Games on Networks:
Direct Complements and Indirect Substitutes

Sergio Currarini* Elena Fumagalli † Fabrizio Panebianco ‡

December 2012

Abstract

Many types of economic and social activities involve significant behavioral complementarities (peer effects) with neighbors in the social network. The same activities often exert externalities, that cumulates in "stocks" affecting agents' welfare and incentives. For instance, smoking is subject to peer effects, and the stock of passive smoke increases the marginal risks of bad health, decreasing the incentives to smoke. In the linear quadratic framework studied by Ballester et al. (2006), we consider contexts where agents' incentives decrease with the "stock" to which neighbors are exposed (agents may, for instance, care about their friends' health). In such contexts, the patterns of strategic interaction differ from the network of social relations, as agents display strategic substitution with distance-two neighbors. We show that behavior is predicted by a weighted Bonacich centrality index, with weights accounting for distance-two relations. We find that both maximal behavior and key-players tend to move to the periphery of the network, and we discuss the effect of close-knit communities and segregated groups on aggregate behavior. We finally discuss the implications for peer effects identification and for the emergence of potential biases in the estimation of social effects.
Keywords: Networks, Peer Effects, Key-player, Centrality, Substitutes, Altruism.

1 Introduction

Socio-economic behavior typically occurs within relational networks of various kinds, describing the pattern of interpersonal, institutional and technological ties, among others. Economic agents typically interact with their direct neighbors in the network, jointly consuming or producing goods, discussing political opinions, sharing information, etc. Consequently, agents who are linked in the network tend to display correlation in behavior. Positive correlation is particularly pervasive in many social and economic contexts, and has been the object of a vast literature in economics. Behavioral complementarities, or "peer effects", have commanded substantial attention in economics because of their amplification of individual shocks in terms of aggregate outcomes (Glaeser et al., 2003), and because of the related difficulty in correctly estimating individual elasticities to such shocks. In many instances where they have shown to be important, peer effects stem from emulation, shared identity and conformity: examples include risky adolescent behavior (smoking, drug use, educational attainments), criminal activities, health related behavior, habits on the workplace (Evans et al., 1992; Gaviria and Raphael, 2001; Kirke, 2004; Christakis and Fowler, 2007; Clark and Loheac, 2007; Poutvara and Siemers, 2008; Fowler and Christakis, 2008; Calvó-Armengol et al., 2009; Fletcher, 2010). More in general, local complementarities arise whenever one's incentives to act increase with neighbors' actions. For instance,

*Universita' Ca' Foscari di Venezia and FEEM. Email: s.currarini@unive.it
†Universite de Lausanne, CH. Email: elena.fumagalli@unil.ch
‡Universita' degli Studi di Milano - Bicocca. Email: fabrizio.panebianco@unimib.it

this is the case when links describe technological complementarities, so that firms face larger demand and larger incentives to produce when neighbor firms increase their output. These complementarities are also present when links transmit local spillovers in investment and firms face higher incentives to invest when other neighbor firms invest more.

The same actions that generate peer effects often exert local externalities, that cumulate in "stocks" affecting agents' welfare. For example, passive smoke, one determinant of health risks, results from the sum of friends' smoking intensities; the price of an intermediate product produced by a neighbor firm depends on the aggregate demand for that product coming from other neighbor firms; the profitability of investing in a neighbor market may depend on how many neighbor firms have access to that market. These stocks not only affect welfare, but typically affect the incentives of agents to act. If, for instance, a smoker perceives that the marginal health risk due to an additional cigarette increases with the stock of passive and active smoke, her incentive to smoke will depend (negatively) on the smoking intensity of her friends, counteracting the peer effect. If she cares about her friends at all, her incentives will also decrease with the stock of smoke her friends are exposed to. Similarly, firms' incentives to produce decrease with the price of their neighbors' intermediary products, and firms' investment will decrease with the investments of firms that share common markets.

In the previous examples, agents end up interacting with agents at distance two in the network, whose actions are strategic substitutes with their own. These additional interactions are captured by augmenting the network of local interactions to include, together with direct complementarities, indirect (i.e., with agents at distance two) substitutability. Equilibrium behavior will depend on both types of interaction, and the behavioral consequences of a change in the network of direct complementarities will relate to the changes in the augmented network. While, for instance, increasing the density of social contacts has the unambiguous effect of increasing the sources of complementarities (direct links) and behavior, the parallel increase of common neighbors, with the associated substitute effects, may well counteract the increase in behaviour.

In this paper we employ the analytical framework developed by Ballester et al. (2006) for linear quadratic games on networks to study the joint effect of direct and indirect interaction. In addition to the traditional peer effects with neighbors, agents suffer from the negative externalities stemming form their neighbors' actions, and, to an extent captured by a parameter, from the stock of externalities to which their neighbors are exposed to. In both cases we assume a quadratic loss function. The quadratic specification implies that an agent's incentive to act decreases with her distance-two neighbors' actions, introducing the indirect strategic substitutability discussed above. For this model, we characterize individual behavior as a function of agents centralities in the augmented network of direct and indirect effects. We then show that equilibrium behavior is predicted by a weighted variant of agents' Bonacich centralities in the original network of direct complementarities, where weights keep track of distance-two relations. We show that agents who occupy a central position in the original network typically fail to be central in the augmented network, due to their intense two-distance relations with other agents and to the implied diffuse strategic substitutability. This tendency of behavior to move towards the periphery of the network is consistent, for instance, with some robust evidence obtained by Christakis and Fowler (2008) on the progressive marginalization of heavy smokers in social networks.

We then study how the presence of indirect substitutes affects the design of network-based policies. We first analyze how changes in the network affect individual and aggregate behavior. For regular networks, we show that the relationship between network density and behavior is non monotonic, with a positive correlation in sparse networks, and a negative correlation when the network becomes dense. Intuitively, in dense networks the strategic substitutions that flow on two-distance neighborhoods (of cardinality approximated by the square of the degree) tend to dominate the complementarities that flow on one-distance relations (of cardinality equal to the degree). For the case of non regular networks, we show that aggregate behavior can be decreased by creating clusters of agents and increased by dissolving existing clusters.

We then examine the problem of identifying key-players (in the definition of Ballester et al., 2006) in this new context. We show that key players in the augmented network tend to be at the periphery of the original network, and that this tendency is stronger the more important is the strength of indirect substitutability compared to peer effects. Finally, we study the effect of policies that affect the degree of segregation in the network, within a context of agents with heterogeneous gains from the action. We show that aggregate behavior is minimal for moderate levels of segregation, where agents have neighbors of mixed types.

In the final part of the paper we derive some implications for the empirical estimation of peer effects. The interpretation of indirect substitutes is mainly that of altruism towards one's friends in the network. We both characterize the sign and the nature of the bias that originates when externalities and altruism are not taken into account and we address the challenges in identification of peer effects due to the reflection problem (Manski, 1993). In particular, we derive conditions for the identification of peer effects to account for the presence of altruism both in the case in which the peer effect is defined as the sum of peers' actions (Liu and Lee, 2010) and in the one in which the peer effect is defined as its average (Bramoullé et al., 2009). In the first case, even without altruism, peer effects cannot be identified if people interact in regular networks. When altruism is included in the model, conditions for identification become more stringent and some other types of networks (e.g. the star) have to be excluded. In addition, we find not including altruism in the model makes the peer effects systematically underestimated. Finally, we discuss the choice of the instruments needed to estimate both the endogenous peer effect and the endogenous effect of altruism.

The paper is organized as follows. Section 2 describes the linear quadratic model. Section 3 characterizes behavior as a function of Bonacich centralities in the augmented network and in the original network. Section 4 addresses various network-based policies. Section 5 discusses empirical implications. Section 6 concludes.

2 The Model

We consider a set N of n agents, organized in a network $\mathbf{g}$ defined by a $n \times n$ matrix $\mathbf{G}$ whose generic element $g_{ij} \in \{0, 1\}$ measures the presence of a social tie (or link) between agents i and j. We limit our analysis to symmetric networks, that is $g_{ij} = g_{ji}$ for all $i, j \in N$. Agents i and j are "neighbors" in $\mathbf{g}$ whenever $g_{ij} = 1$, and the degree d_i of agent i in the network $\mathbf{g}$ denotes the number of neighbors of i in $\mathbf{g}$. A *path* between i and j in $\mathbf{g}$ is a series of distinct agents $i_1, i_2, ..., i_m$ such that $i_1 = i$, $i_m = j$ and $g_{i_p i_{p-1}} = 1$ for all $p = 2, 3, ..., m$. Similarly, we define a *walk* by dropping the requirement of distinct agents. We finally use the convention $g_{ii} = 0$, $\forall i$.

Each agent i chooses an action $x_i \in \mathbb{R}_+$; for each agent i, we denote by

$$Q_i \equiv \left(\sum_{k \in N} g_{ik} x_k + x_i \right)$$

the sum of all actions taken in the neighborhood of i in $\mathbf{g}$ (that is, taken by i and by neighbors of i). Agent i derives the following utility from the vector $\bar{x} \in \mathbb{R}^n_+$ of actions chosen in the network:

$$U_i = \alpha_i x_i - \gamma_0 \frac{x_i^2}{2} + \phi \sum_{j \in N} g_{ij} x_i x_j - \gamma_1 \frac{Q_i^2}{2} - \gamma_2 \sum_{j \in N} g_{ij} \frac{Q_j^2}{2} \tag{1}$$

The first two terms of the function U capture the private benefits from one's own action, which may be a source of heterogeneity when $\alpha_i \neq \alpha_j$ for some i and j. These benefits are the sum of a linear increasing part and a quadratic decreasing part, with intensity measured by the parameter γ_0. The third term, with $\phi > 0$, captures local complementarities that are at work on the links of the network: the marginal incentive to act increases in the aggregate level of actions taken by one's neighbors in $\mathbf{g}$. The parameter ϕ measures the intensity of such complementarities. These first three terms make the functional form studied by Ballester et al. (2006). The fourth term captures the external effects of agents' actions on their neighbors (externalities) and on themselves. The stock Q_i is assumed to affect i's payoffs quadratically. Finally, the last term measures the effect on agent i's payoff of the aggregate stock of actions to which agent i's neighbors are exposed. We provide three illustrative examples in which the parameter γ_2, measuring the intensity of the last effect, takes on different interpretations: altruism, congestion and price effects.

Example 1. Peer effects, externalities and altruism in social networks. Actions exert peer effects between neighbors, together with a negative externality. Consider for instance, the case of smoking and passive smoke to which neighbors are exposed. The parameter ϕ measures the intensity of local behavioral complementarities; the parameter γ_1 measures the concern for the negative externality to which one is exposed; the parameter γ_2 measures the degree of concern that agents have for the amount of negative externality that neighbors are exposed to. This last term can be though of as an element of altruism towards neighbors.

Example 2. Investments with local spillovers and increasing returns. Each firm i chooses a level of investment x_i. Firm i's return from the investment is given by some related economic activity on i's neighbourhood. Investment technology features increasing returns, and there are local spillovers, so that neighbors' investments decrease the marginal cost of investing. The profitability of the economic activity on any given node depends negatively on the activity of other firms on that node. Since the firm activity increases with its investment level, the profitability of firm i's activity on a given neighbor market j decreases with the overall activity of firms that are neighbors of j. Summing up, ϕ measures the increase in incentives to invest due to local spillovers in investment; γ_1 is zero; γ_2 measures the decrease in incentives to invest due to the decreased profitability of the firm economic activity following an increase in rivals activity on common markets.

Example 3. Local complementarities in production. A set of firms produces a different commodity x_i each. Each commodity is sold in a monopolistic market, and is also used as input by neighbor firms. Complementarities in production imply that the demand for a given product increases with the production level of neighbor firms. These complementarities are measured by the term ϕ. In addition, the increase in production by firm neighbors' of j causes an increase in the demand for commodity j and therefore an increase in its price. If i is one of j's neighbors, this implies an increase in the marginal cost of production of i and, therefore, a decrease in the incentives of i to produce. Hence the term γ_2.

