live to the center. The effect of an increase of t is similar.

We can give some parameter values for which each condition is satisfied given that $\theta < 0.707$. For example, if we set $\alpha = 6$, $t = 1$ and $\theta = 0.2$, then: *(i)* if $c < 7.62$, there exists a unique Central equilibrium where $\mathcal{C} = \{1,2,3\}$ and $\mathcal{P} = \varnothing$; *(ii)* if $7.62 < c < 8.86$, there is a unique Core-Periphery equilibrium where $\mathcal{C} = \{1\}$ and $\mathcal{P} = \{2,3\}$; *(iii)* if $c > 8.86$, there exists a unique Peripheral equilibrium where $\mathcal{C} = \varnothing$ and $\mathcal{P} = \{1,2,3\}$.

In each case, we can calculate the equilibrium utility of each agent. For example, if we consider the Central equilibrium, then the equilibrium utility of agent 1 is equal to:

$$U_1^*(v_1^*(0,0,0,g),\mathbf{v}_{-\mathbf{1}}^*,g) = y + \frac{\alpha^2\left(1+2\theta\right)^2}{2\left(1-2\theta^2\right)^2} - c$$

while the equilibrium utilities of agents 2 and 3 are given by

$$U_2^*(v_2^*(0,0,0,g),\mathbf{v}_{-\mathbf{2}}^*,g) = U_3^*(v_3^*(0,0,0,g),\mathbf{v}_{-\mathbf{3}}^*,g) = y + \frac{\alpha^2\left(1+\theta\right)^2}{2\left(1-2\theta^2\right)^2} - c$$

Not surprisingly, agent 1, who is the most central agent in the network, provides a higher effort and thus has a higher utility than the two other agents. At the other extreme, if there is a Peripheral equilibrium, then, to calculate the equilibrium utilities of all agents, one needs to replace α^2 by $(\alpha - t)^2$ and to remove the cost c in the expressions above. Finally, in the Core-Periphery equilibrium, $\mathcal{C} = \{1\}$ and $\mathcal{P} = \{2,3\}$, we obtain:[17]

$$U_1^*(v_1^*(0,1,1,g),\mathbf{v}_{-\mathbf{1}}^*,g) = y + \frac{\left[2\left(\alpha-t\right)\theta+\alpha\right]^2}{2\left(1-2\theta^2\right)^2} - c$$

$$U_2^*(v_2^*(1,0,1,g),\mathbf{v}_{-\mathbf{2}}^*,g) = U_3^*(v_3^*(1,0,1,g),\mathbf{v}_{-\mathbf{3}}^*,g) = y + \frac{\left(\alpha\theta-t+\alpha\right)^2}{2\left(1-2\theta^2\right)^2}$$

It is easily verified that all agents would be better off by living in the center if the cost c is not too large. This result can clearly be generalized for a star network with n agents where there will be 3 types of equilibria as in Proposition 7.

4.2.2 Complete networks: One type of agent

Let us now consider a complete network and, as in the previous section, set $n = 3$ (the generalization to n agents is straightforward). If $\theta < 1/2$, then

$$\mathbf{M} = \left[\mathbf{I} - \theta\mathbf{G}\right]^{-1} = \frac{1}{1-\theta-2\theta^2}\begin{bmatrix} 1-\theta & \theta & \theta \\ \theta & 1-\theta & \theta \\ \theta & \theta & 1-\theta \end{bmatrix} \tag{22}$$

We have the following result.

[17] Inside the utility function, the equilibrium effort $v_i^*(x_i, \mathbf{x_{-i}}, g)$ is written such that the first element in the parenthesis is the location of agent i while the other elements are the locations of all other agents by increasing numbering order, starting from agent 1 if $1 \neq i$. For example, for the star network with three agents, $v_1^*(0,1,1,g)$ is the equilibium effort of agent 1 (the star) for the Core-Periphery equilibrium $\mathcal{C} = \{1\}$ and $\mathcal{P} = \{2,3\}$ since $x_1 = 0$ and $x_2 = x_3 = 1$ while $v_2^*(1,0,1,g)$ is the equilibium effort of agent 2 (peripheral agent) for the same Core-Periphery equilibrium.

Proposition 8 (Locational equilibrium for a complete network) *Consider the complete network with 3 agents and assume that $\theta < 1/2$.*

(*i*) *If*

$$c < \frac{t\,(1-\theta)^2\,(2\alpha + 4\alpha t - t)}{2\left(1-\theta-2\theta^2\right)^2}$$

there exists a unique Central equilibrium where all agents live in the center, i.e. $\mathcal{C} = \{1,2,3\}$ *and* $\mathcal{P} = \varnothing$.

(*ii*) *If*

$$c > \frac{t\,(1-\theta)^2\,(2\alpha + 4\alpha t - t)}{2\left(1-\theta-2\theta^2\right)^2}$$

there exists a unique Peripheral equilibrium where all agents live in the periphery, i.e. $\mathcal{C} = \varnothing$ *and* $\mathcal{P} = \{1,2,3\}$.

This proposition completely characterizes the equilibrium configuration for a complete network. As showed in Proposition 5, there are no multiple equilibria and no core-periphery equilibrium. We can give parameter values for which each condition is satisfied given that $\theta < 0.5$. For example, if take exactly the same parameters as for the star network, i.e. $\alpha = 6$, $t = 1$ and $\theta = 0.2$, then: (*i*) if $c < 21.61$, there exists a unique Central equilibrium where $\mathcal{C} = \{1,2,3\}$ and $\mathcal{P} = \varnothing$; (*ii*) if $c > 21.61$, there exists a unique Peripheral equilibrium where $\mathcal{C} = \varnothing$ and $\mathcal{P} = \{1,2,3\}$.

It is straightforward to generalize this result for a complete network with n agents but also for any *regular network*. Using the argument of the proof, we can state that, for any regular network (i.e. each agent has the same number of links) with n agents, only two equilibria will emerge: the Central and the Peripheral equilibrium. If c is low enough, there will be a unique Central equilibrium while, if c is high enough, there will be a unique Peripheral equilibrium.

Observe that when we compare the star network and the complete network with 3 agents, we see that there is much more clustering in the center for the latter than for the former. Indeed, if we again consider the parameters $\alpha = 6$, $t = 1$ and $\theta = 0.2$, then when $8.86 < c < 21.61$, all the 3 agents live in the center in the complete network while they all reside in the periphery in the star network. This is because there are much more interactions in the *complete* than in the *star* social network because, in the former, everybody interact *directly* with everybody while, in the latter, agents 1 and 2 interact *directly* with the star (agent 1) but only *indirectly* with each other. This is in fact a general result, which is straightforward to prove, which says that the networks that favor more interactions will have more clustering in the center than those that induce less interactions.

4.2.3 Networks with three types of agents

Let us finally consider a network where there are three types of agents so that there are richer equilibrium configurations: individual 1 has 3 links $\{12, 13, 14\}$, individual 2 and 3 have two links each $\{21, 23, 31, 32\}$ and individual 4 has one link $\{41\}$. There are thus three types of agents: type 1 (agent 1), type 2 (agents 2 and 3) and type 3 (agent 4). This network is depicted in Figure 3:

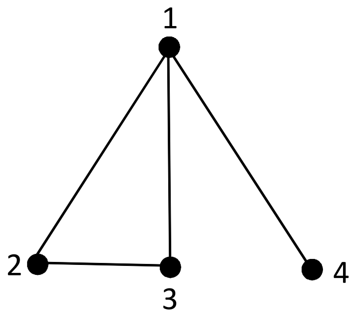

Figure 3: Network with three types of agents

The adjacency matrix is given by:

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

If $\theta < 0.46$ (the largest eigenvalue is 2.17), then

$$\mathbf{M} = [\mathbf{I} - \theta \mathbf{G}]^{-1} = \begin{bmatrix} \frac{1-\theta}{1-\theta-3\theta^2+\theta^3} & \frac{\theta}{1-\theta-3\theta^2+\theta^3} & \frac{\theta}{1-\theta-3\theta^2+\theta^3} & \frac{\theta(1-\theta)}{1-\theta-3\theta^2+\theta^3} \\ \frac{\theta}{1-\theta-3\theta^2+\theta^3} & \frac{1-2\theta^2}{1-4\theta^2-2\theta^3+\theta^4} & \frac{\theta+\theta^2-\theta^3}{1-4\theta^2-2\theta^3+\theta^4} & \frac{\theta^2}{1-\theta-3\theta^2+\theta^3} \\ \frac{\theta}{1-\theta-3\theta^2+\theta^3} & \frac{\theta+\theta^2-\theta^3}{1-4\theta^2-2\theta^3+\theta^4} & \frac{1-2\theta^2}{1-4\theta^2-2\theta^3+\theta^4} & \frac{\theta^2}{1-\theta-3\theta^2+\theta^3} \\ \frac{\theta(1-\theta)}{1-\theta-3\theta^2+\theta^3} & \frac{\theta^2}{1-\theta-3\theta^2+\theta^3} & \frac{\theta^2}{1-\theta-3\theta^2+\theta^3} & \frac{1-\theta-2\theta^2}{1-\theta-3\theta^2+\theta^3} \end{bmatrix} \quad (23)$$

Define

$$\Phi^C(m_{ii}) \equiv t\,(2\alpha - t)\,(m_{ii})^2 + 2t \left(\alpha \sum_{j \in N - \{i\}} m_{ij} \right) m_{ii}$$

where $N = \{1, 2, 3, 4\}$ and the m_{ii}s and m_{ij}s are defined in (23). We have the following result:

Proposition 9 (Locational equilibrium for the network in Figure 3) *Consider the network described in Figure 3 and assume that* $\theta < 0.46$.

(i) If

$$2c < \Phi^{\mathcal{C}}(m_{44})$$

there exists a unique Central equilibrium where all agents live in the center, i.e. $\mathcal{C} = \{1,2,3,4\}$ *and* $\mathcal{P} = \varnothing$.

(ii) If

$$\Phi^{\mathcal{C}}(m_{44}) < 2c < \Phi^{\mathcal{C}}(m_{33}) - 2t^2 m_{34} m_{33}$$

there exists a unique Core-Periphery equilibrium such that $\mathcal{C} = \{1,2,3\}$ *and* $\mathcal{P} = \{4\}$.

(iii) If

$$\Phi^{\mathcal{C}}(m_{33}) - 2t^2 m_{34} m_{33} < 2c < \Phi^{\mathcal{C}}(m_{11}) - 2t^2\left(m_{12} + m_{13} + m_{14}\right) m_{11}$$

there exists a unique Core-Periphery equilibrium such that $\mathcal{C} = \{1\}$ *and* $\mathcal{P} = \{2,3,4\}$.

(iv) If

$$\Phi^{\mathcal{C}}(m_{11}) - 2t^2\left(m_{12} + m_{13} + m_{14}\right) m_{11} < 2c$$

there exists a unique Peripheral equilibrium where all agents live in the periphery, i.e. $\mathcal{C} = \varnothing$ *and* $\mathcal{P} = \{1,2,3,4\}$.

This proposition is a direct application of Proposition 5 and confirms the fact that the number of equilibria is always equal to the number of types of agents plus 1 and that there cannot exist a Core-Periphery equilibrium such that two identical agents live in different parts of the city, e.g. $\mathcal{C} = \{1,2\}$ and $\mathcal{P} = \{3,4\}$. It is also easy to find values for α, t and θ (for $\theta < 0.46$) such that all these conditions hold for a given c.

5 Welfare analysis and subsidy policies

5.1 Exogenous locations

Consider first the case when location choices are exogenous as in Section 3 so that we study welfare policies for a given equilibrium locational configuration.

5.1.1 First-best analysis

We would like to see if the equilibrium outcomes are efficient in terms of *social interactions.* For that, the planner chooses $v_1, ..., v_n$ to maximize total welfare, that is:

$$\max_{v_1,...,v_n} \mathcal{W} = \max_{v_1,...,v_n} \sum_{i=1}^{i=n} U_i(v_i, \mathbf{v_{-i}}, g)$$
$$= \max_{v_1,...,v_n} \left\{ \sum_{i=1}^{i=n} \left[y + \alpha_i v_i - \frac{1}{2} v_i^2 \right] + \theta \sum_{i=1}^{i=n} \sum_{j=1}^{n} g_{ij} v_i v_j \right\}$$

First-order condition gives for each $i = 1, ..., n$:[18]

$$\alpha_i - v_i + \theta \sum_j g_{ij} v_j + \theta \sum_j g_{ji} v_j = 0$$

which implies that (since $g_{ij} = g_{ji}$):[19]

$$v_i^O = \alpha_i + 2\theta \sum_j g_{ij} v_j \tag{24}$$

Using (8), we easily see that:

$$v_i^O = v_i^* + \theta \sum_j g_{ij} v_j \tag{25}$$

where v_i^* is the Nash equilibrium number of visits given in (8). This means that there are *too few visits* at the Nash equilibrium as compared to the social optimum outcome. Equilibrium interaction effort is too low because each agent ignores the positive impact of a visit on the visit choices of others, that is, each agent ignores the positive externality arising from complementarity in visit choices. As a result, the market equilibrium is not efficient and the planner would like to subsidize visits to the interaction center.

5.1.2 Subsidizing social interactions

Letting S_i^O denote the optimal subsidy to per visit, comparison of (24) and (25) implies:

$$S_i^O = \theta \sum_j g_{ij} v_j \tag{26}$$

or in matrix form

$$\mathbf{S}^O = \theta \mathbf{G} \mathbf{v}$$

[18] It is easily checked that there is a unique maximum for each v_i.

[19] The superscript O refers to the "social optimum" outcome while a star refers to the "Nash equilibrium" outcome.

If we add one stage before the visit game is played, the planner will announce the optimal subsidy S_i^O to each agent i such that:

$$\begin{aligned} U_i &= y + \left(\alpha_i + S_i^O\right) v_i - \frac{1}{2} v_i^2 + \theta \sum_j g_{ij} v_i v_j \\ &= y + \alpha_i v_i - \frac{1}{2} v_i^2 + 2\theta \sum_j g_{ij} v_i v_j \end{aligned}$$

By doing so, the planner will restore the first best. Observe that the optimal subsidy is such that

$$\begin{aligned} \mathbf{v}^O &= (\mathbf{I} - \theta \mathbf{G})^{-1} \left(\alpha_i + S_i^O\right) \mathbf{1} \\ &= (\mathbf{I} - 2\theta \mathbf{G})^{-1} \boldsymbol{\alpha} \end{aligned}$$

where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)^{\mathrm{T}}$, which means that

$$\begin{aligned} v_i^O &= \sum_{j=1}^{n} \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]} \left(\alpha_j + S_j^O\right) \\ &= \sum_{j=1}^{n} \sum_{k=0}^{+\infty} 2^k \theta^k g_{ij}^{[k]} \alpha_j \end{aligned}$$

and thus

$$\begin{aligned} U_i^O &= y + \frac{1}{2} \left[b_{\alpha + S^O}(\theta, g)\right]^2 \\ &= y + \frac{1}{2} \left[b_{\alpha}(2\theta, g)\right]^2 \end{aligned}$$

In particular, the optimal subsidy is given by:

$$S_i^O = \theta \sum_j g_{ij} v_j^O = \theta \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{k=0}^{+\infty} 2^k \theta^k g_{ij} g_{jl}^{[k]} \alpha_l \qquad (27)$$

What is interesting here is that the planner will give a larger subsidy to more central agents in the social network. Let us summarize our results by the following proposition.

Proposition 10 (Optimal level of social interactions) *The Nash equilibrium outcome in terms of social interactions is not efficient since there are too few social interactions. If the planner proposes a subsidy $S_i^O = \theta \sum_j g_{ij} v_j$ to each individual i, then the first-best outcome can be restored. In that case, it is optimal for the planner to give higher subsidies to more central agents in the social network.*

5.2 Endogenous locations

As in Section 4, assume now that agents can choose where to locate.

5.2.1 Effort (number of visits) subsidies

Assume that the planner cannot control location but only effort. In that case, in the second stage, she will choose a higher level of effort given by

$$v_i^O = \sum_{j=1}^{n} \sum_{k=0}^{+\infty} 2^k \theta^k g_{ij}^{[k]} \alpha_j \tag{28}$$

and then let the agents choose their location. In other words, the timing is as follows: First, agents choose their location in the city between the center and the periphery and then the government choose effort. The second stage is solved as above and the optimal effort is given by (28). By plugging back this effort into the utility function, we obtain the following equilibrium *optimal* utility of each agent i if she lives in the center ($x = 0$):

$$U_i^O(v_i^*(0, \mathbf{x_{-i}}, g), \mathbf{v_{-i}^*}, g) = y + \frac{1}{2} \left[\sum_{j \in \mathcal{C} - \{i\}} \sum_{k=0}^{+\infty} 2^k \theta^k g_{ij}^{[k]} \alpha + \sum_{j \in \mathcal{P} - \{i\}} \sum_{k=0}^{+\infty} 2^k \theta^k g_{ij}^{[k]} (\alpha - t) + \sum_{k=0}^{+\infty} 2^k \theta^k g_{ii}^{[k]} \alpha \right]^2 - c \tag{29}$$

Similarly, if individual i resides in the periphery ($x = 1$), her equilibrium *optimal* utility is equal to:

$$U_i^O(v_i^*(1, \mathbf{x_{-i}}, g), \mathbf{v_{-i}^*}, g) = y + \frac{1}{2} \left[\sum_{j \in \mathcal{C} - \{i\}} \sum_{k=0}^{+\infty} 2^k \theta^k g_{ij}^{[k]} \alpha + \sum_{j \in \mathcal{P} - \{i\}} \sum_{k=0}^{+\infty} 2^k \theta^k g_{ij}^{[k]} (\alpha - t) + \sum_{k=0}^{+\infty} 2^k \theta^k g_{ii}^{[k]} (\alpha - t) \right]^2 \tag{30}$$

We can characterize the equilibrium location and it is clear that more agents will live in the center compared to the case where they choose themselves their own effort.

Furthermore, if we investigate a *constrained efficient allocation* in which the planner can subsidize interactions (i.e., provide a subsidy of S_i^O per visit by agent i) but cannot directly control location choices, then more agents will live in the center compared to the case without subsidies. Indeed, since all agents devote more effort to interacting with others under the subsidy program (26), the incentives for clustering must be stronger under that allocation than in the Subgame-Perfect Nash equilibrium.

To see that, we can calculate the "incentive" function $\Phi^{kO}(m_{ii})$, which is now given by:

$$\begin{aligned} \Phi^{kO}(m_{ii}) &= 4t\,(2\alpha - t)\,(m_{ii})^2 + 8t \left[\alpha \sum_{j \in \mathcal{C} - \{i\}} m_{ij} + (\alpha - t) \sum_{j \in \mathcal{P} - \{i\}} m_{ij} \right] m_{ii} \\ &= 4\Phi^k(m_{ii}) \end{aligned} \tag{31}$$

where $\Phi^k(m_{ii})$ is given by (42) and is calculated when agents chose both location and effort. Remember that this function determines the location decision of each individual i. Given an equilibrium configuration $k = \mathcal{CP}$, if $\Phi^{kO}(m_{ii}) > 2c$, individual i resides in the center while, if $\Phi^{kO}(m_{ii}) < 2c$, she will reside in the periphery. It is straightforward to write the equivalent of Proposition 5 in the case when the planner chooses effort by only changing the value $\Phi^{C}(m_{ii})$, given in (17), to a new value equal to:

$$\Phi^{CO}(m_{ii}) = 4t\,(2\alpha - t)\,(m_{ii})^2 + 8t\left(\alpha \sum_{j=N-\{i\}} m_{ij}\right) m_{ii} = 4\Phi^{C}(m_{ii})$$

Since $\Phi^{kO}(m_{ii}) > \Phi^{k}(m_{ii})$, it is then straightforward to write the following proposition:

Proposition 11 (Equilibrium versus optimal location choices) *If the planner proposes a per visit subsidy S_i^O to each individual i, then, compared to the Subgame-Perfect Nash equilibrium location choices, more agents live in the center.*

To illustrate this proposition, take again the star network described in Figure 1. We easily obtain:

Proposition 12 (Subsidizing effort in a star-shaped network) *Consider the star-shaped network depicted in Figure 1 and assume that $\theta < 1/\left(2\sqrt{2}\right) = 0.35$ and that the planner chooses (or subsidizes) agents' effort.*

(i) If

$$c < \frac{2t\,(1-\theta)\,(1+\theta)^2\,[2\alpha - (1-\theta)\,t]}{\left(1-2\theta^2\right)^2} \tag{32}$$

there exists a unique Central equilibrium where all agents live in the center, i.e. $\mathcal{C} = \{1, 2, 3\}$ and $\mathcal{P} = \varnothing$.

(ii) If

$$\frac{2t\,(1-\theta)\,(1+\theta)^2\,[2\alpha - (1-\theta)\,t]}{\left(1-2\theta^2\right)^2} < c < \frac{2t\,[2\alpha\,(1+2\theta) - t\,(1+4\theta)]}{\left(1-2\theta^2\right)^2} \tag{33}$$

there exists a unique Core-Periphery equilibrium where the star agent lives in the center while the peripheral agents reside in the periphery, i.e. $\mathcal{C} = \{1\}$ and $\mathcal{P} = \{2, 3\}$.

(iii) If

$$c > \frac{2t\,[2\alpha\,(1+2\theta) - t\,(1+4\theta)]}{\left(1-2\theta^2\right)^2} \tag{34}$$

Let us illustrate Proposition 11. If

$$\frac{t(1-\theta)(1+\theta)^2[2\alpha-(1-\theta)t]}{2(1-2\theta^2)^2} < c < \frac{2t(1-\theta)(1+\theta)^2[2\alpha-(1-\theta)t]}{(1-2\theta^2)^2}$$

then agents 1, 2 and 3 live in the center when the planner chooses effort (Proposition 12) whereas agent 1 lives in the center while agents 2 and 3 reside in the periphery when agents choose effort (Proposition 7). Since the sum of utilities of all agents is the sum of their Katz-Bonacich centralities weighted by α for those who reside in the center and by $\alpha - t$ for those who reside in the periphery, then total welfare is higher when the planner subsidizes effort.

Similarly, if

$$\frac{t[2\alpha(1+2\theta)-t(1+4\theta)]}{2(1-2\theta^2)^2} < c < \frac{2t[2\alpha(1+2\theta)-t(1+4\theta)]}{(1-2\theta^2)^2}$$

then agents 1 lives in the center while agents 2 and 3 reside in the periphery when the planner chooses effort (Proposition 12) whereas all agents live in the periphery when agents choose effort (Proposition 7). The total is also clearly higher when agents' effort is subsidized. As a result, when the planner chooses (or subsidizes) effort, then agents tend to concentrate more in the center of the city than when she don't.