Rewrite (1), in order to isolate terms that depend linearly on agent i's action, terms that depend on the square of agent i's action, terms that depend on the product of agents i and j's actions and terms that are independent from i's actions. We can then frame our problem in terms of the linear-quadratic form studied in Ballester et al. (2006)'s analysis of games on networks. Let $g_{ij}^{[2]}$ denote the generic term of the squared matrix $\mathbf{G^2}$, counting the number of walks of length two from node i to node j in $\mathbf{G}$. Let also h_{-i} denote the sum of all terms that in (1) do not depend on x_i. We rewrite (1) as follows (see Appendix A):

$$U_i = \alpha x_i - \frac{1}{2}\sigma_i x_i^2 + (\phi - \gamma_1 - \gamma_2)\sum_{j \in N} g_{ij} x_i x_j - \gamma_2 (\sum_{k \neq i} g_{ik}^{[2]}) x_i x_k + h_{-i}, \tag{2}$$

where we have denoted by

$$\sigma_i = \gamma_0 + \gamma_1 + \gamma_2 d_i$$

the coefficient that multiplies the square of agent i's action in (1) and we have used the fact that $\sum_j g_{ij}^2 = d_i$ since $g_{ij} \in \{0, 1\}$ for all ij.

Equation (2) shows two key features of our model. First, the sign and the intensity of agents' strategic interaction with neighbors is given by the "net complementarity" parameter $(\phi - \gamma_1 - \gamma_2)$, where the strategic complementarity due to peer effects is corrected by agents' concern for the effect of externalities (γ_1) and for their concern for their neighbors' (γ_2). Due to the convexity of these effects, agents' actions acquire elements of strategic substitutability, which reduce, and possible revert, peer-effects. Second, the parameter γ_2 measures the strategic interdependence (of the substitute type) with distance-two neighbors in the network $\mathbf{g}$.

3 Equilibrium behavior on the Network

We now characterize the Nash Equilibrium of the game with set of players N, strategy set $\mathbb{R}_+$ for each player, and payoff functions given by (1). This section will heavily build on Ballester et al. (2006)'s analysis of linear quadratic games played on networks. Assume by now that α is homogeneous across agents, simplifying the analysis without changing the qualitative results; we study the role of different α's in section 4.3 for the case of two values.

3.1 Existence and Characterization

The first order conditions characterizing an internal equilibrium vector of actions $\bar{x}$ are written in the following matrix form, where each line refers to a specific agent:

$$\alpha \cdot \bar{1} = \left[(\gamma_1 + \gamma_0)\mathbf{I} - (\phi - \gamma_1 - \gamma_2)\mathbf{G} + \gamma_2 \mathbf{G^2} \right] \bar{x}. \tag{3}$$

In obtaining (3) we have used the definition of σ_i and the fact that the main diagonal of $\mathbf{G^2}$ has d_i at the i^{th} row. Dividing by $(\gamma_1 + \gamma_0)$ and factorizing terms we obtain:

$$\frac{\alpha}{(\gamma_1 + \gamma_0)} \cdot \bar{1} = \left[\mathbf{I} - \frac{\phi - \gamma_1 - \gamma_2}{(\gamma_1 + \gamma_0)} \left(\mathbf{G} - \frac{\gamma_2}{\phi - \gamma_1 - \gamma_2} \mathbf{G^2} \right) \right] \bar{x}. \tag{4}$$

3.1.1 A benchmark case: $\gamma_2 = 0$

When $\gamma_2 = 0$, condition (4) reduces to:

$$\frac{\alpha}{\gamma_1 + \gamma_0} \cdot \bar{1} = \left[\mathbf{I} - \frac{(\phi - \gamma_1)}{\gamma_1 + \gamma_0} \mathbf{G} \right] \bar{x}. \tag{5}$$

As long as $(\phi - \gamma_1) > 0$, we can directly apply results from Ballester et al. (2006), that characterize equilibrium behavior *via* the Bonacich centralities of agents in $\mathbf{G}$. For completeness, we introduce here this notion of centrality, which will be used throughout the paper.

Definition 1 *Given the network* $\mathbf{g}$ *with adjacency matrix* $\mathbf{G}$*, the Bonacich centrality matrix of* $\mathbf{g}$ *with parameter* a *is given by:*

$$\mathbf{M}(\mathbf{G}, a) \equiv (I - a\mathbf{G})^{-1}. \tag{6}$$

The matrix $\mathbf{M}(\mathbf{G}, a)$ is well defined if $\mu(\mathbf{G}) < \frac{1}{a}$ with $\mu(\mathbf{G})$ being the largest eigenvalue associated with the matrix $\mathbf{G}$.

Definition 2 *Given the network* $\mathbf{g}$ *with adjacency matrix* $\mathbf{G}$*, the vector of Bonacich centralities of* $\mathbf{g}$ *with parameter* a *is given by:*

$$\mathbf{b}(\mathbf{G}, a) \equiv \mathbf{M}(\mathbf{G}, a) \cdot \bar{1}. \tag{7}$$

We will denote by $b(\mathbf{G}, a)$ the sum of all agents' centralities, that is, the internal product $\bar{1}' \cdot \mathbf{b}(\mathbf{G}, a)$. Using (5)-(7) we immediately obtain the next proposition, due to Ballester et al. (2006):

Proposition 1 *Let $\mu(\mathbf{G}) < \frac{\gamma_1+\gamma_0}{\phi-\gamma_1}$, where $\mu(\mathbf{G})$ denotes the largest Eigenvalue of $\mathbf{G}$. Then the unique interior Nash Equilibrium of the game is given by:*

$$\bar{x} = \frac{\alpha}{\gamma_1 + \gamma_0} \mathbf{b}(\mathbf{G}, \frac{\phi - \gamma_1}{\gamma_1 + \gamma_0}). \tag{8}$$

Proposition 1 establishes a proportional relation between the Nash equilibrium actions and the Bonacich centralities in the network $\mathbf{G}$ with parameter $\frac{\phi-\gamma_1}{\gamma_1+\gamma_0}$. More central agents choose larger actions, due to their higher exposure to the direct and indirect effects of behavioral complementarities. The constraints on the term $\frac{\phi-\gamma_1}{\gamma_1+\gamma_0}$ are meant to preserve the strategic complementarity nature of the game, and to ensure that such complementarities do not cause an unbounded increase in equilibrium actions (the largest Eigenvalue can be thought of as a measure of the networks' connectedness, and in regular networks it coincides with the average degree).

Proposition 1 also shows that the (negative) local externalities stemming from agents' actions counteract the strength of peer effects in the parameter $\phi-\gamma_1$. This parameter acts as a weighting factor applied to the various walks agents have in the network $\mathbf{G}$, and contributes to form their centrality indices: the larger this parameter, the more the centrality index of each agent is affected by his longer walks in the network (relative to shorter paths). Moreover, the externality parameter γ_1 enters the proportionality factor between centralities and actions: the larger γ_1, the weaker the factor.

3.1.2 Equilibrium with $\gamma_2 > 0$

Direct inspection of condition (4) provides insights on the role of the parameter γ_2 in shaping equilibrium behavior. Letting

$$\eta \equiv \frac{\gamma_2}{\phi - \gamma_1 - \gamma_2}, \tag{9}$$

the new matrix

$$\tilde{\mathbf{G}} \equiv \mathbf{G} - \eta \mathbf{G^2} \tag{10}$$

describes the patterns of interaction in the game. In this new matrix, strategic interaction occurs between agents who are connected in $\mathbf{G}$, and between agents that share common neighbors in $\mathbf{G}$ (captured by the network $\mathbf{G^2}$, counting for each pair of agents the number of walks of length two between these agents in $\mathbf{G}$). More precisely, the generic element $\tilde{g}_{ij}$ is given by:

$$\tilde{g}_{ij} = \begin{cases} 0 & \text{if } g_{ij} = 0 \text{ and } g_{ij}^{[2]} = 0 \\ 1 & \text{if } g_{ij} = 1 \text{ and } g_{ij}^{[2]} = 0 \\ -\eta g_{ij}^{[2]} & \text{if } g_{ij} = 0 \text{ and } g_{ij}^{[2]} > 0 \\ 1 - \eta g_{ij}^{[2]} & \text{if } g_{ij} = 1 \text{ and } g_{ij}^{[2]} > 0 \end{cases}$$

Note that $\tilde{\mathbf{G}}$ always contains negative terms; in fact, diagonal terms are given by:

$$\tilde{g}_{ii} = -\eta d_i.$$

Sufficient conditions for an interior positive equilibrium can be obtained applying Ballester et al. (2006)'s normalization to the matrix $\tilde{\mathbf{G}}$. Let $\bar{c} = max\{\tilde{g}_{ij}\} = 1$ and $\underline{c} = min\{\tilde{g}_{ij}\}$ denote respectively the maximal complementarity and substitutability in $\tilde{\mathbf{G}}$, and let $\theta = -min\{0, \underline{c}\} > 0$. Let also $\lambda = \overline{c} + \theta$ denote the range between the maximal and minimal elements in $\tilde{\mathbf{G}}$. Define then the new matrix $\mathbf{C}$ whose generic element is as follows:

$$c_{ij} = \frac{\tilde{g}_{ij} + \theta}{\lambda} \in [0, 1]. \tag{11}$$

The system in (3) can be rewritten in terms of the matrix $\mathbf{C}$ as:

$$\alpha \bar{1} = \left[(\gamma_1 + \gamma_0) I + \frac{\gamma_2}{\eta} (\theta \cdot U - \lambda \mathbf{C}) \right] \bar{x}.$$

The following proposition applies the results of Ballester et al. (2006) to the present setting.

Proposition 2 *Consider a symmetric network* $\mathbf{g}$ *with adjacency matrix* $\mathbf{G}$*, and the matrix* $\mathbf{C}$ *defined as in (11). Let* $\frac{\eta(\gamma_1+\gamma_0)}{\lambda\gamma_2} > \mu(\mathbf{C})$*. The unique Nash equilibrium of the game is given by:*

$$\bar{x} = \frac{\alpha \eta \ \mathbf{b}(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)})}{\eta(\gamma_1 + \gamma_0) + \gamma_2 \theta b(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)})}. \tag{12}$$

Note that, when $\gamma_2 = 0$, $g_{ij} = \tilde{g}_{ij}$, $\lambda = 1$ and $\theta = 0$, this implies that $\mathbf{C} = \mathbf{G}$ and expression (8) obtains again.

3.2 Centrality and behavior

The issue we address in this section is how equilibrium behavior relates to agents' positions in the social network $\mathbf{G}$. Two aspects of behavior are of particular interest: the relation between individual actions and individual positions in the network and the relation between the network architecture and aggregate actions.

The following proposition shows that the equilibrium with $\gamma_2 > 0$ is given by a weighted variant of the Bonacich centrality vector of the network $\mathbf{G}$, in which weights are functions of both the Bonacich centrality matrix for $\mathbf{G}$ and of the topology of common neighborhoods in $\mathbf{G}$. We first define the notion of weighted centrality.

Definition 3 *Let* $\mathbf{G}$ *be an* $n \times n$ *adjacency matrix, and let* $\bar{w}$ *be a* $n \times 1$ *positive vector. The weighted Bonacich centrality vector for* $\mathbf{G}$ *with parameter* a *and with weights vector* $\bar{w}$ *is defined as follows:*

$$\mathbf{b}(\mathbf{G}, a, \bar{w}) = \mathbf{M}(\mathbf{G}, a) \cdot \bar{w} \tag{13}$$

Proposition 3 *Let* $\mathbf{M}$ *the Bonacich centrality matrix* $\mathbf{M}(\mathbf{G}, \frac{\phi-\gamma_1}{\gamma_1+\gamma_0})$*. The vector of equilibrium actions in (12) is given by:*

$$\bar{x} = \frac{\alpha}{\gamma_1 + \gamma_0} \mathbf{b}(\mathbf{G}, \frac{\phi - \gamma_1}{\gamma_1 + \gamma_0}, \bar{w}) \tag{14}$$

where the vector of weights is given by:

$$\bar{w} = [I + \frac{\gamma_2}{\gamma_1 + \gamma_0} \mathbf{M}(\mathbf{G} + \mathbf{G^2})]^{-1} \cdot \bar{1}. \tag{15}$$

Behavior is still related to the original centralities in $\mathbf{G}$, but the (discounted) paths used to compute centralities are weighted in a way that depends on the number of common neighbors that agents have. We can better understand the effect of altruism on equilibrium actions by rewriting (14) as follows:

$$\bar{x} = \frac{\alpha}{\gamma_1 + \gamma_0} \mathbf{b}\left(\mathbf{G}, \frac{\phi - \gamma_1}{\gamma_1 + \gamma_0}\right) - \frac{\gamma_2}{\gamma_1 + \gamma_0} \mathbf{M}(\mathbf{G} + \mathbf{G^2})\bar{x} \tag{16}$$

The first term in (16) is the equilibrium vector of actions when $\gamma_2 = 0$, while the second term measures the correction of equilibrium behavior that is due to γ_2. The term $\mathbf{M}(\mathbf{G} + \mathbf{G^2})\bar{x}$ can be viewed as the multiplication of the Bonacich matrix $\mathbf{M}$ times a vector $\bar{z} = (\mathbf{G} + \mathbf{G^2})\bar{x}$. The generic element

$$z_i = \sum_j [(g_{ij} + g_{ij}^{[2]})x_j]$$

measures the aggregate equilibrium actions that agents j in i's neighborhood are exposed to. The correction of equilibrium behavior is therefore higher for those agents whose high centrality (in $\mathbf{G}$) comes from paths that lead to agents who are exposed to large amounts of actions in equilibrium. Intuitively, given convexity of damages, such agents are those with larger associated marginal costs due to the negative externality.