5.2.2 Location subsidies

Let us now consider a model where the planner subsidizes location but not effort. Since there are more interactions when agents live in the center and since interactions increase utility, then the planner could subsidize the location cost c in the center.[20] In other words, she can give a per-cost subsidy σ so that the cost of locating in the center would be $(1-\sigma)c$ instead of c. The timing is now as follows. In the first stage, the planner announces the subsidy to agents locating in the center. In the second stage, agents decide where to locate while, in the last stage, their decide their effort level. As for the subsidy effort, this will clearly generate more clustering in the center but the mechanism is different since, in the latter, the effect is direct while, in the former, it is indirect. In that case, equilibrium efforts will still be determined by (11) while location decisions will be characterized by Proposition 5 where c has to be replaced by $(1-\sigma)c$.

In this model, it is clear that, if the planner wants to reach the first best in terms of location, she will subsidize c so that all agents will live in the center. This maximizes aggregate interactions and thus total welfare. For example, in the case of the star network described in Figure 1, we

[20] It is easily verified that a policy that subsidizes t (marginal transport cost) is equivalent to a policy that subsidizes c. Therefore, we focus our analysis on a subsidy of c.

have shown that if $\alpha = 6$, $t = 1$ and $\theta = 0.2$, then: (*i*) if $c < 7.62$, there exists a unique Central equilibrium; (*ii*) if $7.62 < c < 8.86$, there is a unique Core-Periphery equilibrium; (*iii*) if $c > 8.86$, there exists a unique Peripheral equilibrium. As a result, if, for all agents, $(1 - \sigma)\, c \leq 7.62$, which is equivalent to $\sigma \geq 1 - (7.62/c)$, then the first best is reached and all workers reside in the center. For example, if $c = 20$, then planner needs to subsidize 61.9 percent of the cost of living in the center of all agents. Interestingly, this result depends on the network structure. For the complete network with 3 agents, we have seen that, with exactly the same parameters, $\alpha = 6$, $t = 1$ and $\theta = 0.2$, then: (*i*) if $c < 21.61$, there exists a unique Central equilibrium; (*ii*) if $c > 21.61$, there exists a unique Peripheral equilibrium. In that case, we need to subsidize $\sigma \geq 1 - (21.61)\,/c$ percent of c for all agents to reach the first best. Thus, for the complete network, if $c = 20$, the planner does *not* need to subsidy any worker to reach the first best in efforts since $20 < 21.61$. Using this reasoning and looking at Proposition 5, the optimal subsidy for any network with n agents is given by:

$$\sigma^O > 1 - \frac{\Phi^{\mathcal{C}}(m_{nn})}{2c} \tag{35}$$

where, from (17), we have:

$$\Phi^C(m_{nn}) \equiv t\,(2\alpha - t)\,(m_{nn})^2 + 2t\left(\alpha \sum_{j=N-\{n\}} m_{nj}\right) m_{nn} \tag{36}$$

Observe that equation (35) gives the subsidy for the agent n who has the lowest centrality in the network. Indeed, if the planner gives a $c-$subsidy of $1 - \left[\Phi^{\mathcal{C}}(m_{nn})/2c\right]$ to all agents, the first best will be reached since all individuals will be induced to reside in the center. This is clearly a sufficient condition. The planner could also discriminate between agents and gives a different subsidy to each agent so that the higher is the centrality of an agent in a network, the lower is the subsidy. In that case, the subsidy to be given to each agent i will be equal to:

$$\sigma_i^O > 1 - \frac{\Phi^{\mathcal{C}}(m_{ii})}{2c} \tag{37}$$

for all $i = 1, ..., n$, where $\Phi^{\mathcal{C}}(m_{ii})$ is defined by (17).

Observe also that if $\Phi^{\mathcal{C}}(m_{nn}) > 2c$, meaning that $1 - \frac{\Phi^{\mathcal{C}}(m_{nn})}{2c} < 0$, the condition (35) is always satisfied. This is because, in this case, we do not need to subsidize any worker to obtain a Central equilibrium because $\Phi^{\mathcal{C}}(m_{nn}) > 2c$ is precisely the condition for which a Central equilibrium exists and is unique (see Proposition 5(*i*)). Assuming that, when a worker is indifferent between residing in the center and the periphery, she always chooses to live in the center,[21] then the subsidy (37) can be written as:

$$\sigma_i^O = \max\left\{0, 1 - \frac{\Phi^{\mathcal{C}}(m_{ii})}{2c}\right\} \tag{38}$$

[21] This assumption is made for the sake of the presentation. We could clearly no assume it and, instead, have a

5.2.3 Effort versus location subsidy

Let us now study both the effort and location subsidies. We have seen that if the planner only subsidizes effort, then the optimal subsidy is given by (27), that is

$$S_i^O = \theta \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{k=0}^{+\infty} 2^k \theta^k g_{ij} g_{jl}^{[k]} \alpha_l$$

This optimal subsidy clearly depends on agents' locations. As a result, the first best when both locations and efforts are taken into account should be when the effort subsidy is S_i^O and all agents live in the center. The timing is now as follows: First, the planner announces the location and the effort subsidies. Second, agents choose their location. Third, agents choose efforts.

Denote $\mathbf{M}^O = (\mathbf{I} - 2\theta\mathbf{G})^{-1}$ so that the element of the ith and jth of $\mathbf{M}^O$ is m_{ij}^O.[22] Then, using the same reasoning as above, the location subsidy and the effort subsidy that guarantee that the first best (when both locations and efforts are taken into account) is achieved is determined in the following proposition:

Proposition 13 (First best with effort and location subsidies) *Assume $2\theta\rho(\mathbf{G}) < 1$ and consider any network of n agents. If the location subsidy and the effort subsidy for each agent i are such that*

$$\begin{cases} \sigma_i^O = \max\left\{0, 1 - \frac{2\Phi^{\mathcal{C}}(m_{ii}^O)}{c}\right\} \\ S_i^O = \theta\alpha \sum_{j=1}^{n} \sum_{l=1}^{n} \sum_{k=0}^{+\infty} 2^k \theta^k g_{ij} g_{jl}^{[k]} \end{cases}$$

then all agents live in the center and provide optimal interaction efforts (number of visits) and therefore the first best is achieved.

This proposition implies that, to reach the first best, it is optimal for the the planner to give *higher effort subsidies* but *lower location subsidies* to more central agents in the social network.

If we consider the *star network* of Figure 1, it is readily verified that, if $\theta < 1/\left(2\sqrt{2}\right) = 0.35$,

---

subsidy equal to:

$$\sigma_i^O = \max\left\{0, 1 - \frac{\Phi^{\mathcal{C}}(m_{ii})}{2c} + \varepsilon\right\}$$

where ε is very small.

[22] Remember that, when the per-effort subsidy S_i^O is given to each agent i, she provides an optimal effort v_i^O, which is defined by:

$$\mathbf{v}^O = (\mathbf{I} - 2\theta\mathbf{G})^{-1} \boldsymbol{\alpha} = \mathbf{M}^O \boldsymbol{\alpha}$$

then, if all agents live in the center (i.e. $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$), we have:

$$\mathbf{M}^O \boldsymbol{\alpha} = (\mathbf{I} - 2\theta\mathbf{G})^{-1} \boldsymbol{\alpha} = \frac{\alpha}{(1-8\theta^2)} \begin{pmatrix} 1+4\theta \\ 1+2\theta \\ 1+2\theta \end{pmatrix}$$

so that

$$\begin{pmatrix} v_1^O \\ v_2^O \\ v_3^O \end{pmatrix} = \frac{\alpha}{(1-8\theta^2)} \begin{pmatrix} 1+4\theta \\ 1+2\theta \\ 1+2\theta \end{pmatrix}$$

Since the optimal effort subsidy for each agent i is given by:

$$S_i^O = \theta \sum_j g_{ij} v_j^O$$

or in matrix form

$$\mathbf{S}^O = \theta \mathbf{G} \mathbf{v}^O$$

We have

$$S_1^O = \theta \left(v_2^O + v_3^O \right) = 2\alpha\theta \left(\frac{1+2\theta}{1-8\theta^2} \right)$$

$$S_2^O = S_3^O = \theta \, v_1^O = \alpha\theta \left(\frac{1+4\theta}{1-8\theta^2} \right)$$

Not surprisingly, the planner gives more effort subsidy to more central agents since $S_1^O > S_2^O$.

Now, let us calculate the location subsidy. The matrix $\mathbf{M}^O$ is given by

$$\mathbf{M}^O = (\mathbf{I} - 2\theta\mathbf{G})^{-1} = \frac{1}{(1-8\theta^2)} \begin{pmatrix} 1 & 2\theta & 2\theta \\ 2\theta & 1-4\theta^2 & 4\theta^2 \\ 2\theta & 4\theta^2 & 1-4\theta^2 \end{pmatrix}$$

As a result,

$$\begin{aligned} \Phi^C(m_{11}^O) &= t\,(2\alpha - t)\left(m_{11}^O\right)^2 + 2t\alpha\left(m_{12}^O + m_{13}^O\right) m_{11}^O \\ &= \frac{t\,[2\alpha\,(1+4\theta) - t]}{\left(1-8\theta^2\right)^2} \end{aligned}$$

Thus, the subsidy given to agent 1 is equal to:

$$\sigma_1^O = \max\left\{ 0, 1 - \frac{t\,[2\alpha\,(1+4\theta) - t]}{2c\left(1-8\theta^2\right)^2} \right\}$$

Simarly, we have:

$$\begin{aligned}\Phi^C(m_{22}^O) &= \Phi^C(m_{33}^O) = t\,(2\alpha - t)\left(m_{22}^O\right)^2 + 2t\alpha\left(m_{21}^O + m_{23}^O\right)m_{22}^O \\ &= \frac{t\left(1-4\theta^2\right)(1+2\theta)\left[2\alpha - t\,(1-2\theta)\right]}{\left(1-8\theta^2\right)^2}\end{aligned}$$

Thus, the subsidy $\sigma_2^O = \sigma_3^O$ to give to agents 2 and 3 is:

$$\sigma_2^O = \max\left\{0, 1 - \frac{t\left(1-4\theta^2\right)(1+2\theta)\left[2\alpha - t\,(1-2\theta)\right]}{2c\left(1-8\theta^2\right)^2}\right\}$$

As expected, it is easily verified that $\sigma_1^O < \sigma_2^O$, i.e. the planner gives less location subsidy to more central agents.

Take again $\alpha = 6$, $t = 1$ and $\theta = 0.2$, then the planner needs to give the following subsidies:

$$S_1^O = 4.94 \text{ and } S_2^O = S_3^O = 3.18$$

and

$$\sigma_1^O = \max\left\{0, 1 - \frac{22.28}{c}\right\} \text{ and } \sigma_2^O = \sigma_3^O = \max\left\{0, 1 - \frac{14.50}{c}\right\}$$

to reach the first best. If, for example, $c = 20$, then the planner does not need to subsidize agent 1 but need to subsidy 27.5 percent of the location cost c to live in the center for agents 2 and 3.

Consider now the *complete network* with the three agents residing in the center. It is easily verified that, if $\theta < 0.25$,

$$S_1^O = S_2^O = S_3^O = S^O = \frac{2\alpha\theta\,(1+2\theta)}{\left(1-2\theta-8\theta^2\right)}$$

$$\sigma_1^O = \sigma_2^O = \sigma_3^O = \sigma^O = \max\left\{0, 1 - \frac{t\,(1-2\theta)\left[2\alpha\,(1+2\theta) - t\,(1-2\theta)\right]}{2c\left(1-2\theta-8\theta^2\right)^2}\right\}$$

If we take the same parameter values, $\alpha = 6$, $t = 1$ and $\theta = 0.2$, then $S^O = 12$ and

$$\sigma^O = \max\left\{0, 1 - \frac{17.36}{c}\right\}$$

If we compare the two networks, it easily verified that the planner needs to subsidize much more the social effort of all agents in the complete network (there are more network externalities in the complete network compared to the star network) while, for a given c and for location subsidies, she needs to subsidize less agent 1 but more agents 2 and 3 in the star network. In terms of network design, this means that the planner would not always like to choose the *complete network*, even though it is the network that generates most interactions among all possible networks. The optimal network will clearly depend on parameters values and will be, in general, difficult to determine.

6 Spatial mismatch and policy issues

There is an important literature in urban economics showing that, in the United States, distance to jobs is harmful to workers, in particular, black workers. This is known as the "spatial mismatch hypothesis". Indeed, first formulated by Kain (1968), the spatial mismatch hypothesis states that, residing in urban segregated areas distant from and poorly connected to major centres of employment growth, black workers face strong geographic barriers to finding and keeping well-paid jobs. In the US context, where jobs have been decentralized and blacks have stayed in the central parts of cities, the main conclusion of the spatial mismatch hypothesis is that distance to jobs is the main cause of their high unemployment rates. Since Kain's study, hundreds of others have been conducted trying to test the spatial mismatch hypothesis (see, in particular, the literature surveys by Ihlanfeldt and Sjoquist, 1998; Ihlanfeldt, 2006; Gobillon et al., 2007; Zenou, 2008). The usual approach is to relate a measure of labor-market outcomes, typically employment or earnings, to another measure of job access, typically some index that captures the distance between residences and centres of employment. The general conclusions are: (*i*) poor job access indeed worsens labor-market outcomes, (*iii*) black and Hispanic workers have worse access to jobs than white workers, and (*iii*) racial differences in job access can explain between one-third and one-half of racial differences in employment. Interpret the model in terms of black and white workers.

Our model can shed new light on the "spatial mismatch hypothesis" debate by putting forward the importance of, not only the geographical space (distance to jobs), but also the social space in explaining the adverse labor-market outcomes of black workers.

Let us interpret our model in the following way. There are two locations, a center, where all jobs are located and all interactions take place, and a periphery.[23] Here an interaction between two individuals means that they exchange job information with each other and thus each visit to the center implies a job-information exchange with someone. As above, v_i is the number of visits that individual i makes to the center in order to obtain information about jobs and each visit results in one interaction. We do not explicitly model the labor market. We just assume that the higher is the number and quality[24] of interactions, the higher is the quality of job information and the higher is the probability of being employed.[25] In other words, each time a person goes to the center, she interacts with someone and obtains a piece of job information, which is proportional to

[23] Observe that, in the context of real-world cities, the center does not necessarily mean the physical center of the city but the place where jobs and interactions take place.

[24] In equilibrium, more central workers provide higher quality job information because they interact more with others than less central workers.

[25] This is the basic idea behind most network models of the labor market such as Calvó-Armengol (2004), Calvó-Armengol and Jackson (2004), Calvó-Armengol and Zenou (2005) and Ioannides and Soetevent (2006).

the network centrality of the individual she meets. This leads to a positive relationship between v_i, the individual number of visits to the center, and e_i, the employment rate of each individual i. Underlying this idea is some form of information imperfection in which networks serve at least partially to mitigate these imperfections.[26]

There are two types of workers: black and white individuals. The only difference between black and white workers is their position in the network. We assume that whites have a more central position (in terms of Katz-Bonacich centrality) in the network than blacks. This captures the idea of the "old boy network" where whites grew up together, went through school together, socialized together during adolescence and early adulthood, and entered the labor force together (Wial, 1991).[27] There is strong evidence that indicates that labor-market networks are partly race based, operating more strongly within than across races (Ioannides and Loury, 2004; Hellerstein et al., 2011) and that the social network of black workers is of lower quality than that of whites (Frijters et. al., 2005; Fernandez and Fernandez-Mateo, 2006; Battu et al., 2011).

To understand the interpretation of the current model, consider the network with three types of agents displayed in Figure 3 and assume that individuals 1, 2 and 3 are white workers while individual 4 is a black worker. We have shown in Section 4.2.3 that if

$$\Phi^{\mathcal{C}}(m_{44}) < 2c < \Phi^{\mathcal{C}}(m_{33}) - 2t^2 m_{34} m_{33}$$

there exists a unique Core-Periphery equilibrium such that $\mathcal{C} = \{1, 2, 3\}$ and $\mathcal{P} = \{4\}$. In the labor-market interpretation of this model, white workers will experience a higher employment rate than the black worker because they will have much more information about jobs. In other words, the white workers, especially individual 1, will interact much more with other workers than the black worker because the latter will visit less often the center and will gather little information about jobs. In this model, it is assumed that any worker i can give information about job but the quality of information she gives is proportional her v_i, the number of visits she makes to the center or equivalently the number of interactions she has with others. As stated above, the employment probability of each worker is then proportional to the information she has gathered in equilibrium.

In this interpretation of the model, we have shown that black workers make less visits to the center (Proposition 1) and thus interact less with other workers in the network, in particular, with very central agents than whites. We have also shown that black workers will choose to locate further away from jobs than white workers (Proposition 5) precisely because they interact less with

[26] See Ioannides and Loury (2004) and Topa (2011) for a review of the evidence on labor market networks.

[27] Calvó-Armengol and Jackson (2004) show that an equilibrium with a clustering of workers with the same status is likely to emerge since, in the long run (i.e. steady state), employed workers tend to be friends with employed workers. In this model, if because of some initial condition some black workers are unemployed, then in steady-state they will still be unemployed because both their strong and weak ties will also be unemployed.

central workers. At the extreme, we could have an equilibrium where all white workers live in the center while all black workers reside in the periphery (as in the example above where $\mathcal{C} = \{1, 2, 3\}$ and $\mathcal{P} = \{4\}$). This would imply that whites will interact with others much more than blacks and that whites will interact more with whites (since they will have a very high effort v_i in equilibrium) than with blacks. Blacks will just interact less and thus will have much less information about jobs. This will clearly have dramatic consequences in the labor market and will explain why black workers experience a lower employment rate than white workers. Indeed, less central agents in the network (i.e. black workers who do not have an old-boy network) will reside further away from jobs (i.e. in the periphery) than more central agents (whites) and thus will have adverse labor-market outcomes. In other words, the *lack of good job contacts would be here a structural consequence of the social isolation of inner-city neighborhoods*.[28] Importantly, the *causality goes from the social space to the geographical space* so that it is the *social mismatch* (i.e. their "bad" location in the social network) of black workers that leads to their *spatial mismatch* (i.e. their "bad" location in the geographical space). Observe that the network structure is crucial in our model. For example, in a complete network (or any regular network), there will be no effect since black and white workers will be totally identical. As a result, the more the network is heterogenous and asymmetric, the worse are the labor-market outcomes for black workers.[29]

Interestingly, Zenou (2013) has developed a model where the causality goes the other way around. In his model, which is quite different since the labor market is explicitly model but the social network is just captured by dyads, it is the spatial mismatch of black workers (due to housing discrimination) that leads to their social mismatch (i.e. less interaction with white weak ties) and thus their adverse labor-market outcomes.

For the policy implications of each model, it is crucial to know the sense of causality. If, as in Zenou (2013), it is the geographical space that causes the social mismatch of black workers, then the policies should focus on workers' geographical location, as in the spatial mismatch literature. In that case, *neighborhood regeneration policies* would be the right tool to use. Such policies have been implemented in the US and in Europe through the enterprise zone programs and the empowerment zone programs (e.g. Papke, 1994; Bondonio and Greenbaum, 2007; Ham et al., 2011; Busso et al.,

[28] Observe that we interpret here the "periphery" location of our model as an "inner-city neighborhood" because what characterizes the latter is not its location in the city but the fact that it is disconnected (or badly connected) from job centers.

[29] There is evidence showing that, though networks are a popular method of finding a job for the ethnic minorities, they are not necessarily the most effective in terms of gaining employment (Frijters et al., 2005; Fernandez and Fernandez-Mateo, 2006; Battu et al., 2011). Our labor-network model could explain this fact since minority workers tend to have less central positions in the network and tend to interact less with white workers (who have more central positions) because they live further away from interaction centers.

2012). The enterprise zone policy consists in designating a specific urban (or rural) area, which is depressed, and targeting it for economic development through government-provided subsidies to labor and capital. The aim of the empowerment zone program is to revitalize distressed urban communities and it represents a nexus between social welfare policy and economic development efforts. By implementing these types of policies, one brings jobs to people and thus facilitates the flows of job information in depressed neighborhoods. Another way of reducing the spatial mismatch of black workers would be to implement a transportation policy that subsidizes workers' commuting costs (Pugh, 1998). In the United States, a number of states and counties have used welfare block grants and other federal funds to support urban transportation services for welfare recipients. For example, programs helping job takers (especially African Americans) obtain a used car – a secured loan for purchase, a leasing scheme, a revolving credit arrangement – may offer real promise and help low-skill workers obtain a job by commuting to the center where jobs are located.

If, on the contrary, as in the current model, it is the social space that causes the spatial mismatch of black workers, then the policies should focus on workers' social isolation. Policies that promote social integration and thus increase the interracial interactions between black and white workers would also have positive effects on the labor-market outcomes of minority workers. Such policies, like the Moving to Opportunity (MTO) programs (Katz et al., 2001; Rosenbaum and Harris, 2001; Kling et al., 2005), have been implemented in the United States. By giving housing assistance to low-income families, the MTO programs help them relocate to better and richer neighborhoods. For example, Rosenbaum and Harris (2001) show that: "After moving to their new neighborhoods, the Section 8 respondents (treated group) were far more likely to be actively participating in the labor force (i.e. working or looking for a job), while for MTO respondents, a statistically significant increase is evident only for employment per se." Another way of reducing the unemployment rate of minorities in the context of our model is to observe that *institutional connections* can be engineered to create connections between job seekers and employers in ways that parallel social network processes. For example, scholars like Granovetter (1979) and Wilson (1996) have called for poverty reduction programs to "create connections" between employers and poor and disadvantaged job seekers.

This is ultimately an empirical question of causality – whether people that are central in the network move to the city, or do people that are less connected move to the city and then become more central. Such an empirical test is crucial but one would need either a natural experiment with an exogenous shock or convincing instruments to break the sense of causality. In the labor-market interpretation, the key issues is whether black workers first choose to live in geographically isolated neighborhoods (or are forced to live there because of housing discrimination) and then become isolated in the social space because of the lack of contacts with white workers, or do black workers mainly prefer to interact with other black individuals and as a consequence locate in areas where few whites live, which are isolated from jobs. In any case, we believe that the social and the geographical space are intimately related and policies should take into account both of them if they want to be successful.

Finally, using the current model, one could also examine the role of the network structure of the labor-market outcomes of black workers. We have seen that labor-market networks are partly race based, operating more strongly within than across races (Ioannides and Loury, 2004; Hellerstein et al., 2011) and that more regular social networks lead to less geographical segregation and better labor-market outcomes for black workers. Are social networks regular or very asymmetric? How do black and white workers interact in a network? Our welfare analysis developed in Section 5 will also be useful. In Section 5.2.3, we have seen that there is a trade off between subsidizing effort (or the number of visits to the center) and location (or the transportation cost t). In terms of the policies discussed above, subsidizing effort in the social space or location in the geographical space will be equivalent to the MTO programs in the sense that it favors interactions between people by locating them close to each other.

7 Concluding remarks

This paper provides what we believe to be the first analysis of the interaction between position in a social network and position in a geographic space, or between social and physical distance. We have developed a model in which agents who are more central in a social network, or are located closer to an interaction center, choose higher levels of interaction effort in equilibrium. As a result, the level of interactivity in the economy as a whole rises with density of links in the social network and with the degree to which agents are clustered in physical space. When agents can choose geographic locations, there is a tendency for those who are more central in the social network to locate closer to the interaction center.