Example [Ballester et al. (2006)] Consider the network used by Ballester et al. (2006).

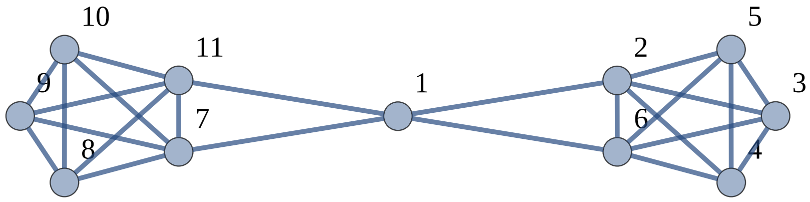

Figure 1: From Ballester et al. (2006).

There are basically three types of agents in this network, that we refer to as types 1, 2 and 3, from the names of the corresponding representative nodes. Figure 2 records the equilibrium action of these types (x_1, x_2, x_3) as a function of the parameter γ_2 within a given interval.

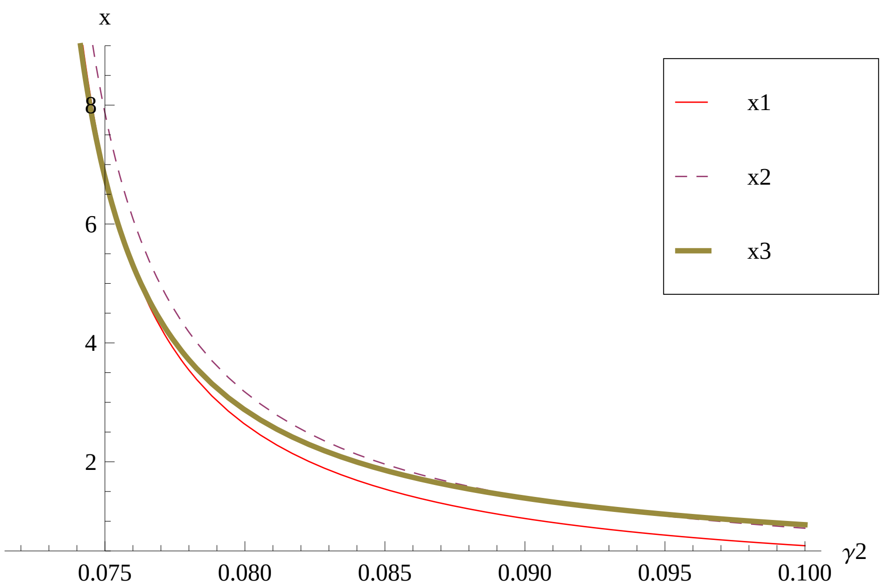

Figure 2: Equilibrium actions with varying degrees of altruism ($\gamma_0 = 0, \gamma_1 = 0.5, \alpha = 0.6, \phi = 1$).

Figure 2 makes clear that increases in γ_2 may have a strong impact on the magnitude and the ranking of centralities and equilibrium actions. In particular, the most central agent is type 2 for low levels of γ_2, and switches to type 3 for higher levels of γ_2. Note that while type 2 is more connected, and therefore more central when γ_2 is small, an equilibrium where type 2's action is high is not sustainable when γ_2 is high. This happens because in such equilibrium agent 1, with whom type 2 agents are linked, would be exposed to large amount of externalities, and this would substantially increase the marginal cost of type 2. This causes a switch to an equilibrium in which type 3, more peripheral in the network **G**, chooses the largest action. The switch between type 1 and 3, occurring at low levels of γ_2, is explained along similar lines.

The next proposition formally qualifies the shift of behavior towards less central agents in the network **G** we observed in the above example.

Proposition 4 *The marginal effect of γ_2 on equilibrium behavior is proportional to the effect of a marginal increase in ϕ on the Bonacich centralities of the network* **G***, weighted by the nodes' degrees. Formally:*

$$\frac{\partial \bar{x}^*}{\partial \gamma_2} = -\frac{\partial}{\partial \phi}\bar{b}(\mathbf{G}, \frac{\phi - \gamma_1}{\gamma_1 + \gamma_0}, d), \tag{17}$$

Moreover, the magnitude of the effect of γ_2 increases with ϕ.

Proposition 4 provides insights on the effect of γ_2 on the vector of equilibrium actions, and in particular on the ranking of its entries. We first note that if the network is almost regular (i.e, degrees vary little across nodes), the effect of γ_2 on behavior is essentially proportional to the effect of ϕ. We know from Ballester et al. (2006) that this effect, alone, potentially alters the ordering of agents' equilibrium actions. When agents' degrees vary substantially, the ranking is also affected by the weighting vector. In particular, the decrease in action is the largest for those agents who have a large discounted sum of paths to nodes with a large degree in **G**. This implies that those agents whose high centrality in **G** is due to paths towards highly connected agents will experience a large decrease in behavior as a consequence of γ_2. This qualifies the generic intuition we had from previous examples: marginalization of behavior as a result of γ_2 occurs when very central agents in **G** are characterized by a large number of paths towards agents with large degree; from (17), these agents will suffer the largest decrease in behavior as a result of an increase in γ_2. This can be observed in the three different networks in figure 3 (equilibrium choices are reported in table 4). These are non regular networks with 5 nodes, each obtained by increasing the number of connections between peripheral agents, starting from a star.

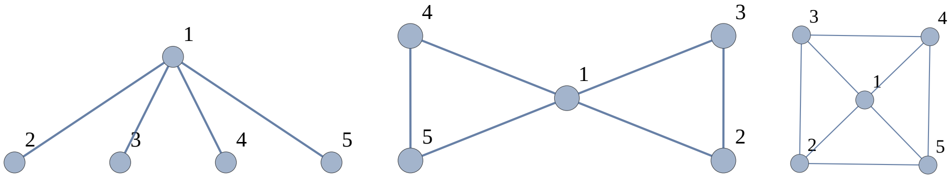

Figure 3: Star, Papillon and Connected Star

When $\gamma_2 = 0$, agent 1 is the most central in all networks. Note also that the three networks differ with respect to the degree of agent's 1 neighbors. Proposition 4 implies that the impact of γ_2 on agent 1's behavior is stronger when agent 1's neighbors are more connected. This is indeed what we observe: in the papillon and the connected star, agent 1 reduces by a large amount her equilibrium action, and the actions' ordering in equilibrium are reversed as γ_2 increases, while this does not occur in the star network.

Table 1: Effect of altruism on equilibrium actions ($\gamma_0 = 0$, $\gamma_1 = 0.9$, $\phi = 1$, $\alpha = 2$)

Network	Players	$\gamma_2 = 0$	$\gamma_2 = 0.06$
Star	1	3.38	2.01
	2-5	2.59	1.82
Papillon	1	3.53	1.63
	2-5	2.95	1.69
Connected Star	1	3.73	1.34
	2-5	3.39	1.45

4 Network-Based Policy

In this section we wish to assess the implication of the indirect strategic interaction discussed above for various kinds of network based policies considered in the literature. We will look at policies that affect the structure of the network by either changing the number and the pattern of connections or by deleting key nodes from the network. We mainly refer to an interpretation of the parameter γ_2 in terms of altruism that agents have towards their neighbors. This in order to ease the exposition of the main ideas, and because of the prevalent importance of network based policy in the context of interpersonal networks (see, for instance the work by Ballester et al. (2010) on criminal networks and by Christakis and Fowler (2007) and Christakis and Fowler (2008) on health related behavior).

4.1 Changing the Network

4.1.1 Network Density

Let us first look at changes in the number of links in the network. To get a first rough intuition, let us compare behavior in the three networks of figure 3. For each network, we consider the sum of equilibrium actions, first with $\gamma_2 = 0$ and then with $\gamma_2 > 0$.

Table 2: Effect of altruism on aggregate actions ($\gamma_0 = 0$, $\gamma_1 = 0.9$, $\phi = 1$, $\alpha = 2$)

Network	$\gamma_2 = 0$	$\gamma_2 = 0.06$
Star	13.77	9.31
Papillon	15.29	8.42
Connected Star	17.29	7.17

We see that, while the introduction of altruism does not affect the ranking of individual actions within each network, altruism reverts the trends in individual and aggregate behavior as we add connections. Without altruism, increased connections imply increased individual and aggregate behaviors (as a pure effect of increased complementarities), while with altruism both individual and aggregate behaviors decrease with connectivity. The simultaneous creation of distance-two interactions (of a strategic substitute type) in the network $\tilde{\mathbf{G}}$ decreases the incentives to act, although the centrality of all agents in the network $\mathbf{G}$ is increased by the new connections,

We can perform a more systematic analysis of the effect of network density on behavior by focusing on the class of regular networks, in which density is proportional to the common degree d. Here, equilibrium behavior is characterized by the following first order condition for each agent i (a simplified version of condition (3)):

$$\alpha - \sigma_i x_i + (\phi - \gamma_1 - \gamma_2)\sum_{j\in N} g_{ij}x_j - \gamma_2 \sum_{k\neq i} g_{ik}^{[2]} x_k = 0. \tag{18}$$

In a symmetric equilibrium, $x_i^* = x_j^*$ for all i, j. Moreover, in a regular graph of degree d, $\sum_{k\in N} g_{ik}^{[2]} = d(d-1)$. Using the expression for the term σ_i, we can rewrite (18) as follows:

$$\alpha - x^*[\gamma_1 + \gamma_0 - d(\phi - \gamma_1 - \gamma_2) + \gamma_2 d^2] = 0. \tag{19}$$

When $[\gamma_1 + \gamma_0 - d(\phi - \gamma_1 - \gamma_2) + \gamma_2 d^2] < 0$, no positive action is consistent with equilibrium (a simplified version of the constraint on the largest Eigenvalue of the matrix $\mathbf{G}$, which in regular network coincides with the degree). When $[\gamma_1 + \gamma_0 - d(\phi - \gamma_1 - \gamma_2) + \gamma_2 d^2] > 0$, the unique positive symmetric equilibrium is given by:

$$x^* = \frac{\alpha}{\gamma_1 + \gamma_0 - d(\phi - \gamma_1 - \gamma_2) + \gamma_2 d^2}. \tag{20}$$

The effect of network density on behavior is measured by the first derivative of (20) with respect to d:

$$\frac{\partial x^*}{\partial d} = \frac{\alpha(\phi - \gamma_1 - \gamma_2 - 2d\gamma_2)}{[\gamma_1 + \gamma_0 - d(\phi - \gamma_1 - \gamma_2) + \gamma_2 d^2]^2}.$$

The sign of the effect of density on behavior is determined by the following regions:

$$\begin{cases} d < \frac{\phi-\gamma_1-\gamma_2}{2\gamma_2} & \Rightarrow & \frac{\partial x^*}{\partial d} > 0 \\ d = \frac{\phi-\gamma_1-\gamma_2}{2\gamma_2} & \Rightarrow & \frac{\partial x^*}{\partial d} = 0 \\ d > \frac{\phi-\gamma_1-\gamma_2}{2\gamma_2} & \Rightarrow & \frac{\partial x^*}{\partial d} < 0 \end{cases}$$

Note that when $\gamma_1 + \gamma_2 > \phi$, behavior is always decreasing with network density and approaching zero for very large degrees. When instead $\gamma_1 + \gamma_2 < \phi$, equilibrium behavior follows a non monotonic pattern, reaching a maximum for $d^{max} = \frac{\phi-\gamma_1-\gamma_2}{2\gamma_2}$. Before d^{max}, behavior increases with network density as a result of the prevailing force of the net peer effect; after d^{max}, behavior monotonically decreases as a result of altruism, and tends to zero for very large degrees. Note also that, if $\gamma_1 < \phi$, then d^{max} is always decreasing in both γ_1 and γ_2. This leads to the following:

Proposition 5 *Let* $\mathbf{g}$ *be a regular network with identical agents. When* $\gamma_1 + \gamma_2 \geq \phi$*, equilibrium behavior always decreases with the degree. When* $\gamma_1 + \gamma_2 < \phi$*, equilibrium behavior is a non monotonic function of the degree, increasing for low degrees (*$d < d^*$*) and decreasing for high degrees (*$d > d^*$*). The threshold* d^* *is decreasing with* γ_2*, and* $d^* \to \infty$ *when* $\gamma_2 \to 0$*.*

While in the absence of altruism the "social multiplier" associated with the net peer effects generates a positive and monotonic relation between network density and behavior, this is not the case with altruism. The non monotonic relation is due to the presence of direct and indirect strategic interactions in the network. A larger degree affects behavior through the growth of an agent's direct neighbors (complements) and through the growth of neighbors of distance two (substitutes). Direct connections grow with d, while distance-two connections grow with d^2, possibly taking over and causing a decrease in overall behavior. Figure 4 gives a graphical representation of the relationship between degree and behavior for different levels of γ_2.