There are many potential extensions and applications of the work described here. First, we have assumed that all interactions occur at a single, exogenous interaction center. In reality, interactions in cities occur at many sites, and whether a site becomes a focal point for interactions is of course endogenous. As in all models of complementarity, there is an interesting coordination problem in the endogenous determination of the location of an interaction center in this model.

Second, we have developed a model where efforts are strategic complements, i.e. $\theta > 0$. It would be interesting to assume instead that efforts are strategic substitutes, i.e. $\theta < 0$ (as in Bramoullé and Kranton, 2007), or, following Bramoullé et al. (2012), extend parts of the analysis to allow for a larger set of parameters (i.e. the entire range of θ). The analysis would certainly be much more complicated since some agents will free ride on others and provide zero effort. The interpretation of the results will also be different.

Finally, in the present paper, we focus on fixed networks. Where we live crucially determines whom we interact with and indeed a main attraction of big cities is that they provide more opportunities to meet new people. So network endogeneity could be an important feature of the interplay between social networks and geographic locations. As a result, it would be interesting to extend the model to have an endogenous network formation. We believe that it will be very difficult to obtain clear-cut results since it is well-known that there are combinatorial (coordination) equilibrium multiplicity in standard network-formation models (see e.g. Jackson and Wolinsky, 1996). Here agents also choose effort and location so coordination problems will even be more severe. One possible way out could be to use the Cabrales et al. (2011) approach where network formation is not the result of an earmarked socialization process so that there is a generic socialization effort. Another way out would be to use the dynamic model of König et al. (2012) where, at each period of time, one agent is chosen at random to form a link with others. It is clear that endogeneizing the structure of the network would be a very difficult task given that agents also choose effort and location. We leave this at a future research area.

References

[1] Akerlof, G. (1997), "Social distance and social decisions," *Econometrica* 65, 1005-1027.

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

[3] Ballester, C., Calvó-Armengol, A. and Y. Zenou (2006), "Who's who in networks. Wanted: The key player," *Econometrica* 74, 1403-1417.

[4] Battu, H., Seaman, P. and Y. Zenou (2011), "Job contact networks and the ethnic minorities," *Labour Economics* 18, 48-56.

[5] 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.

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

[7] Benabou, R. (1993), "Workings of a city: Location, education, and production," *Quarterly Journal of Economics* 108, 619-652.

[8] Benabou, R. (1996), "Equity and efficiency in human capital investment: The local connection," *Review of Economic Studies* 63, 237-264.

[9] Bonacich, P. (1987), "Power and centrality: A family of measures," *American Journal of Sociology* 92, 1170-1182.

[10] Bondonio, D. and R.T. Greenbaum (2007), "Do local tax incentives affect economic growth? What mean impact miss in the analysis of enterprise zone policies," *Regional Science and Urban Economics* 37, 121-136.

[11] Bramoullé, Y. and R. Kranton (2007), "Public goods in networks," *Journal of Economic Theory* 135, 478-494.

[12] Bramoullé, Y., Kranton, R. and M. D'Amours (2012), "Strategic Interaction and Networks," Unpublished manuscript, Duke Univerity.

[13] Brueckner, J.K. (2006), "Friendship networks," *Journal of Regional Science* 46, 847-865.

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

[15] Brueckner, J.K., Thisse, J-F. and Y. Zenou (2002), "Local labor markets, job matching and urban location," *International Economic Review* 43, 155-171.

[16] Busso, M., Gregory, J. and P. Kline (2012), "Assessing the incidence and efficiency of a prominent place based policy," *American Economic Review*, forthcoming.

[17] Cabrales, A., Calvó-Armengol, A. and Y. Zenou (2011), "Social interactions and spillovers," *Games and Economic Behavior* 72, 339-360.

[18] Calvó-Armengol, A. (2004), "Job contact networks", *Journal of Economic Theory* 115, 191-206.

[19] Calvó-Armengol, A. and M.O. Jackson (2004), "The effects of social networks on employment and inequality," *American Economic Review* 94, 426-454.

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

[21] Calvó-Armengol, A. and Y. Zenou (2005) "Job matching, social network and word-of-mouth communication", *Journal of Urban Economics* 57, 500-522.

[22] De Bartolome, C. (1990), "Equilibrium and inefficiency in a community model with peer group effects," *Journal of Political Economy* 98, 110-33.

[23] Durlauf, S.N. (1996), "A theory of persistent income inequality," *Journal of Economic Growth* 1, 75-93.

[24] Epple, D. and R.E. Romano (1998), "Competition between private and public schools, vouchers, and peer-group effects," *American Economic Review* 88, 33-62.

[25] Fernandez, R.M. and I. Fernandez-Mateo (2006), "Networks, race, and hiring," *American Sociological Review* 71, 42-71.

[26] Frijters, P., Shields, M.A. and S. Wheatley Price (2005), "Job search methods and their success: A comparison of immigrants and natives in the UK," *Economic Journal* 115, F359-F376.

[27] Fujita, M. and J.-F. Thisse (2002), *Economics of Agglomeration. Cities, Industrial Location, and Regional Growth*, Cambridge: Cambridge University Press.

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

[29] Gaspar, J. and E.L. Glaeser (1997), "Information technology and the future of cities," *Journal of Urban Economics* 43, 136-156.

[30] Gobillon, L., Selod, H. and Y. Zenou (2007), "The mechanisms of spatial mismatch," *Urban Studies* 44, 2401-2427.

[31] Ghiglino, C. and A. Nocco (2012), "When Veblen meets Krugman," Unpublished manuscript, University of Essex.

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

[33] Glaeser, E.L. Sacerdote, B., and J. Scheinkman (1996), "Crime and social interactions," *Quarterly Journal of Economics* 111, 508-548.

[34] Glaeser, E.L., Sacerdote, Bruce I. and J.A. Scheinkman (2003), "The social multiplier," *Journal of the European Economic Association* 1, 345-353.

[35] Glaeser, E.L. and J.A. Scheinkman (2001), "Measuring social interactions," In: S. Durlauf and P. Young (Eds.), *Social Dynamics*, Cambridge, MA: MIT Press, pp. 83-131.

[36] Goyal, S. (2007), *Connections: An Introduction to the Economics of Networks*, Princeton: Princeton University Press.

[37] Granovetter, M.S. (1979), "Placement as brokerage: Information problems in the labor market for rehabilitated workers," In: D. Vandergoot and J.D. Worrall (Eds.), *Placement in Rehabilitation: A Career Development Perspective*, Baltimore, MD: University Park Press, pp. 83-101.

[38] Grauwin, S., Goffette-Nagot, F. and P. Jensen (2012), "Dynamic models of residential segregation: An analytical solution," *Journal of Public Economics* 96, 124-141.

[39] Ham, J.C., Swenson, C., Imrohoroglu, A. and H. Song (2011), "Government programs can improve local labor markets: Evidence from State Enterprise Zones, Federal Empowerment Zones and Federal Enterprise Community," *Journal of Public Economics* 95, 779-797.

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

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

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

[43] Ihlanfeldt, K.R. (2006), "A primer on spatial mismatch within urban labor markets," In: R. Arnott and D. McMillen (Eds.), *A Companion to Urban Economics*, Boston: Blackwell Publishing, pp. 404-417.

[44] Ihlanfeldt, K. R. and D. Sjoquist (1998), "The spatial mismatch hypothesis: A review of recent studies and their implications for welfare reform," *Housing Policy Debate* 9, 849-892.

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

[46] Ioannides, Y.M. and D.L. Loury (2004), "Job information networks, neighborhood effects, and inequality," *Journal of Economic Literature* 42, 1056-1093.

[47] Ioannides, Y.M. and A.R. Soetevent (2006), "Wages and employment in a random social network with arbitrary degree distribution," *American Economic Review Papers and Proceedings* 96, 270-274.

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

[49] Jackson, M.O and A. Wolinsky (1996), "A strategic model of social and economic networks," *Journal of Economic Theory* 71, 44-74.

[50] Jackson, M.O. and Y. Zenou (2013), "Games on networks,“ In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier, forthcoming.

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

[52] Jaffe, A.B., Trajtenberg, M. and R. Henderson (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations," *Quarterly Journal of Economics* 108, 577-598.

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

[54] Kain, J. (1968), "Housing segregation, negro employment, and metropolitan decentralization," *Quarterly Journal of Economics* 82, 175-197.

[55] Katz, L. (1953), "A new status index derived from sociometric analysis," *Psychometrika* 18, 39-43.

[56] Katz, L.F., Kling, J.R. and J.B. Liebman (2001), "Moving to opportunity in Boston: Early results of a randomized mobility experiment," *Quarterly Journal of Economics* 116, 607-654.

[57] Kling, J.R., Ludwig, J. and L.F. Katz (2005), "Neighborhood effects on crime for female and male youth: Evidence from a randomized housing voucher experiment," *Quarterly Journal of Economics* 120, 87-130.

[58] König, M., Tessone, C. and Y. Zenou (2012), "Nestedness in networks: A theoretical model and some applications", Unpublished manuscript, Stockholm University.

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

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

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

[62] Monderer, D. and L.S. Shapley (1996), "Potential games," *Games and Economic Behavior* 14, 124-143.

[63] Moretti, E. (2004), "Workers' education, spillovers, and productivity: Evidence from plant-level production functions," *American Economic Review* 94, 656-690.

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

[65] Mossay P. and P.M. Picard (2012), "Spatial segregation and urban structure," Unpublished manuscript, University of Luxemburg.

[66] Papke, L. (1994), "Tax policy and urban development: evidence from the Indiana enterprise zone program," *Journal of Public Economics* 54, 37-49.

[67] Patacchini, E. and Y. Zenou (2012), "Ethnic networks and employment outcomes," *Regional Science and Urban Economics* 42, 938-949.

[68] Picard, P.M. and T. Tabuchi (2010), "Self-organized agglomerations and transport costs," *Economic Theory* 42, 565-589.

[69] Pugh, M. (1998), "Barriers to work: the spatial divide between jobs and welfare recipients in metropolitan areas", Washington, DC: The Brookings Institution.

[70] Putnam, R.D. (1993), *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon and Schuster.

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

[72] Romer, P.M. (1990), "Endogenous technological change," *Journal of Political Economy* 98, S71-S102.

[73] Rosenbaum, E. and L.E. Harris (2001), "Residential mobility and opportunities: Early impacts of the Moving to Opportunity demonstration program in Chicago," *Housing Policy Debate* 12, 321-346.

[74] Rosenthal, S.S. and W.C. Strange (2003), "Geography, industrial organization, and agglomeration," *Review of Economics and Statistics* 85, 377-393.

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

[76] Schelling, T.C. (1971), "Dynamic models of segregation," *Journal of Mathematical Sociology* 1, 143-186.

[77] Topa, G. (2001), "Social interactions, local spillovers and unemployment," *Review of Economic Studies* 68, 261-295.

[78] Topa, G. (2011), "Labor markets and referrals," In: J. Benhabib, A. Bisin and M.O. Jackson (Eds.), *Handbook of Social Economics, Vol. 1B*, Amsterdam: Elsevier Science, pp. 1193-1221.

[79] Verdier, T. and Y. Zenou (2004), "Racial beliefs, location and the causes of crime," *International Economic Review* 45, 727-756.

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

[81] Wial, H. (1991), "Getting a good job: Mobility in a segmented labor market," *Industrial Relations* 30, 396-416.

[82] Wilson, W.J. (1996), *When Work Disappears: The World of the New Urban Poor*, New York: Knopf.

[83] Zenou, Y. (2008), "The spatial mismatch hypothesis," In: L. Blume and S. Durlauf (Eds.), *The New Palgrave, A Dictionary of Economics, Second Edition*, London: MacMillan Press.

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

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

[86] Zhang, J. (2004), "Residential segregation in an all-integrationist world," *Journal of Economic Behavior and Organization* 54, 533-550.

Appendix

Proof of Proposition 2. Observe that this game is a potential game (as defined by Monderer and Shapley, 1996)[30] with potential function:[31]

$$\begin{aligned} P(\mathbf{v},g,\theta) &= \sum_{i=1}^{n} u_i(\mathbf{v},\mathbf{g}) - \frac{\theta}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}v_iv_j \\ &= \sum_{i=1}^{n} \alpha\, v_i - \frac{1}{2}\sum_{i=1}^{n} v_i^2 + \frac{\theta}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}v_iv_j, \end{aligned}$$

or in matrix form:

$$\begin{aligned} P(\mathbf{v},g,\theta) &= \alpha\mathbf{v}^\top\mathbf{1}-\frac{1}{2}\mathbf{v}^\top\mathbf{v} + \mathbf{v}^\top\frac{\theta}{2}\mathbf{G}\mathbf{v} \\ &= \alpha\mathbf{v}^\top\mathbf{1}-\frac{1}{2}\mathbf{v}^\top\left(\mathbf{I}-\theta\mathbf{G}\right)\mathbf{v}. \end{aligned}$$

It is well-known (see e.g., Monderer and Shapley, 1996) that solutions of the program $\max_{\mathbf{v}} P(\mathbf{v},g,\theta)$ are a subset of the set of Nash equilibria. This program has a unique interior solution if the potential function $P(\mathbf{v},g,\theta)$ is strictly concave on the relevant domain. The Hessian matrix of $P(\mathbf{v},g,\theta)$ is easily computed to be $-\left(\mathbf{I}-\theta\mathbf{G}\right)$. The matrix $\mathbf{I}-\theta\mathbf{G}$ is positive definite if for all non-zero $\mathbf{v}$

$$\mathbf{v}^\top\left(\mathbf{I}-\theta\mathbf{G}\right)\mathbf{v} > 0 \Leftrightarrow \theta < \left(\frac{\mathbf{v}^\top\mathbf{G}\mathbf{v}}{\mathbf{v}^\top\mathbf{v}}\right)^{-1}.$$

By the Rayleigh-Ritz theorem, we have $\rho(\mathbf{G}) = \sup_{\mathbf{v}\neq\mathbf{0}}\left(\frac{\mathbf{v}^\top\mathbf{G}\mathbf{v}}{\mathbf{v}^\top\mathbf{v}}\right)$. Thus a necessary and sufficient condition for having a strict concave potential is that $\theta\rho(\mathbf{G}) < 1$, as stated in the Proposition. ∎

Proof of Proposition 4. Define $\mathcal{C}$ as the set of all *central* agents (i.e. all individuals who live in the center) and $\mathcal{P}$ as the set of all *peripheral* agents (i.e. all individuals who live in the periphery). Observe that $\mathcal{C}-\{i\}$ and $\mathcal{P}-\{i\}$ denotes respectively the set of all *central* agents but i and the set of all *peripheral* agents but i. The condition for which, a given i prefers to live in the

[30] A game is a potential game if there is a function $P : X \rightarrow \mathbb{R}$ such that, for each $i \in N$, for each $x_{-i} \in X_{-i}$, and for each $x_i, z_i \in X_i$,

$$u_i(x_i,x_{-i}) - u_i(z_i,x_{-i}) = P(x_i,x_{-i}) - P(z_i,x_{-i})$$

[31] Here the potential $P(\mathbf{v},g,\theta)$ is constructed by taking the sum of all utilities, a sum that is corrected by a term which takes into account the network externalities exerted by each player i.

center is:

$$y + \frac{1}{2}\left[\sum_{j\in\mathcal{C}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}\alpha + \sum_{j\in\mathcal{P}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}(\alpha - t) + \sum_{k=0}^{+\infty}\theta^k g_{ii}^{[k]}\alpha\right]^2 - c \tag{39}$$

$$\geq y + \frac{1}{2}\left[\sum_{j\in\mathcal{C}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}\alpha + \sum_{j\in\mathcal{P}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]}(\alpha - t) + \sum_{k=0}^{+\infty}\theta^k g_{ii}^{[k]}(\alpha - t)\right]^2$$

where the expression on the left-hand side of the inequality is the utility of i if she resides in the center while the expression on the right-hand side of the inequality is the utility of i if she resides in the periphery. Indeed, given $\mathcal{P}$ and $\mathcal{C}$, an individual i's utility is her income y plus her Katz-Bonacich centrality squared minus c if she lives in the center ($x = 0$) and minus zero if she resides in the periphery ($x = 1$). The difficulty here is to calculate the Katz-Bonacich centrality of individual i. If she decides to reside in the center (resp. the periphery), then the total number of people living in the center is composed of all individuals $j \neq i$ living in the center, i.e. the cardinal of the set $\mathcal{C} - \{i\}$, plus individual i (resp. without individual i) while the total number of people living in the periphery is composed of all individuals $j \neq i$ living in the periphery, i.e. the cardinal of the set $\mathcal{P} - \{i\}$, (resp. plus individual i). To calculate the Katz-Bonacich centrality of all these agents, we proceed as follows. For all individuals $j \neq i$ living in the center (resp. in the periphery), we calculate all the paths that are not self-looped (i.e. the off diagonals of the matrices $\mathbf{G} = [g_{ij}^{[1]}] \equiv [g_{ij}]$, $\mathbf{G}^2 = [g_{ij}^{[2]}]$, etc.) and we weigh them by α (resp. by $\alpha - t$). For individual i, we take the self-loop paths (i.e. the diagonals of the matrices $\mathbf{G} = [g_{ij}^{[1]}] \equiv [g_{ij}]$, $\mathbf{G}^2 = [g_{ij}^{[2]}]$, etc.) and we weigh them by α if she lives in the center and by $\alpha - t$ if she resides in the periphery.

Denote by

$$\begin{aligned} A(\mathcal{P},\mathcal{C}) &= \alpha\sum_{k=0}^{+\infty}\theta^k g_{ii}^{[k]} + \alpha\sum_{j\in\mathcal{C}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]} + (\alpha - t)\sum_{j\in\mathcal{P}-\{i\}}\sum_{k=0}^{+\infty}\theta^k g_{ij}^{[k]} \\ &= \alpha m_{ii} + \alpha\sum_{j\in\mathcal{C}-\{i\}} m_{ij} + (\alpha - t)\sum_{j\in\mathcal{P}-\{i\}} m_{ij} \end{aligned}$$

The inequality (39) is then equivalent to:

$$[A(\mathcal{P},\mathcal{C})]^2 - 2c \geq \left[A(\mathcal{P},\mathcal{C}) - t\sum_{k=0}^{+\infty}\theta^k g_{ii}^{[k]}\right]^2$$

Since $A(\mathcal{P},\mathcal{C}) = \alpha m_{ii} + \alpha\sum_{j\in\mathcal{C}-\{i\}} m_{ij} + (\alpha - t)\sum_{j\in\mathcal{P}-\{i\}} m_{ij}$ and since $m_{ii} = \sum_{k=0}^{+\infty}\theta^k g_{ii}^{[k]}$, this is equivalent to:

$$\Phi(m_{ii}) \equiv t(2\alpha - t)(m_{ii})^2 + 2t\left(\alpha\sum_{j\in\mathcal{C}-\{i\}} m_{ij} + (\alpha - t)\sum_{j\in\mathcal{P}-\{i\}} m_{ij}\right) m_{ii} \geq 2c \tag{40}$$

To summarize, taking as given $\mathcal{P}$ and $\mathcal{C}$, any individual i for which $\Phi(m_{ii}) \geq 2c$ will reside in the center while any individual i for which $\Phi(m_{ii}) \leq 2c$ will reside in the periphery. Let us study $\Phi(m_{ii}) - 2c$, which is a second-degree equation. We know that $\alpha > t$, which implies that $2\alpha > t$. This means that $\Phi(m_{ii}) - 2c$ is a convex function with $\Phi(0) - 2c = -2c$. The discriminant of $\Phi(m_{ii}) - 2c$ is given by:

$$\Delta = 4t^2 \left[\left(\alpha \sum_{j \in \mathcal{C}-\{i\}} m_{ij} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} m_{ij} \right)^2 + \frac{2c}{t} (2\alpha - t) \right]$$

There are two roots that are given by:

$$m_{ii}^{(1)} = \frac{-2t \left(\alpha \sum_{j \in \mathcal{C}-\{i\}} m_{ij} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} m_{ij} \right) - \sqrt{\Delta}}{2t (2\alpha - t)}$$

$$m_{ii}^{(2)} = \frac{-2t \left(\alpha \sum_{j \in \mathcal{C}-\{i\}} m_{ij} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} m_{ij} \right) + \sqrt{\Delta}}{2t (2\alpha - t)}$$

It is clear that $m_{ii}^{(1)} < 0$ and $m_{ii}^{(2)} > 0$. Figure 2 describes the equilibrium configuration. All individuals i with a $m_{ii} > m_{ii}^{(2)}$ will reside in the center of the city (i.e. $x = 0$) while all individuals i with a $m_{ii} < m_{ii}^{(2)}$ will live at the periphery of the city (i.e. $x = 1$). ∎

Proof of Proposition 5.

First, remember that $\Phi^{\mathcal{C}}(m_{nn})$ is defined by (17), i.e.

$$\Phi^{C}(m_{ii}) \equiv t (2\alpha - t) (m_{ii})^2 + 2t \left(\alpha \sum_{j \in N-\{i\}} m_{ij} \right) m_{ii} \tag{41}$$

since $\mathcal{C} = N$. From the proof of Proposition 4, remember also that $\Phi(m_{ii})$ is defined by (40), that is:

$$\Phi^{k}(m_{ii}) \equiv t (2\alpha - t) (m_{ii})^2 + 2t \left(\alpha \sum_{j \in \mathcal{C}-\{i\}} m_{ij} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} m_{ij} \right) m_{ii} \tag{42}$$

We have added the superscript k to show which type of equilibrium we are considering. For example, $k = \mathcal{C}$ when $\mathcal{C} = N$ and $\mathcal{P} = \varnothing$ while $k = \mathcal{P}$ when $\mathcal{C} = \varnothing$ and $\mathcal{P} = N$.