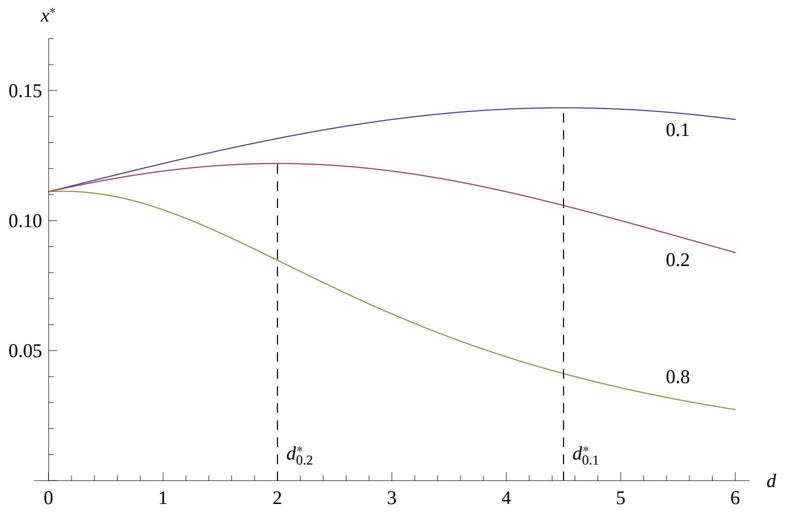

Figure 4: Degree and behavior: $\gamma_2 \in \{0.1,\ 0.2,\ 0.8\}$, $d^*_{0.2} \equiv d^*|_{\gamma_2=0.2}$, $d^*_{0.1} \equiv d^*|_{\gamma_2=0.1}$, $\phi = 1, \alpha = 2, \gamma_0 = 0, \gamma_1 = 0.3$

4.1.2 Adding and Severing Links

Policies that focus on decreases in behavior are of particular interest in the present setting of negative externalities. Here we ask whether we can alter the structure of a generic network to decrease the overall incentives of agents to act. The following result, due to Ballester et al. (2006), rigorously states the intuitive idea that increasing channels of complementarities increase aggregate behavior $\mathbf{x}^*$.

Theorem 1 (Ballester et al. (2005)) *Let $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{G}}'$ be symmetric and such that $\tilde{g}_{ij} \geq \tilde{g}'_{ij}$ for all i, j and $\tilde{g}_{ij} > \tilde{g}'_{ij}$ for at least one ij. If $\frac{\eta(\gamma_1+\gamma_0)}{\gamma_2} > \lambda\mu(\mathbf{C})$ and $\frac{\eta(\gamma_1+\gamma_0)}{\lambda\gamma_2} > \mu(\mathbf{C}')$, then $\mathbf{x}^*(\tilde{\mathbf{G}}) > \mathbf{x}^*(\tilde{\mathbf{G}}')$.*

In the present context, the relevant policy problem is affecting the network $\mathbf{G}$ to obtain the desired ordering on the induced network $\tilde{\mathbf{G}}$. The following proposition finds sufficient conditions under which this can be done.

Proposition 6 *Consider $\rho \in \mathbb{N}$ such that $(\phi - \gamma_1 - \gamma_2) \leq \rho\gamma_2$. Consider network $\mathbf{G}'$ obtained from $\mathbf{G}$ by fully connecting an independent set of nodes Z in $\mathbf{G}$, and such that $|Z| = \rho + 2$. Then $\tilde{\mathbf{G}}' < \tilde{\mathbf{G}}$ and $\bar{x}(\tilde{\mathbf{G}}') < \bar{x}(\tilde{\mathbf{G}})$.*

A sufficient condition to reduce aggregate behavior is therefore the presence of sparse sets of agents unconnected in network $\mathbf{G}$. The number of such individuals is inversely related to the degree of altruism. Behavior is reduced by creating very dense relations among these agents, so that new direct ties come together with enough new indirect ones. Note how this result differs from the case without altruism, where second order effects are null and the creation of clustered communities unambiguously increases behavior. As a final remark, note that a converse argument also applies. Since clustered communities magnify the second order effects, the effect of disconnecting such communities is always to increase behavior.

4.2 Key Players and Policy Targets

One important class of network based policies is the identification of players (or groups of players) who, if targeted, would trigger a maximal change in aggregate behavior. Such *key-players* are of crucial importance in various health related policies and in policies that try to reduce crime (see Ballester et al., 2010). Ballester et al. (2006) define the key-player as the node of the network whose removal produces the largest reduction in aggregate behavior, and show that the key-player is the node with the highest *intercentrality* in the network.

Theorem 2 (Ballester et al. 2006) *If* $\frac{\eta(\gamma_1+\gamma_0)}{\lambda\gamma_2} > \mu(\mathbf{C})$*, the key player is the agent with the highest intercentrality index, measured by* $c_i = b_i^2/m_{ii}$*.*

As for the notion of Bonacich centrality, the ordering of intercentralities is potentially affected by the degree of altruism. In particular, the same marginalization of central players we observed in section 3.2 seems to characterize the intercentrality ordering. Consider, for instance, the case of a "line" network (figure 5). We can identify 3 types of agents depending on their position:

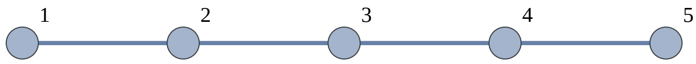

Figure 5: Line network

type A (agents 1 and 5), type B (agents 2 and 4) and type C (agent 3). Table 3 provides the ordering of centralities and intercentralities for different values of γ_2.

Table 3: Key Player - Line network

γ_2	b_i	c_i
0	$C > B > A$	$C > B > A$
0.05	$A > B > C$	$A > B > C$

Parametrization: $\phi = 1$, $\gamma_1 = 0.9$, $\alpha = 2$

With no altruism, the key-player is type C. Consider now altruism. Type C is responsible for several distance-two relations, that keep aggregate behavior low by creating strategic substitutability. Removing type C from the network has therefore little effect in decreasing behavior. Type A, in contrast, does not generate any distance-two relations, and A's removal results therefore more effective in decreasing behavior.

Inspection of the interconnected cliques network studied by Ballester at al. (2006) and in section 3 of this paper (see figure 1) shows that not only key players move towards the periphery, but also that inter-centrality and Bonacich centrality need not move together as γ_2 grows. Table 4 reports the ordering of Bonacich centralities and inter-centralities for this network.

Table 4: Key Player - Ballester et al. (2006) network

γ_2	b_i	c_i
0	$2 > 1 > 3$	$2 > 1 > 3$
0.002	$2 > 1 > 3$	$2 > 3 > 1$
0.01	$2 > 3 > 1$	$2 > 3 > 1$
0.02	$3 > 2 > 1$	$3 > 2 > 1$

Parametrization: $\phi = 1$, $\gamma_1 = 0.95$, $\alpha = 2$

For low but positive levels of γ_2, type 3 agents have higher inter-centrality than type 1 agent, but are less central. In fact, while agent 1 is critical for several two-distance relations, type 3 agents are not. It follows that removing agent 1 would remove several sources of substitutability, thereby offsetting the negative effect on behavior due to the removal of direct connections. For larger levels of γ_2, the two measures go alongside and produce the same ranking.

While the above notion of key-player rests on the assumption that agents can be removed from the network, policies are often based on measures that affect agents' incentives without removing them from society. Suppose, for instance, that the policy maker aims at lowering the private benefit from agents' own actions (that is, lowering the parameter α in our model). The following definition identifies the agent whose marginal reduction in such incentives would bring about the largest reduction in aggregate behavior.

Definition 4 *The α-key player is the agent i such that $\frac{\partial \mathbf{x}}{\partial \alpha^i}$ is maximal.*

To characterize the position of the α-key player, we need to reformulate the equilibrium characterization for the case of heterogenous α's. Following Calvó-Armengol et al. (2009), the vector of equilibrium actions is.

$$\bar{x} = \eta \mathbf{b}(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)}, \alpha) - \frac{\gamma_2 \theta b(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)}, \alpha)}{(\gamma_1+\gamma_0)[\eta(\gamma_1+\gamma_0) + \gamma_2\theta b(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)})]} \mathbf{b}(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)}) \quad (21)$$

Proposition 7 *The α-key player is the agent with the highest Bonacich centrality in the network $\mathbf{C}$.*

The difference between the key-player, with highest intercentrality, and the α-key player, with highest Bonacich centrality, has to do with the different effect of policies that remove a node compared to policies that affect a node's behavior. Intercentrality corrects centrality by "neglecting" those paths of complementarities connecting a node to itself. This happens because once the key-player is removed from the network, the effect of the policy on her own action is not considered. The α-key player, in contrast, affects aggregate behavior before and after the policy intervention. For this reason, the reflection of players' action on their own incentives to act due to the network's complementarities is important, and Bonacich centrality, rather than inter-centrality, matters.

4.3 Policies that affect segregation

When agents are heterogeneous along some preference dimension, both aggregate behavior, its distribution across individuals and societal welfare may depend on the extent to which social interaction is segregated. In such cases, policies can increase welfare by affecting the patterns of interaction between heterogeneous agents. In particular, policies may reduce segregation by moderating the effects of homophily, or induce segregation by clustering agents with similar habits (e.g.smoking bans and smoking areas). In this section we look at the effect of such policies on behavior in the simple case of regular networks. Heterogeneity is captured by the terms α_i, which may now differ across agents. The degree of segregation in a given regular network with degree d is instead captured by the parameter q, common to all agents, and measuring the fraction of an agent's neighbors that are of the same type as she is. For simplicity, we will also assume that agents come in two typologies: those with high preferences α^H for the action and those with low preferences α^L, with $\alpha^H > \alpha^L$. Populations of the two types are assumed of equal sizes.