Let us first show that, in any equilibrium, agents of the same type (i.e. with the same Katz-Bonacich centrality) have to reside in the same part of the city, i.e. either at $x = 0$ or at $x = 1$. Assume, on the contrary, that two agents with the same centrality reside in different parts of the city, i.e., agent i resides at $x = 0$ (center) while agent i' resides at $x = 1$ (periphery) with

$m_{ii} = m_{i'i'}$ and $m_{ij} = m_{i'j}$, $\forall j$. For this to be an equilibrium, it has to be that $\Phi^{\mathcal{CP}}(m_{ii}) > 2c$ and $\Phi^{\mathcal{CP}}(m_{i'i'}) < 2c$, which implies that $\Phi^{\mathcal{CP}}(m_{ii}) > \Phi^{\mathcal{CP}}(m_{i'i'})$. Using (40), this inequality is equivalent to:

$$\begin{aligned} & t\,(2\alpha - t)\,(m_{ii})^2 + 2t\left(\alpha \sum_{j\in\mathcal{C}-\{i\}} m_{ij} + (\alpha - t)\sum_{j\in\mathcal{P}-\{i\}} m_{ij}\right) m_{ii} \\ > \; & t\,(2\alpha - t)\,(m_{i'i'})^2 + 2t\left(\alpha \sum_{j\in\mathcal{C}-\{i'\}} m_{i'j} + (\alpha - t)\sum_{j\in\mathcal{P}-\{i'\}} m_{i'j}\right) m_{i'i'} \end{aligned}$$

Since $m_{ii} = m_{i'i'}$, this can be written as

$$\alpha \sum_{j\in\mathcal{C}-\{i\}} m_{ij} + (\alpha - t)\sum_{j\in\mathcal{P}-\{i\}} m_{ij} > \alpha \sum_{j\in\mathcal{C}-\{i'\}} m_{i'j} + (\alpha - t)\sum_{j\in\mathcal{P}-\{i'\}} m_{i'j}$$

Since $\sum_{j\in\mathcal{C}-\{i\}} m_{ij} = \sum_{j\in\mathcal{C}-\{i'\}} m_{i'j}$ and $\sum_{j\in\mathcal{P}-\{i\}} m_{ij} = \sum_{j\in\mathcal{P}-\{i'\}} m_{i'j}$, this inequality is equivalent to:

$$-t \sum_{j\in\mathcal{P}-\{i\}} m_{ij} > 0$$

which is clearly impossible.

Given these results, we will now determine the equilibria by construction, starting from an equilibrium where all individuals live in the center and then looking at equilibrium where, one by one, we move agents from the center to the periphery, starting with agents who have the lowest centrality in the network, that is agent n.

let us consider the case when the number of types is the same as the number of agents.

(i) Let us start with the Central equilibrium where all agents locate in the center, i.e. $\mathcal{C} = N$ and $\mathcal{P} = \varnothing$. Since $\Phi^k(m_{ii})$ is increasing in m_{ii} and since $m_{nn} = \min_i m_{ii}$, we only need to impose that

$$\Phi^{\mathcal{C}}(m_{nn}) > 2c$$

where $\mathcal{C} = N$.

Let us now move one agent at a time by always taking the agent with the lowest centrality in the center and moving her to the periphery. If two agents have the same centrality, then we have to move them together because, below, we show that there cannot be an equilibrium for which two identical agents live in different parts of the city. We also show below that there cannot be other equilibria, i.e. it is not possible to have an agent in the periphery that has a strictly higher centrality than an agent in the center.

(ii) Let us thus move agent n from the center to the periphery to obtain the Core-Periphery with $\mathcal{C} = N - \{n\}$ and $\mathcal{P} = \{n\}$ that we denote $\mathcal{CP}1$. First, observe that

$$\begin{aligned}
\Phi^{\mathcal{CP}1}(m_{n-1n-1}) &= t\,(2\alpha - t)\,(m_{n-1n-1})^2 + 2t\left(\alpha \sum_{j\in\mathcal{C}-\{n-1\}} m_{n-1j} + (\alpha - t)\sum_{j\in\mathcal{P}-\{n-1\}} m_{n-1j}\right) m_{n-1n-1} \\
&= t\,(2\alpha - t)\,(m_{n-1n-1})^2 + 2t\left(\alpha \sum_{j\in N-\{n-1\}} m_{n-1j} - \alpha m_{n-1n} + (\alpha - t)\, m_{n-1n}\right) m_{n-1n-1} \\
&= t\,(2\alpha - t)\,(m_{n-1n-1})^2 + 2t\left(\alpha \sum_{j\in N-\{n-1\}} m_{n-1j}\right) m_{n-1n-1} - 2t^2 m_{n-1n} m_{n-1n-1} \\
&= \Phi^{\mathcal{C}}(m_{n-1n-1}) - 2t^2 m_{n-1n} m_{n-1n-1}
\end{aligned}$$

and

$$\begin{aligned}
\Phi^{\mathcal{CP}1}(m_{nn}) &= t\,(2\alpha - t)\,(m_{nn})^2 + 2t\left(\alpha \sum_{j\in\mathcal{C}-\{n\}} m_{nj} + (\alpha - t)\sum_{j\in\mathcal{P}-\{n\}} m_{nj}\right) m_{nn} \\
&= t\,(2\alpha - t)\,(m_{nn})^2 + 2t\left(\alpha \sum_{j\in N-\{n\}} m_{nj}\right) m_{nn} \\
&= \Phi^{\mathcal{C}}(m_{nn})
\end{aligned}$$

We have thus shown that

$$\Phi^{\mathcal{CP}1}(m_{n-1n-1}) = \Phi^{\mathcal{C}}(m_{n-1n-1}) - 2t^2 m_{n-1n} m_{n-1n-1} \text{ and } \Phi^{\mathcal{CP}1}(m_{nn}) = \Phi^{\mathcal{C}}(m_{nn})$$

This is because when we move agent n from the center to the periphery, her weight changes from α to $\alpha - t$. If we compare $\Phi^{C}(m_{ii})$, given by (17), and $\Phi^{\mathcal{CP}1}(m_{n-1n-1})$ given by (40) for $k = \mathcal{CP}1$ and $i = n-1$, we see that the only difference is the weight $-t$ given to agent n who now lives in the periphery. Quite naturally, when we compare $\Phi^{\mathcal{CP}1}(m_{nn})$ and $\Phi^{\mathcal{C}}(m_{nn})$, there is no difference because we only look at non self-loops.

We need to show that $\Phi^{\mathcal{CP}1}(m_{n-1n-1}) > 2c$ and $\Phi^{\mathcal{CP}1}(m_{nn}) < 2c$. This is equivalent to

$$\Phi^{\mathcal{C}}(m_{n-1n-1}) - 2t^2 m_{n-1n} m_{n-1n-1} > 2c \text{ and } \Phi^{\mathcal{C}}(m_{nn}) < 2c$$

which is equivalent to

$$\Phi^{\mathcal{C}}(m_{nn}) < 2c < \Phi^{\mathcal{C}}(m_{n-1n-1}) - 2t^2 m_{n-1n} m_{n-1n-1}$$

(iii) From the previous equilibrium, let us now move agent $n-1$ from the center to the periphery to obtain the Core-Periphery with $\mathcal{C} = N - \{n-1, n\}$ and $\mathcal{P} = \{n-1, n\}$ that we denote $\mathcal{CP}2$.

First, observe that

$$\Phi^{\mathcal{CP}2}(m_{n-2n-2}) = \Phi^{\mathcal{C}}(m_{n-2n-2}) - 2t^2\left(m_{n-2n-1} + m_{n-2n}\right)m_{n-2n-2}$$

and

$$\Phi^{\mathcal{CP}2}(m_{n-1n-1}) = \Phi^{\mathcal{C}}(m_{n-1n-1}) - 2t^2 m_{n-1n}m_{n-1n-1}$$

We need to show that $\Phi^{\mathcal{CP}2}(m_{n-2n-2}) > 2c$ and $\Phi^{\mathcal{CP}2}(m_{n-1n-1}) < 2c$. This is equivalent to

$$\Phi^{\mathcal{C}}(m_{n-2n-2}) - 2t^2\left(m_{n-2n-1} + m_{n-2n}\right)m_{n-2n-2} > 2c \text{ and } \Phi^{\mathcal{C}}(m_{n-1n-1}) - 2t^2 m_{n-1n}m_{n-1n-1} < 2c$$

This is equivalent to

$$\Phi^{\mathcal{C}}(m_{n-1n-1}) - 2t^2 m_{n-1n}m_{n-1n-1} < 2c < \Phi^{\mathcal{C}}(m_{n-2n-2}) - 2t^2\left(m_{n-2n-1} + m_{n-2n}\right)m_{n-2n-2}$$

(iv) From the previous equilibrium, let us now move agent $n-2$ from the center to the periphery to obtain the Core-Periphery with $\mathcal{C} = N - \{n-2, n-1, n\}$ and $\mathcal{P} = \{n-2, n-1, n\}$ that we denote $\mathcal{CP}3$. First, observe that

$$\Phi^{\mathcal{CP}3}(m_{n-3n-3}) = \Phi^{\mathcal{C}}(m_{n-3n-3}) - 2t^2\left(m_{n-3n-2} + m_{n-3n-1} + m_{n-3n}\right)m_{n-3n-3}$$

and

$$\Phi^{\mathcal{CP}3}(m_{n-2n-2}) = \Phi^{\mathcal{C}}(m_{n-2n-2}) - 2t^2\left(m_{n-2n-1} + m_{n-2n}\right)m_{n-2n-2}$$

We need to show that $\Phi^{\mathcal{CP}3}(m_{n-3n-3}) > 2c$ and $\Phi^{\mathcal{CP}3}(m_{n-2n-2}) < 2c$. This is equivalent to

$$\Phi^{\mathcal{C}}(m_{n-3n-3}) - 2t^2\left(m_{n-3n-2} + m_{n-3n-1} + m_{n-3n}\right)m_{n-3n-3} > 2c$$

and

$$\Phi^{\mathcal{C}}(m_{n-2n-2}) - 2t^2\left(m_{n-2n-1} + m_{n-2n}\right)m_{n-2n-2} < 2c$$

which is equivalent to

$$\Phi^{\mathcal{C}}(m_{n-2n-2}) - 2t^2\left(m_{n-2n-1} + m_{n-2n}\right)m_{n-2n-2} < 2c < \Phi^{\mathcal{C}}(m_{n-3n-3}) - 2t^2\left(\sum_{j\in\mathcal{P}}^{n} m_{n-3j}\right)m_{n-3n-3}$$

where $\mathcal{P} = \{n-2, n-1, n\}$.

(v) We can continue like that until we reach a Peripheral equilibrium with $\mathcal{C} = \varnothing$ and $\mathcal{P} = N$.

(vi) From the previous equilibrium (i.e., $\mathcal{C} = \{1\}$ and $\mathcal{P} = N - \{1\}$), let us finally move agent 1 from the center to the periphery to obtain the Peripheral equilibrium with $\mathcal{C} = \varnothing$ and $\mathcal{P} = N$. Observe that

$$\Phi^{\mathcal{P}}(m_{11}) = \Phi^{\mathcal{C}}(m_{11}) - 2t^2\left(\sum_{j\in\mathcal{P}-\{1\}}^{n} m_{1j}\right)m_{11}$$

We need to show that $\Phi^{\mathcal{P}}(m_{11}) < 2c$. This is equivalent to

$$\Phi^{\mathcal{C}}(m_{11}) - 2t^2 \left(\sum_{j \in \mathcal{P}-\{1\}}^{n} m_{1j} \right) m_{11} < 2c$$

Since all the conditions are mutually exclusive, we have shown that, for each condition, there exists a unique corresponding equilibrium as defined in the Proposition. One can always find a set of parameters α, t, c, and θ that satisfies each condition. We give some examples in Section 4.2.

If the number of types is less than the number of agents, then each step described above has to be made by type and not by individual. The conditions on parametrs will be exactly the same.

Let us show that it is not possible to have any other equilibrium such that an agent who resides in the periphery has a higher centrality than an agent who lives in the center. Without loss of generality, take the Core-periphery equilibrium $\mathcal{CP}3$, described in (iv), where $\mathcal{C} = N - \{n-2, n-1, n\}$ and $\mathcal{P} = \{n-2, n-1, n\}$. Is it possible to have an equilibrium where $\mathcal{C} = N - \{n-3, n-1, n\}$ and $\mathcal{P} = \{n-3, n-1, n\}$? For this equilibrium to be true, it has to be (at least) that $\Phi^{\mathcal{CP}3}(m_{n-3n-3}) < 2c$ and $\Phi^{\mathcal{CP}3}(m_{n-2n-2}) > 2c$, which implies that $\Phi^{\mathcal{CP}3}(m_{n-2n-2}) > \Phi^{\mathcal{CP}3}(m_{n-3n-3})$. Using (40), it is easily verified that $\Phi^{\mathcal{CP}3}(m_{ii})$ is increasing in m_{ii}. But since, by definition, $m_{n-3n-3} > m_{n-2n-2}$, which implies that $\Phi^{\mathcal{CP}3}(m_{n-2n-2}) < \Phi^{\mathcal{CP}3}(m_{n-3n-3})$, a contradiction. This reasoning can be applied to any equilibrium.