The type-symmetric equilibrium levels for types H and L are (see appendix C for derivations):

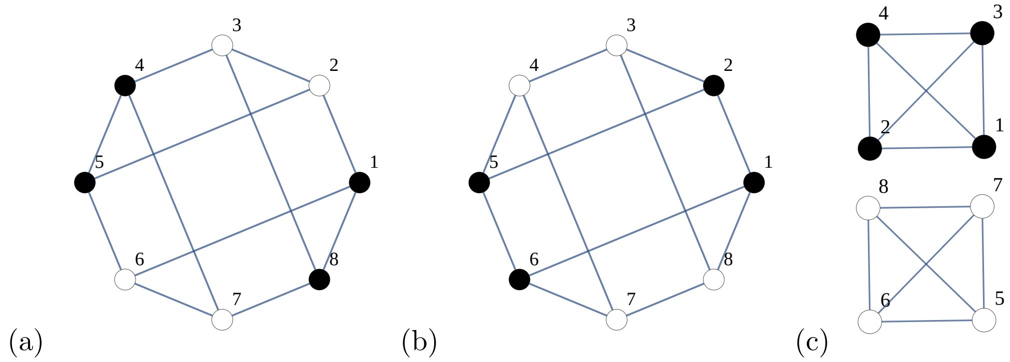

Figure 6: Three networks with increasing degrees of segregation.

$$\begin{cases} x^H = \frac{1}{2}\left[\frac{\alpha^H+\alpha^L}{\gamma_1+\gamma_0-d(\phi-\gamma_1-\gamma_2)+\gamma_2 d^2} + \frac{\alpha^H-\alpha^L}{\gamma_1+\gamma_0+\gamma_2 d^2(1-2q)^2+d(1-2q)(\phi-\gamma_1-\gamma_2)}\right] \\ x^L = \frac{1}{2}\left[\frac{\alpha^H+\alpha^L}{\gamma_1+\gamma_0-d(\phi-\gamma_1-\gamma_2)+\gamma_2 d^2} + \frac{\alpha^L-\alpha^H}{\gamma_1+\gamma_0+\gamma_2 d^2(1-2q)^2+d(1-2q)(\phi-\gamma_1-\gamma_2)}\right] \end{cases} \tag{22}$$

Equilibrium behavior of each type is the sum of two terms. The first common term coincides with the equilibrium behavior if all agents had preferences $\frac{\alpha_H+\alpha_L}{2}$. The second term measures how types' actions are spread around this mean. Symmetry of the spread implies that the average behavior is not affected by q. In the example of Figure 7, the spread is increasing for low q and decreasing for high q, reaching its maximum at an intermediate level of segregation $\bar{q} = \frac{\phi-\gamma_1-\gamma_2(1-2d)}{4d\gamma_2} > \frac{1}{2}$. As proposition 8 shows, this non monotonic relation occurs in the range of degrees for which the average action is a decreasing function of the degree:

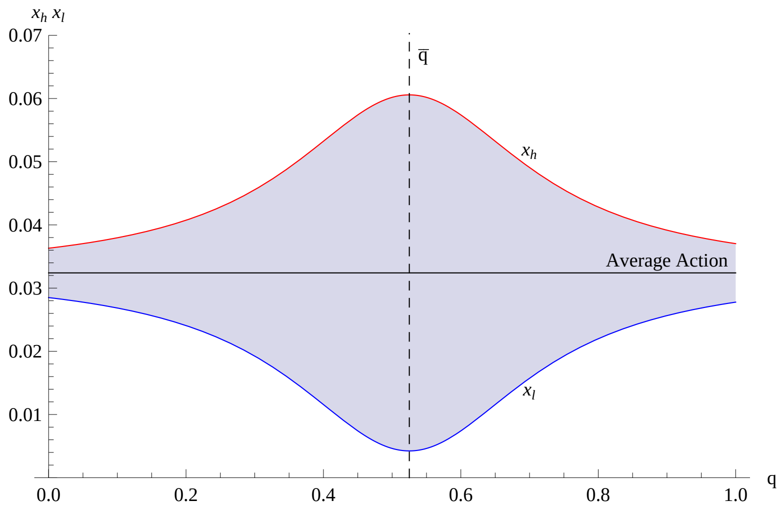

Figure 7: Spread: Parametrization: $\gamma_0 = 0, \gamma_1 = 0.85$, $\gamma_2 = 0.1$, $\phi = 1$, $\alpha^H = 1.5$, $\alpha^L = 1$, $d = 10$.

Proposition 8 *When $d < d^*$, the spread between x^H and x^L is monotonically increasing in q, where d^* is as in proposition 5. When $d > d^*$, the spread is non monotone in q, reaching its maximum at $\bar{q} \in (1/2, 1]$, where $\bar{q}$ is decreasing in γ_1, γ_2 and d. Moreover, the maximal spread is independent of the degree.*

Proposition 8 implies that increases in segregation are first followed by increased heterogeneity in behavior, and, as segregation increases, by homogenization of behavior. In a model of pure peer effects (i.e, with $\gamma_2 = 0$), segregation would unambiguously increase H types' equilibrium action and decrease L types' action. From the expression of $\bar{q}$, we also note that the presence of externalities and of convex damages cannot, alone, imply the non monotone relation between segregation and spread, since $\bar{q} > 1$ for very low levels of γ_2.

The crucial role of altruism for this non monotonicity becomes clear once we consider the forces at work as q increases. At low levels of q, H agents are mainly surrounded by L agents and *viceversa*. Thus, given that H agents always choose higher actions than those chosen by L agents, the stock of externality experienced by H agents from their neighbors is smaller than the one experienced by L agents. An increase in q has the effect of replacing L agents with H agents in the neighborhood of H agents. This naturally tends to drive H actions up *via* the net peer effect. For low q's, the lower stock experienced by the new H neighbors compared to the replaced L agents reinforces the peer effect in driving the H action up. As q increases, H agents tend to have more and more H neighbors, and therefore to be recipients of larger and larger stocks of externality compared to L agents. For large enough q, replacing L neighbors with H neighbors increases the overall marginal damage for H agent (this occurring sooner the larger γ_2), and when this outweigh the peer effect, H actions start decreasing. Key to the above argument is the fact that, while peer effects apply to flows of individual actions, altruism applies to stocks of actions in each agent's neighborhood. While H actions always exceed L actions, the stock in H neighborhoods is smaller for small q and larger for large q's than the stock in L neighborhoods. Hence the non monotonicity result follows.

5 Implications for empirical work on peer effects

In this final section we wish to discuss how the explicit consideration of externalities and altruism modifies the procedure for the estimation of peer effects in social networks proposed in the recent econometric literature on the subject (Bramoullé et al., 2009; Lee et al., 2010; Liu and Lee, 2010; Liu et al., 2012).[1] We start by considering the FOC in (2), allowing for a possibly heterogeneity in α_i:

$$\alpha_i - \sigma_i x_i + (\phi - \gamma_1 - \gamma_2) \sum_{j \in N} g_{ij} x_j - \gamma_2 \sum_{k \neq i} g_{ik}^{[2]} x_k = 0 \tag{23}$$

where, as previously defined, $\sigma_i = \gamma_0 + \gamma_1 + \gamma_2 d_i$. Letting $\sigma = \gamma_0 + \gamma_1$, (23) can be rewritten as follows:

$$\alpha_i - \sigma x_i + (\phi - \gamma_1 - \gamma_2) \sum_{j \in N} g_{ij} x_j - \gamma_2 \sum_{k \in N} g_{ik}^{[2]} x_k = 0, \tag{24}$$

where α_i accounts for a set of observable personal characteristics (z_i),[2] average friends' characteristics $(\frac{1}{d_i} \sum g_{ij} z_i)$ and a random error term ϵ. The FOC to be estimated can now be written as:

$$x_i = \frac{\theta}{\sigma} z_i + \frac{\kappa}{\sigma} \frac{1}{d_i} \sum_{j \in N} g_{ij} z_i + \frac{(\phi - \gamma_1 - \gamma_2)}{\sigma} \sum_{j \in N} g_{ij} x_j - \frac{\gamma_2}{\sigma} \sum_{k \in N} g_{ik}^{[2]} x_k + \epsilon. \tag{25}$$

Note that the parameters are identified up to a normalization, since every coefficient is divided by a factor σ measuring the concavity of agent's utility function. Which parameters we

[1] To keep the model simple and to make it comparable to the previous literature we do not include in the specification neither a constant term nor a network fixed effect

[2] Assume without loss of generality that we include in the model just one demographic characteristic.

are able to identify will depend on which (nested) model we want to estimate, that is whether or not γ_0 and γ_1 are different from zero. In particular, under the appropriate identification conditions, to be discussed in section 5.2, when $\gamma_0 = 0$, the model identifies both $\frac{\phi}{\gamma_1}$ and $\frac{\gamma_2}{\gamma_1}$, capturing the pure net peer effects and the degree of altruism.

Define $\mathbf{G}^*$ to be the row normalized matrix $\mathbf{G}$, with $g^*_{ij} = \frac{1}{d_i} g_{ij}$. As previously defined, $\mathbf{G^2}$ is the matrix counting the number of two-distance walks between agents, with diagonal terms d_i. Calling $\rho = \frac{\theta}{\sigma}$, $\zeta = \frac{\kappa}{\sigma}$ we obtain the following matrix form specification.

$$x = \beta_1 \mathbf{G}x + \beta_2 \mathbf{G^2} x + \rho z + \zeta \mathbf{G}^* z + \epsilon \tag{26}$$

The action x is determined by the sum of the actions chosen by peers ($\mathbf{G}x$), the actions chosen by two-distance neighbors ($\mathbf{G^2}x$), own demographics (z), own neighbors' demographics ($\mathbf{G}^*z$) and a random error term ϵ.

The next subsections are structured as follows. In 5.1 we characterize the biases, due to externalities and altruism, that arise in the estimation of peer effects, in 5.2 we derive new conditions for identification of the model with altruism. Finally, in 5.3, we present the optimal set of instruments.

5.1 Bias in the estimation of peer effects

Even when altruism is not present, the value of the parameter β_1 can be interpreted as the peer effect modified by the effect of externalities. Such a value would be smaller than the one we would obtain if externalites did not play any role.

Let us now consider the case of altruism ($\gamma_2 > 0$), and its effect for the estimation of $\hat{\beta}_1$. Suppose that we do not include $\mathbf{G^2}x$ in the model in equation (26) (as in Bramoullé et al., 2009; Lee et al., 2010; Liu and Lee, 2010), thus estimating the following:

$$x = \beta_1 \mathbf{G}x + \rho z + \zeta \mathbf{G}^* z + \epsilon \tag{27}$$

Using the usual omitted variable bias formula (see, for example, Angrist and Pischke, 2008), the coefficient of the peer effect $\hat{\beta}_1$ can be written as the sum of the real effect β_1 and a bias, derived from the correlation between the omitted variable $\mathbf{G^2}x$ and the included explanatory one $\mathbf{G}x$:

$$\frac{Cov(x, \mathbf{G}x)}{Var(\mathbf{G}x)} = \beta_1 + \beta_2 \delta_{\mathbf{G^2}x, \mathbf{G}x} \tag{28}$$

where $\delta_{\mathbf{G^2}x, \mathbf{G}x}$ is the coefficient from a regression of $\mathbf{G^2}x$ on $\mathbf{G}x$. The theoretical model suggests the patterns of substitutability and complementarity between the actions of the agents in the network. In particular, we expect that altruistic agents decrease their equilibrium choices when the stock of negative externalities their friends are exposed to increases. For such a reason we expect the coefficient associated to second order neighbors β_2 to be negative (see equation 25). However, the quantities chosen by friends and by second order neighbors are between them strategic complements (because of peer effects) and thus $\delta_{\mathbf{G^2}x, \mathbf{G}x}$ is positive. Thus, the omitted variable bias $\beta_2 \delta_{\mathbf{G^2}x, \mathbf{G}x}$ is always negative and the peer effects in (27) systematically underestimated. Moreover, the larger the complementarities between first and second order neighbors' choices, the larger the bias.

5.2 Identification

As shown by Manski (1993), identification in a model with peer effect is difficult due to the reflection problem. However, when networks are not complete so that people do not interact in groups and data on the network interaction is available, identification can be achieved under some conditions. This section expands the results provided in Bramoullé et al. (2009) by looking at the identification issue in a model with and without altruism, by focusing on two relevant cases: $\mathbf{G} = \mathbf{G}^*$ and $\mathbf{G} \neq \mathbf{G}^*$. The first case is the one that received most theoretical attention. In particular, the empirical peer effect literature starting from Manski (1993) considers peer effect to be the result of the average behavior around each agent (see Bramoullé et al., 2009; Lee et al., 2010). In this framework both neighbors' behaviors and their characteristics influence the agents by means of the same row-normalized social interaction matrix. The recent theoretical literature on peer effects, in contrast, considers the aggregate of neighbors' actions as the source of local complementarities. Thus, if we assume that neighbors' aggregate personal characteristics influence agents' choices, again $\mathbf{G} = \mathbf{G}^*$. This case, in which $\mathbf{G} = \mathbf{G}^*$ and the matrix is not row normalized, is studied by Liu and Lee (2010). The case in which $\mathbf{G} \neq \mathbf{G}^*$, i.e. peer effects result from aggregate neighbors' behavior (as in our theoretical model, i.e. $\mathbf{G}$ not row-normalized), while the average of neighbors' demographic matters (i.e. $\mathbf{G}^*$ is row-normalized) does not fall in the previous two categories, and is explored in Case 2.

Case 1. $\mathbf{G} = \mathbf{G}^*$

Proposition 9 (Bramoullé et al. (2009)) *If there is no altruism, the model in (27) identifies* β_1*, if* $\zeta + \beta_1\rho \neq 0$ *and* $\mathbf{I}, \mathbf{G}$ *and* $\mathbf{G^2}$ *are linearly independent.*

This sufficient condition states that, when demographics have some explanatory value ($\zeta + \beta_1\rho \neq 0$), the peer effect cannot be identified in fully connected networks. Consider now the case with altruism in (26):

Proposition 10 *Let* $\mathbf{G} = \mathbf{G}^*$*. If* $\mathbf{I}, \mathbf{G}, \mathbf{G^2}, \mathbf{G^3}$ *are linearly independent, the net peer effect* β_1 *and the effect of altruism* β_2 *in (26) are identified if* $\beta_1\zeta + \frac{\zeta^2}{\rho} + \beta_2\rho \neq 0$ *and* $\rho \neq 0$.[3]

Notice that, with respect to the previous case with no altruism, when $\beta_2 \neq 0$ more restrictive conditions are required in order to identify parameters β_1 and β_2. The conditions on the networks are the same found by Bramoullé et al. (2009) in a model without altruism and with network fixed effects.

Case 2. $\mathbf{G} \neq \mathbf{G}^*$

Consider the model without altruism in (27). Adapting to our framework the results from Bramoullé et al. (2009), the following holds:

Proposition 11 *Consider the model (27) and let* $\mathbf{G} \neq \mathbf{G}^*$*. If* $\mathbf{I}, \mathbf{G}, \mathbf{G}^*, \mathbf{G}\mathbf{G}^*$ *are linearly independent and if* $\rho \neq 0$ *or* $\zeta \neq 0$*, the net peer effect* β_1 *is identified .*

It is now important to identify which classes of networks are ruled out by the above sufficient conditions for indentification. Note first that $\mathbf{I}$ is linearly dependent with $\mathbf{G}$ and also with $\mathbf{G}^*$ only in the empty network. Note then that $\mathbf{G}$ and $\mathbf{G}^*$ are linearly dependent only in regular networks, where d is the common degree and $\mathbf{G} = \frac{1}{d}\mathbf{G}^*$. Finally, let us consider when $\mathbf{G}\mathbf{G}^*$ is linearly independent from both $\mathbf{G}$ and $\mathbf{G}^*$. Since $\mathbf{G}\mathbf{G}^*$ keeps track of weighted distance-two paths, while $\mathbf{G}$ and $\mathbf{G}^*$ just consider distance-one neighbors, a necessary condition for linear dependence is that all triangles in $\mathbf{G}$ close, leading to the complete network. However, a complete network has already been excluded as it belongs to the class of regular networks. It follows that in order to identify the peer effect in model (27), all regular networks must be excluded.

Let us then consider the model with altruism in equation (26).

[3]Note that this condition can be written as $\zeta + \beta_1\rho + \frac{\beta_2\rho^2}{\zeta}$, which is a modification of the one found by Bramoullé et al. (2009)

Proposition 12 *Consider model (26) and let* $\mathbf{G} \neq \mathbf{G}^*$. *If* $\mathbf{I}, \mathbf{G}, \mathbf{G^2}, \mathbf{G}^*, \mathbf{GG}^*, \mathbf{G^2G}^*$ *are linearly independent and if* $\rho \neq 0$ *or* $\zeta \neq 0$*, the peer effect* β_1 *and the effect of altruism* β_2 *are identified.*

The introduction of altruism restricts the set of networks that enable the identification of both peer effect and altruism. In fact, together with regular networks, other classes of networks must be excluded. These include the star, where $\mathbf{G^2G}^* = (n-1)\mathbf{G}^*$, and the network in figure 3b, in which $\mathbf{GG}^* = \frac{1}{2}\mathbf{G^2}$.

5.3 Choice of the instruments

Let us write the complete model to be estimated as follows:

$$x_r = \beta_1 \mathbf{G_r} x_r + \beta_2 \mathbf{G_r^2} x_r + \mathbf{Z_r}^* \delta + \epsilon_r \tag{29}$$

where r is the number of networks in the dataset, nr the number of individuals in the network, $x_r = (x_{1,r}, ..., x_{nr,r})'$, $z_r = (z_{1,r}, ..., z_{nr,r})'$, ϵ_r=($\epsilon_{1,r}$,......, $\epsilon_{nr,r}$), $\mathbf{Z}_r^* = (z_r, \mathbf{G}^* z_r)$ and $\delta = (\zeta, \rho)'$, β_1 captures the peer effect, β_2 the effect of altruism.

Defining $\mathbf{M}_{1r} = [I - \beta_1 \mathbf{G_r}]^{-1}$, it is easy to see that the variables $\mathbf{G_r} x_r$ and $\mathbf{G_r^2} x_r$ are endogenous because they are the result of the same maximization process.

Following Liu et al. (2012), we derive the explicit expression for the two endogenous variables (see Appendix B), that can be rewritten as follows:

$$\begin{aligned} E(\mathbf{G}x_r) = \rho \mathbf{G_r} \sum_{j=0}^{\infty} [\beta_2 \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j z_r + \rho\beta_1 \sum_{j=0}^{\infty} [\beta_2 \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r^2} z + \\ \zeta \sum_{j=0}^{\infty} [\beta_2 \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r} \mathbf{G_r^*} z_r \end{aligned} \tag{30}$$

$$\begin{aligned} E(\mathbf{G^2}x_r) = \rho \mathbf{G_r^2} \sum_{j=0}^{\infty} [\beta_2 \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j z_r + \rho\beta_1 \sum_{j=0}^{\infty} [\beta_2 \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r^3} z_r \\ + \zeta \sum_{j=0}^{\infty} [\beta_2 \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j \sum_{j=0}^{\infty} (\beta_1 \mathbf{G_r})^j \mathbf{G_r^2} \mathbf{G_r^*} z_r \end{aligned} \tag{31}$$

The endogenous variable $\mathbf{G}x_r$ is still correlated with $\mathbf{G}z_r, \mathbf{G^2}z_r, \mathbf{GG}^*z_r$ (and some higher terms) used in Liu et al. (2012). In addition, $\mathbf{G^2}x_r$ is correlated with $\mathbf{G^2}z_r, \mathbf{G}^3z_r, \mathbf{G^2G}^*z_r$ (and some higher terms) but not with $\mathbf{G}z_r$. Given that both $\mathbf{G_r}x_r$ and $\mathbf{G_r^2}x_r$ are endogenous variables, the rank condition valid for identification is modified with respect to the case in which just one endogenous variable is present. Call now W the total set of exogenous variables, i.e. exogenous variables Z included in the model (demographics and friends characteristics) and instruments Q, and V the set of all explanatory variables, i.e. W and the endogenous $\mathbf{G}x_r$ and $\mathbf{G}^2x_r$. Thus, the usual rank condition can be split in two parts: (Wooldridge, 2002):

1. rank $E(W'W) = l$
2. rank $E(W'V) = k$.

Where l is the number of exogenous variables W and k the total number of the explanatory variables. Notice that identification is not achieved if the fitted values of the first stages $\hat{\mathbf{G}x_r}$ and $\mathbf{G}\hat{^2}x_r$ are perfectly collinear. Write both $\hat{\mathbf{G}x_r}$ and $\mathbf{G}\hat{^2}x_r$ as a linear combination of two instruments ($Q_1 = \mathbf{G}z_r$ and $Q_2 = \mathbf{G}^2z_r$) multiplied by the coefficents obtained in the two (different) first stages, $\hat{\mathbf{G}x_r} = b_1Q_1 + b_2Q_2$ and $\hat{\mathbf{G}^2x_r} = c_1Q_1 + c_2Q_2$. If $c_1 = 0$ after controlling for Q_2, and if b_1 and b_2 are both different from zero, then the fitted values of the two endogenous variables cannot be perfectly collinear and $\hat{\beta}_1$ and $\hat{\beta}_2$ are identified and consistent.[4] Note that the condition $c_1 = 0$ is not necessary for identification, but it just *ex ante* rules out the presence of multicollinearity in the set of instruments.

6 Conclusions

When social relations generate both peer effects and local externalities which are, to some extent, internalized by agents, the network of strategic interaction generically differs from the network of social relations. In particular, in our model agents that share neighbors end up displaying strategic substitutability in the underlying game. We have shown that this has implications for the relation between agents' centralities in social relations and agents' equilibrium behavior, inducing peripheral agents to take larger actions than central ones. We have also shown that key-players also move towards the periphery of the network, and that the optimal use of other policy instruments is affected. In terms of empirical research, externalities and altruism have been shown to have implications for the estimation of social effects, restricting the set of networks for which identification is possible, and introducing biases in the estimation. The strategic substitution between distance-two neighbors is a general feature of problems where actions accumulate into stocks of externalities, and it is a new and previously unnoticed characteristic of network models, that applies more generally than in the linear quadratic model studied in this paper. Similar mechanisms could be introduced, for instance, in the "network games" framework of Galeotti et al. (2010), where limited information on the network is assumed, but statistical information about agents at distance two is available.

References

Angrist, J. and Pischke, J. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2005). Who's who in the networks. wanted: the key player. Technical report, CEPR, Discussion Paper No. 5329.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2006). Who's who in the networks. *Econometrica*, 74(5):1403–1417.

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2010). Delinquent networks. *Journal of the European Economic Association*, 8(1):34–61.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, vol. 150(1):41–55.

Calvó-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer effects and social networks in education. *Review of Economic Studies*, 76(4):1239–1267.

Christakis, N. and Fowler, J. (2007). The spread of obesity in a large social network over 32 years. *The new england journal of medicine*, 357:370–9.

Christakis, N. and Fowler, J. (2008). The collective dynamics of smoking in a large social network. *The new england journal of medicine*, 358:2249–58.

[4] However, if Q_2 is not strong enough and Q_1 and Q_2 are not jointly relevant (i.e they are weak instruments), the estimation of $\hat{\beta}_2$ could be severely biased in small samples.

Clark, E. and Loheac, Y. (2007). ”it wasn't me, it was them!” social influence in risky behavior by adolescents. *Journal of health economics*, 26:763–784.

Evans, W., Oates, W., and Schab, R. (1992). Measuring peer group effect: a study of teenage behaviour. *The journal of political economy*, 100, No 5:966–991.

Fletcher, J. (2010). Social interaction and smoking: evidence using multiple student cohorts, instrumental variables and school fixed effects. *Health Economics*, 19:466–484.

Fowler, J. and Christakis, N. (2008). Dynamic spread of happiness in large social networks: longitudinal analysis over 20 years in the framingham heart study. *British Medical Journal*, 337.

Galeotti, A., Goyal, S., Jackson, M., Vega-Redondo, F., and Yariv, L. (2010). Network games. *Review of Economic Studies*, 77(1):218–244.

Gaviria, A. and Raphael, S. (2001). School-based peer effects and juvenile bahavior. *The review of economics and statistics*, 83(2):257–268.

Glaeser, E., Sacerdote, B., and Scheinkman, J. (2003). The social multiplier. *Journal of the European Economic Association*, 1(2-3):345–353.

Kirke, D. (2004). Chain reactions in adolescent's cigarette, alcohol and drug use: similarity through peer influence or the patterning of tioes in peer networks? *Social Networks*, 26:3–28.

Lee, L., Liu, X., and Lin, X. (2010). Specification and estimation of social interaction models with network structures. *Econometrics Journal*, 13:145176.

Liu, X. and Lee, L. (2010). Gmm estimation of social interaction models with centrality. *Journal of Econometrics*, 159(1):99–115.

Liu, X., Patacchini, E., Zenou, Y., and Lee, L. (2012). Criminal networks: Who is the key player? Technical report, Working Papers 2012.39, Fondazione Eni Enrico Mattei.

Manski, C. (1993). Identification of endogenous social effects: the reflection problem. *Review of Economic Studies*, 60:531–542.