Let us finally show that the number of equilibria is equal to the number of types of agents plus one (the type of each agent is defined by her Katz-Bonacich centrality). Denote the number of types of agents by $\nu \leq n$. Since we have shown that, in any equilibrium, two identical agents (i.e. with the same Katz-Bonacich centrality) have to reside in the same part of the city, it has to be that the number of Core-Periphery equilibria is equal to $\nu - 1$. If we add the Central equilibrium and the Peripheral equilibrium, then the number of all equilibria is equal to $\nu + 1$. ∎

Proof of Proposition 6. Remember from the proof of Proposition 4 that $m_{ii}^{(2)}$ is defined by $\Lambda(m_{ii}, c) \equiv \Phi(m_{ii}) - 2c$, which is given by (40) or

$$\Lambda(m_{ii}, c) \equiv t\,(2\alpha - t)\,(m_{ii})^2 + 2t \left(\alpha \sum_{j \in \mathcal{C}-\{i\}} m_{ij} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} m_{ij} \right) m_{ii} - 2c = 0$$

Let us differentiate $m_{ii}^{(2)}$ with respect to c. It is straightforward to see that:

$$\frac{\partial m_{ii}^{(2)}}{\partial c} > 0$$

which means that when c increases, $m_{ii}^{(2)}$ rises. Thus, the set $\mathcal{P}$ increases and more people live in the periphery.

Let us differentiate $m_{ii}^{(2)}$ with respect to t. For that, let us differentiate $\Lambda(m_{ii}, c)$. We obtain

$$\frac{\partial m_{ii}^{(2)}}{\partial t} = -\frac{2(\alpha - t)(m_{ii})^2 + 2\alpha m_{ii} \sum_{j \in \mathcal{C}-\{i\}} m_{ij} + 2(\alpha - 2t) m_{ii} \sum_{j \in \mathcal{P}-\{i\}} m_{ij}}{2t(2\alpha - t) m_{ii} + 2t\alpha \sum_{j \in \mathcal{C}-\{i\}} m_{ij} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} m_{ij}} < 0$$

Let us finally differentiate $m_{ii}^{(2)}$ with respect to θ. First, let us write $\Lambda(m_{ii}, c)$ in terms of θ by noticing that $m_{ii} = \sum_{k=0}^{+\infty} \theta^k g_{ii}^{[k]}$ and $m_{ij,j\neq i} = \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]}$. We have:

$$\begin{aligned}
\Phi(\theta) \quad \equiv \quad & t(2\alpha - t)\left(\sum_{k=0}^{+\infty} \theta^k g_{ii}^{[k]}\right)^2 + 2t\left(\alpha \sum_{j \in \mathcal{C}-\{i\}} \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]}\right) \sum_{k=0}^{+\infty} \theta^k g_{ii}^{[k]} \\
& -2c
\end{aligned}$$

This implies that

$$\begin{aligned}
\Lambda'(\theta, c) \quad = \quad & 2t(2\alpha - t)\left(\sum_{k=0}^{+\infty} k\theta^{k-1} g_{ii}^{[k]}\right) \\
& +2t\left(\alpha \sum_{j \in \mathcal{C}-\{i\}} \sum_{k=0}^{+\infty} k\theta^{k-1} g_{ij}^{[k]} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} \sum_{k=0}^{+\infty} k\theta^{k-1} g_{ij}^{[k]}\right) \sum_{k=0}^{+\infty} \theta^k g_{ii}^{[k]} \\
& +2t\left(\alpha \sum_{j \in \mathcal{C}-\{i\}} \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} \sum_{k=0}^{+\infty} \theta^k g_{ij}^{[k]}\right) \sum_{k=0}^{+\infty} k\theta^{k-1} g_{ii}^{[k]}
\end{aligned}$$

Since $\mathbf{G}$ and all its powers are positive matrices, and the coefficients θ^k increase with θ, it immediately follows that the infinite series result in a matrix with all entries larger or equal than the infinite series with the initial value of θ. As a result, $\Lambda'(\theta, c) > 0$.

Let us now totally differentiate $\Lambda(m_{ii}, c)$. We obtain:

$$\frac{\partial m_{ii}^{(2)}}{\partial \theta} = -\frac{\Lambda'(\theta, c)}{2t(2\alpha - t) m_{ii} + 2t\alpha \sum_{j \in \mathcal{C}-\{i\}} m_{ij} + (\alpha - t) \sum_{j \in \mathcal{P}-\{i\}} m_{ij}} < 0$$

Proof of Proposition 7.

From Proposition 5, we know that more central agents (here agent 1) cannot locate further away from the center than less central agents (here agents 2 and 3) and that agents 2 and 3 have to live in the same part of the city, which implies that, for example, a Core-Periphery equilibrium where $\mathcal{C} = \{1, 2\}$ and $\mathcal{P} = \{3\}$ cannot exist. As a result, there will only exist 3 equilibria.

(i) Let us first show under which condition there exists a unique *central equilibrium* for which all individuals live in the center, i.e. $\mathcal{C} = \{1, 2, 3\}$ and $\mathcal{P} = \varnothing$.

Using Proposition 5, we only need to show that $\Phi^C(m_{33}) > 2c$. Using (41), we have:

$$\Phi^C(m_{33}) = t(2\alpha - t)(m_{33})^2 + 2\alpha t(m_{31} + m_{32}) m_{33}$$

Using (18), we obtain:

$$\begin{aligned}\Phi^{C}(m_{33}) &= t(2\alpha - t)\left(\frac{1-\theta^2}{1-2\theta^2}\right)^2 + 2\alpha t\left(\frac{\theta}{1-2\theta^2} + \frac{\theta^2}{1-2\theta^2}\right)\left(\frac{1-\theta^2}{1-2\theta^2}\right) \\ &= \frac{t(1-\theta)(1+\theta)^2[2\alpha-(1-\theta)t]}{(1-2\theta^2)^2}\end{aligned}$$

As a result, the condition $\Phi^{C}(m_{33}) > 2c$ is equivalent to

$$c < \frac{t(1-\theta)(1+\theta)^2[2\alpha-(1-\theta)t]}{2(1-2\theta^2)^2}$$

(ii) Let us now show that there exists a *core-periphery equilibrium* for which individuals 1 lives in the center while individuals 2 and 3 reside in the periphery. This means that $\mathcal{C} = \{1\}$ while $\mathcal{P} = \{2,3\}$. Using Proposition 5, the condition for the core-periphery equilibrium $\mathcal{C} = \{1\}$, $\mathcal{P} = \{2,3\}$ to exist and to be unique is given by:

$$\Phi^{\mathcal{C}}(m_{22}) < 2c < \Phi^{\mathcal{C}}(m_{11}) - 2t^2(m_{12}+m_{13})m_{11}$$

Since $\Phi^{\mathcal{C}}(m_{33}) = \Phi^{\mathcal{C}}(m_{22})$, the value of $\Phi^{\mathcal{C}}(m_{33})$ is given above. Let us determine $\Phi^{\mathcal{C}}(m_{11}) - 2t^2(m_{12}+m_{13})m_{11}$. Using (41), we have:

$$\begin{aligned}\Phi^{C}(m_{11}) &= t(2\alpha - t)(m_{11})^2 + 2t\left(\alpha\sum_{j=N-\{1\}} m_{1j}\right)m_{11} \\ &= t(2\alpha - t)(m_{11})^2 + 2t\alpha(m_{12}+m_{13})m_{11}\end{aligned}$$

Using (18), we obtain:

$$\begin{aligned}\Phi^{C}(m_{11}) &= t(2\alpha - t)\left(\frac{1}{1-2\theta^2}\right)^2 + 4t\alpha\left(\frac{\theta}{1-2\theta^2}\right)\left(\frac{1}{1-2\theta^2}\right) \\ &= \frac{t[2\alpha(1+2\theta)-t]}{(1-2\theta^2)^2}\end{aligned}$$

and thus

$$\begin{aligned}\Phi^{\mathcal{C}}(m_{11}) - 2t^2(m_{12}+m_{13})m_{11} &= \frac{t[2\alpha(1+2\theta)-t]}{(1-2\theta^2)^2} - 2t^2\left(\frac{2\theta}{1-2\theta^2}\right)\left(\frac{1}{1-2\theta^2}\right) \\ &= \frac{t[2\alpha(1+2\theta)-t(1+4\theta)]}{(1-2\theta^2)^2}\end{aligned} \tag{43}$$

As a result, the condition $\Phi^{\mathcal{C}}(m_{33}) < 2c < \Phi^{\mathcal{C}}(m_{11}) - 2t^2(m_{12}+m_{13})m_{11}$ can be written as:

$$\frac{t(1-\theta)(1+\theta)^2[2\alpha-(1-\theta)t]}{2(1-2\theta^2)^2} < c < \frac{t[2\alpha(1+2\theta)-t(1+4\theta)]}{2(1-2\theta^2)^2}$$

(iii) Let us finally show that there exists a unique *peripheral equilibrium* for which all individuals live in the center, i.e. $\mathcal{C} = \varnothing$ and $\mathcal{P} = \{1, 2, 3\}$. Using Proposition 5, the condition is

$$\Phi^{\mathcal{C}}(m_{11}) - 2t^2 (m_{12} + m_{13}) m_{11} < 2c$$

which, using (43), is equivalent to

$$c > \frac{t \left[2\alpha (1 + 2\theta) - t (1 + 4\theta)\right]}{2 \left(1 - 2\theta^2\right)^2}$$

We have thus proven all the statements made in the Proposition. ∎

Proof of Proposition 8.

From Proposition 5, we know that identical agents have to live in the same part of the city. In the complete network, this implies that cannot be a Core-Periphery equilibrium and thus there must only be two equilibria.

(i) Let us first show that there exists a unique *central equilibrium* for which all individuals live in the center. This means that $\mathcal{C} = \{1, 2, 3\}$ while $\mathcal{P} = \varnothing$. Using Proposition 5, we need to show that $\Phi^C(m_{33}) > 2c$. Using (41), we have:

$$\Phi^C(m_{33}) = t (2\alpha - t) (m_{33})^2 + 2\alpha t (m_{31} + m_{32}) m_{33}$$

Using (22), we have:

$$\Phi^C(m_{33}) = \frac{t (1 - \theta)^2 (2\alpha + 4\alpha t - t)}{\left(1 - \theta - 2\theta^2\right)^2}$$

The condition $\Phi^C(m_{33}) > 2c$ can thus be written as

$$c < \frac{t (1 - \theta)^2 (2\alpha + 4\alpha t - t)}{2 \left(1 - \theta - 2\theta^2\right)^2}$$

(ii) Let us now show that there exists a unique *peripheral equilibrium* for which all individuals live in the periphery. This means that $\mathcal{P} = \{1, 2, 3\}$ while $\mathcal{C} = \varnothing$. Using Proposition 5, the condition is

$$c > \frac{t (1 - \theta)^2 (2\alpha + 4\alpha t - t)}{2 \left(1 - \theta - 2\theta^2\right)^2}$$

and the result is proven. ∎

Proof of Proposition 9.

Apply Proposition 5 to this network and the results follows. ∎