Poutvara, P. and Siemers, L. (2008). Smoking and social interaction. *Journal of Health Economics*, 27(6):1503–1515.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Appendix A: Utility Function

In this Appendix we give all the computations necessary in order to rewrite the utility function as in equation (2).

$$U_i = \alpha_i x_i - \gamma_0 \frac{x_i^2}{2} + \phi \sum_{j \in N} g_{ij} x_i x_j - \gamma_1 \frac{1}{2} (\sum_{j \in N} g_{ij} x_j + x_i)^2 - \gamma_2 \frac{1}{2} \sum_{j \neq i} (\sum_{k \in N} g_{jk} x_k + x_j)^2 \quad (32)$$

Now

$$(\sum_{j \in N} g_{ij} x_j + x_i)^2 = x_i^2 + (\sum_{j \in N} g_{ij} x_j)^2 + 2 \sum_{j \in N} g_{ij} x_i x_j = x_i^2 + h_{-i} + 2 \sum_{j \in N} g_{ij} x_i x_j$$

Consider the second part of (32)

$$\sum_{j \neq i} (\sum_{k \in N} g_{jk} x_k + x_j)^2 = \sum_{j \neq i} [\sum_{k \in N} g_{jk} x_k + x_j]^2 =$$

$$\sum_{j \neq i} [x_j^2 + (\sum_{k \in N} g_{jk} x_k)^2 + 2 \sum_{k \in N} g_{jk} x_j x_k] =$$

$$\sum_{j \neq i} x_j^2 + \sum_{j \neq i} (\sum_{k \in N} g_{jk} x_k)^2 + 2 \sum_{j \neq i} \sum_{k \in N} g_{jk} x_j x_k =$$

$$\sum_{j \neq i} x_j^2 + \sum_{j \neq i} (\sum_{k \in N} g_{jk} x_k)^2 + 2 \sum_{j \neq i} \sum_{k \in N} g_{jk} x_j x_k$$

Now, consider the second term of the last equation

$$(\sum_{k \in N} g_{jk} x_k)^2 = (\sum_{k \neq i} g_{jk} x_k + g_{ji} x_i)^2 = (\sum_{k \neq i} g_{jk} x_k)^2 + (g_{ji} x_i)^2 + 2 \sum_{k \neq i} g_{jk} g_{ji} x_k x_i$$

so that

$$\sum_{j \neq i} (\sum_{k \in N} g_{jk} x_k)^2 = \sum_{j \neq i} (\sum_{k \neq i} g_{jk} x_k)^2 + \sum_{j \neq i} (g_{ji} x_i)^2 + \sum_{j \neq i} 2 \sum_{k \neq i} g_{jk} g_{ji} x_k x_i =$$

$$= h_{-i} + d_i x_i^2 + 2 \sum_{k \neq i} g_{ik}^{[2]} x_i x_k$$

Note now that

$$2 \sum_{j \neq i} \sum_{k \in N} g_{jk} x_j x_k = 2 \sum_{j \neq i} g_{ji} x_j x_i + 2 \sum_{j \neq i} \sum_{k \neq i} g_{jk} x_j x_k = 2 \sum_{j \neq i} g_{ij} x_i x_j + h_{-i}$$

Thus the utility is given by

$$U_i = \alpha_i x_i - \gamma_0 \frac{x_i^2}{2} + \phi \sum_{j \in N} g_{ij} x_i x_j - \frac{1}{2} \gamma_1 [x_i^2 + h_{-i} + 2 \sum_{j \in N} g_{ij} x_i x_j] -$$

$$\frac{1}{2} \gamma_2 [\sum_{j \neq i} x_j^2 + h_{-i} + d_i x_i^2 + 2 \sum_{k \neq i} g_{ik}^{[2]} x_i x_k + 2 \sum_{j \neq i} g_{ij} x_i x_j + h_{-i}]$$

so that

$$U_i = \alpha_i x_i - \gamma_0 \frac{x_i^2}{2} + (\phi - \gamma_1 - \gamma_2) \sum_{j \in N} g_{ij} x_i x_j - \frac{1}{2} [\gamma_1 + \gamma_2 d_i] x_i^2 - \gamma_2 \sum_{k \neq i} g_{ik}^{[2]} x_i x_k + h_{-i}$$

that becomes

$$U_i = \alpha_i x_i - \frac{1}{2}\sigma_i x_i^2 + (\phi - \gamma_1 - \gamma_2)\sum_{j\in N} g_{ij}x_i x_j - \gamma_2 \sum_{k\neq i} g_{ik}^{[2]} x_i x_k + h_{-i}$$

Appendix B: Choice of the instruments

$$E(\mathbf{G}x_r) = \beta_2 \mathbf{G_r M_{1r} G_r^2} x + \rho \mathbf{G_r} z_r + \rho\beta_1 \mathbf{G_r M_{1r} G_r} z_r + \zeta \mathbf{G_r M_{1r} G_r^*} z_r \tag{33}$$

and defining $\mathbf{M_{2r}} = (\mathbf{I} - \beta_2 \mathbf{M_{1r} G_r^2})^{-1}$, we get

$$E(\mathbf{G}x_r) = \rho \mathbf{G_r M_{2r}} z_r + \rho\beta_1 \mathbf{G_r M_{2r} M_{1r} G_r} z_r + \zeta \mathbf{G_r M_{2r} M_{1r} G_r^*} z_r \tag{34}$$

substituting $\mathbf{M_{1r}} = \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j$ we get

$$E(\mathbf{G}x_r) = \rho \mathbf{G_r M_{2r}} z_r + \rho\beta_1 \mathbf{M_{2r}} \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2} z + \zeta \mathbf{M_{2r}} \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r G_r^*} z \tag{35}$$

$$E(\mathbf{G^2}x_r) = \rho \mathbf{G_r^2 M_{2r}} z_r + \rho\beta_1 \mathbf{M_{2r}} \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^3} z_r + \zeta \mathbf{M_{2r}} \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2 G_r^*} z_r \tag{36}$$

Substituting $\mathbf{M_{2r}} = \sum_{j=0}^{\infty}[\beta_2 \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j$ into (35) and (36)

$$\begin{aligned} E(\mathbf{G}x_r) = {} & \rho \mathbf{G_r} \sum_{j=0}^{\infty}[\beta_2 \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j z_r + \rho\beta_1 \sum_{j=0}^{\infty}[\beta_2 \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2} z + \\ & \zeta \sum_{j=0}^{\infty}[\beta_2 \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r G_r^*} z_r \end{aligned} \tag{37}$$

$$\begin{aligned} E(\mathbf{G^2}x_r) = {} & \rho \mathbf{G_r^2} \sum_{j=0}^{\infty}[\beta_2 \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j z_r + \rho\beta_1 \sum_{j=0}^{\infty}[\beta_2 \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^3} z_r \\ & + \zeta \sum_{j=0}^{\infty}[\beta_2 \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2}]^j \sum_{j=0}^{\infty}(\beta_1 \mathbf{G_r})^j \mathbf{G_r^2 G_r^*} z_r \end{aligned} \tag{38}$$

Appendix C: Proofs

Proof of Proposition 2. Let U denote the $n \times n$ matrix of ones. We can rewrite the first order conditions (3) as follows:

$$\alpha\bar{1} = \left[(\gamma_1 + \gamma_0)\mathbf{I} + \frac{\gamma_2}{\eta}(\theta \cdot \mathbf{U} - \lambda \mathbf{C})\right]\bar{x}.$$

Rearranging terms we get:

$$\frac{\alpha}{\gamma_1 + \gamma_0}\bar{1} - \frac{\gamma_2}{\eta(\gamma_1 + \gamma_0)}\theta \mathbf{U}\bar{x} = \left[\mathbf{I} - \frac{\gamma_2}{\eta(\gamma_1 + \gamma_0)}\lambda \mathbf{C}\right]\bar{x}.$$

By writing $\mathbf{U}\bar{x} = \mathbf{x}\bar{1}$

$$\frac{\alpha}{\gamma_1+\gamma_0}\bar{1} - \frac{\gamma_2}{\eta(\gamma_1+\gamma_0)}\theta\mathbf{x}\bar{1} = \left[\mathbf{I} - \frac{\gamma_2}{\eta(\gamma_1+\gamma_0)}\lambda\mathbf{C}\right]\bar{x}.$$

A sufficient condition for the matrix $\left[I - \frac{\gamma_2}{\eta(\gamma_1+\gamma_0)}\lambda\mathbf{C}\right]$ to admit a positive inverse is that $1 > \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)}\mu(\mathbf{C})$, with $\mu(\mathbf{C})$ being the largest eigenvalue of the $\mathbf{C}$ matrix. Under this restriction we write:

$$\left[\mathbf{I} - \frac{\gamma_2}{\eta(\gamma_1+\gamma_0)}\lambda\mathbf{C}\right]^{-1}\left(\frac{\alpha}{\gamma_1+\gamma_0} - \frac{\gamma_2}{\eta(\gamma_1+\gamma_0)}\theta\cdot\mathbf{x}\right)\bar{1} = \mathbf{I}\bar{x}. \tag{39}$$

Using the definition of Bonacich centrality vector, we can now write:

$$\frac{\alpha}{\gamma_1+\gamma_0}\mathbf{b}\left(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)}\right) - \frac{\gamma_2}{\eta(\gamma_1+\gamma_0)}\theta\mathbf{x}\mathbf{b}\left(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)}\right) = \mathbf{I}\bar{x} \tag{40}$$

In order to ease notation, from now on we drop the argument of the centrality vectors. Premultiplying by $\bar{1}'$ we get:

$$\frac{\alpha}{\gamma_1+\gamma_0}b - \frac{\gamma_2}{\eta(\gamma_1+\gamma_0)}\theta b\mathbf{x} = \mathbf{x}$$

and thus

$$\mathbf{x} = \frac{\alpha\eta b}{\eta(\gamma_1+\gamma_0)+\gamma_2\theta b}$$

substituting this into (40) we get the result of the proposition. ■

Proof of Proposition 3. Let us rewrite the FOC (3) as follows:

$$\alpha\mathbf{I}\bar{1} - \gamma_2(\mathbf{G}+\mathbf{G^2})\bar{x} = ((\gamma_1+\gamma_0)\mathbf{I} - (\phi-\gamma_1)\mathbf{G})\bar{x} \tag{41}$$

from which we obtain:

$$\alpha\mathbf{I}\bar{1} - \gamma_2(\mathbf{G}+\mathbf{G^2})\bar{x} = (\gamma_1+\gamma_0)(\mathbf{I} - \frac{(\phi-\gamma_1)}{\gamma_1+\gamma_0}\mathbf{G})\bar{x} \tag{42}$$

Recalling that $(\mathbf{I} - \frac{(\phi-\gamma_1)}{\gamma_1+\gamma_0}\mathbf{G})^{-1}$ is the Bonacich centrality matrix $\mathbf{M}(\mathbf{G}, \frac{(\phi-\gamma_1)}{\gamma_1+\gamma_0})$, we can write (dropping the arguments for simplicity):

$$\frac{\alpha}{\gamma_1+\gamma_0}\mathbf{M}\bar{1} - \mathbf{M}\frac{\gamma_2}{\gamma_1+\gamma_0}(\mathbf{G}+\mathbf{G^2})\bar{x} = \bar{x} \tag{43}$$

Rearranging terms we get:

$$\frac{\alpha}{\gamma_1+\gamma_0}\mathbf{M}\bar{1} = [\mathbf{I} + \frac{\gamma_2}{\gamma_1+\gamma_0}\mathbf{M}\ (\mathbf{G}+\mathbf{G^2})]\bar{x} \tag{44}$$

If the matrix $[\mathbf{I} + \gamma_2\mathbf{M}(\mathbf{G}+\mathbf{G^2})]$ is invertible, we write:

$$\frac{\alpha}{\gamma_1+\gamma_0}[\mathbf{I} + \frac{\gamma_2}{\gamma_1+\gamma_0}\mathbf{M}(\mathbf{G}+\mathbf{G^2})]^{-1}\mathbf{M}\bar{1} = \bar{x}. \tag{45}$$

Since all matrices are symmetric, reorganizing terms and using the definition of weighted centrality we obtain the desired expression. ■

Proof of Proposition 6. Consider first a node $k \notin Z$ such that $g_{kz} = 0$ for all $z \in Z$. We have $\tilde{g}_{ki} = \tilde{g}'_{ki}$ for all $i \in N$. Consider then a node $k \notin Z$ such that $g_{ki} = 1$ for at least one $i \in Z$. We have that $\tilde{g}'_{ki} <$˙ki*and*˙kz$\leq \tilde{g}_{kz}$ for all $z \in Z$. Consider now any two nodes $i, j \in Z$, for which, by construction, $g'_{ij} - g_{ij} = 1$. We also have $g'^{[2]}_{ij} - g^{[2]}_{ij} = \rho + 2 - 2$, since all nodes in Z are now linked with each other. Thus $\tilde{g}'_{ij} - \tilde{g}_{ij} = 1 - \frac{\rho\gamma_2}{\phi-\gamma_1-\gamma_2} \leq 0$ since we have assumed that $(\phi - \gamma_1 - \gamma_2) \leq \rho\gamma_2$. Thus, $\tilde{g}'_{ij} \leq \tilde{g}_{ij}$ for all $i, j \in Z$ with at least one strict inequality. ■

Proof of Proposition 7. For ease of notation call

$$A = \frac{\gamma_2\theta}{(\gamma_1+\gamma_0)(\eta(\gamma_1+\gamma_0)+\gamma_2\theta b(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)}))}$$

Consider now equation (21) and call $b(\mathbf{C}, \frac{\lambda\gamma_2}{\eta(\gamma_1+\gamma_0)}, \alpha) \equiv b_\alpha$. We have that

$$\frac{\partial x^i}{\partial \alpha^j} = \frac{1}{\gamma_1+\gamma_0}[\eta m^{ij} - Ab^i \frac{\partial b_\alpha}{\partial \alpha^j}],\ \forall i, j$$

note that

$$\frac{\partial b_\alpha}{\partial \alpha^j} = \sum_i m^{ij}$$

Recall that, given symmetry of matrix $\mathbf{M}$, we have that $\sum_i m^{ij} = b^j$, so that

$$\frac{\partial x^i}{\partial \alpha^j} = \frac{1}{\gamma_1+\gamma_0}[\eta m^{ij} - Ab^i b^j]$$

Now

$$\frac{\partial \mathbf{x}}{\partial \alpha^j} = \sum_i \frac{\partial x^i}{\partial \alpha^j}$$

so that

$$\frac{\partial \mathbf{x}}{\partial \alpha^j} = \frac{1}{\gamma_1+\gamma_0}\sum_i[\eta m^{ij} - Ab^i b^j] = \frac{1}{\gamma_1+\gamma_0}[\eta - Ab]b^j$$

so that the key player is the agent j with the highest b^j. ■

Proof of Proposition 4. We start by considering the matrix:

$$\mathbf{M}\left(\mathbf{G} - \frac{\gamma_2}{(\phi-\gamma_1-\gamma_2)}\mathbf{G^2}, \frac{(\phi-\gamma_1-\gamma_2)}{\gamma_1+\gamma_0})\right) \equiv [\mathbf{I} - \frac{(\phi-\gamma_1-\gamma_2)}{\gamma_1+\gamma_0}(\mathbf{G} - \frac{\gamma_2}{(\phi-\gamma_1-\gamma_2)}\mathbf{G^2}]^{-1} \quad (46)$$

which by (4) determines equilibrium behavior up to proportionality factor.
This can be rewritten as:

$$\sum_{k=0}^{\infty}\mathbf{G}^k[\frac{(\phi-\gamma_1-\gamma_2)}{\gamma_1+\gamma_0}\mathbf{I} + \frac{\gamma_2}{\gamma_1+\gamma_0}\mathbf{G}]^k = \sum_{k=0}^{\infty}\frac{1}{(\gamma_1+\gamma_0)^k}\mathbf{G}^k[(\phi-\gamma_1-\gamma_2)\mathbf{I} - \gamma_2\mathbf{G}]^k \quad (47)$$

Applying the binomial expansion to the second term we get:

$$\sum_{k=0}^{\infty}\frac{1}{(\gamma_1+\gamma_0)^k}\mathbf{G}^k\sum_{i=0}^{k}\binom{k}{i}(\phi-\gamma_1-\gamma_2)^i(-\gamma_2^{k-i})\mathbf{G}^{k-i} \quad (48)$$

from which

$$\sum_{k=0}^{\infty}\frac{1}{(\gamma_1+\gamma_0)^k}\sum_{i=0}^{k}\binom{k}{i}(\phi-\gamma_1-\gamma_2)^i(-\gamma_2^{k-i})\mathbf{G}^{2k-i} \tag{49}$$

The derivative of this with respect to γ_2 evaluated at the point $\gamma_2 = 0$ is:

$$\lim_{\gamma_2\to 0}\frac{\sum_{k=0}^{\infty}\frac{1}{(\gamma_1+\gamma_0)^k}\sum_{i=0}^{k}\binom{k}{i}(\phi-\gamma_1-\gamma_2)^i(-\gamma_2)^{k-i}\mathbf{G}^{2k-i}-\sum_{k=0}^{\infty}\frac{1}{(\gamma_1+\gamma_0)^k}(\phi-\gamma_1)^k\mathbf{G}^k}{\gamma_2} \tag{50}$$

$$\lim_{\gamma_2\to 0}\frac{\sum_{k=0}^{\infty}\frac{1}{(\gamma_1+\gamma_0)^k}[\sum_{i=0}^{k}\binom{k}{i}(\phi-\gamma_1-\gamma_2)^i(-\gamma_2)^{k-i}\mathbf{G}^{2k-i}-(\phi-\gamma_1)^k\mathbf{G}^k]}{\gamma_2} \tag{51}$$

Note now that: for $k = i$ we have $\binom{k}{i} = 1$, $\gamma_2^{k-i} = 1$ and $\mathbf{G}^{2k-i} = \mathbf{G}^k$; for $k - i \geq 2$ we have $\frac{\gamma_2^{k-i}}{\gamma_2} = 0$; for $k - i = 1$ we have $\binom{k}{i} = k$, $(-\gamma_2^{k-i}) = -\gamma_2$ and $\mathbf{G}^{2k-i} = \mathbf{G}^{k+1}$. Summing up we obtain:

$$\begin{aligned}
-\sum_{k=0}^{\infty}\frac{1}{(\gamma_1+\gamma_0)^k}k(\phi-\gamma_1)^{k-1}\mathbf{G}^{k+1} &= -\mathbf{G}\sum_{k=0}^{\infty}[\frac{\partial}{\partial(\phi-\gamma_1)}(\phi-\gamma_1)^k]\frac{1}{(\gamma_1+\gamma_0)^k}\mathbf{G}^k \\
&= -\mathbf{G}\frac{\partial}{\partial(\phi-\gamma_1)}\sum_{k=0}^{\infty}\frac{1}{(\gamma_1+\gamma_0)^k}(\phi-\gamma_1)^k\mathbf{G}^k \\
&= -\mathbf{G}\frac{\partial\mathbf{M}(\mathbf{G},\frac{\phi-\gamma_1}{\gamma_1+\gamma_0})}{\partial\phi}.
\end{aligned} \tag{52}$$

Summing up we obtain:

$$\frac{\partial}{\partial\gamma_2}\mathbf{M}\left(\tilde{\mathbf{G}},\frac{(\phi-\gamma_1-\gamma_2)}{\gamma_1+\gamma_0}\right)|_{\gamma_2=0}=-\mathbf{G}\frac{\partial\mathbf{M}(\mathbf{G},\frac{\phi-\gamma_1}{\gamma_1})}{\partial\phi}=-\frac{\partial\mathbf{M}(\mathbf{G},\frac{\phi-\gamma_1}{\gamma_1+\gamma_0})}{\partial\phi}\mathbf{G} \tag{53}$$

where the last equality comes from symmetry of all involved matrices. Post multiplying the first and last term in the above equalities by $\bar{1}$ we finally get:

$$\frac{\partial}{\partial\gamma_2}\mathbf{M}\left(\tilde{\mathbf{G}},\frac{(\phi-\gamma_1-\gamma_2)}{\gamma_1+\gamma_0}\right)|_{\gamma_2=0}\cdot\bar{1}=-\frac{\partial}{\partial\phi}\bar{b}(\mathbf{G},\frac{\phi-\gamma_1}{\gamma_1+\gamma_0},d), \tag{54}$$

which proves the first part of the proposition.

Turning now to the second part of the proposition, and using the following expression of the centrality matrix for $\mathbf{G}$:

$$\mathbf{M}(\mathbf{G},\frac{\phi-\gamma_1}{\gamma_1+\gamma_0})=\sum_{k=0}^{\infty}(\frac{\phi-\gamma_1}{\gamma_1+\gamma_0})^k\mathbf{G}^k, \tag{55}$$

Using (54) and taking derivatives with respect to ϕ:

$$\frac{\partial^2}{\partial\gamma_2\partial\phi}\bar{b}\left(\tilde{\mathbf{G}},\frac{(\phi-\gamma_1-\gamma_2)}{\gamma_1+\gamma_0}\right)|_{\gamma_2=0}=-\frac{\partial^2}{\partial\phi^2}\bar{b}(\mathbf{G},\frac{\phi-\gamma_1}{\gamma_1+\gamma_0},d)=-\frac{1}{(\gamma_1+\gamma_0)^2}\sum_{k=0}^{\infty}[\mathbf{G}^k k(k-1)(\frac{\phi-\gamma_1}{\gamma_1+\gamma_0})^{k-2}]\cdot d<0$$

■

Proof of Proposition 8. Each agent has dq neighbors of own type and $d(1-q)$ neighbors of different type. Moreover, let $t \in L, H$ and consider an agents of type t. $dq(dq-1)$ is the number of agents of type t (other then self) connected with neighbors of type t; $d(1-q)[d(1-q)-1]$ is the number of agents of type t connected with neighbors of type different from t; $dqd(1-q)$ is the number of agents of type different from t connected with neighbors of type t; $d(1-q)dq$ is the number of agents of different from t connected with neighbors of different from t. Consequently, by imposing symmetry on the FOC of each type, we get

$$\begin{cases} \alpha^H - \sigma x^H + dq(\phi-\gamma_1-\gamma_2)x^H + d(1-q)(\phi-\gamma_1-\gamma_2)x^L \\ -\gamma_2\{dq[dq-1]+d(1-q)[d(1-q)-1]\}x^H - \gamma_2\{dqd(1-q)+d(1-q)dq\}x^L = 0 \\ \\ \alpha^L - \sigma x^L + dq(\phi-\gamma_1-\gamma_2)x^L + d(1-q)(\phi-\gamma_1-\gamma_2)x^H \\ -\gamma_2\{dq[dq-1]+d(1-q)[d(1-q)-1]\}x^L - \Gamma_2\{dqd(1-q)+d(1-q)dq\}x^H = 0 \end{cases} \tag{56}$$

and the equilibrium in (22) is derived.

In order to prove that the maximal spread in independent from d, simply note that $\bar{q}$ is independent from the level os segregation. Thus call the spread S and notice that

$$S(q = \frac{\phi-\gamma_1-\gamma_2(1-2d)}{4d\gamma_2}) = -\frac{4(\alpha^H-\alpha^L)\gamma_2}{\gamma_1^2 - 4\gamma_0\gamma_2 + (\gamma_2-\phi)^2 - 2\gamma_1(\gamma_2+\phi)}$$

Proof of Proposition 10.

The reduced form is:

$$x = (\mathbf{I} - \beta_1\mathbf{G} - \beta_2\mathbf{G^2})^{-1}(\rho\mathbf{I} + \zeta\mathbf{G})z + (\mathbf{I} - \beta_1\mathbf{G} - \beta_2\mathbf{G^2})^{-1}\epsilon$$

Consider two sets of structural parameters $(\alpha, \beta_1, \rho, \zeta, \beta_2)$ and $(\alpha', \beta_1', \rho', \zeta', \beta_2')$. If they lead to the same reduced form, it means that $(\mathbf{I}-\beta_1\mathbf{G}-\beta_2\mathbf{G^2})^{-1}(\rho\mathbf{I}+\zeta\mathbf{G}) = (\mathbf{I}-\beta_1'\mathbf{G}-\beta_2'\mathbf{G^2})^{-1}(\rho'\mathbf{I}+\zeta'\mathbf{G})$. Premultiply the second equality by $(\mathbf{I}-\beta_1\mathbf{G}-\beta_2\mathbf{G^2})(\mathbf{I}-\beta_1'\mathbf{G}-\beta_2'\mathbf{G^2})$ we get

$$(\mathbf{I} - \beta_1'\mathbf{G} - \beta_2'\mathbf{G^2})(\rho\mathbf{I} + \zeta\mathbf{G}) = (\mathbf{I} - \beta_1\mathbf{G} - \beta_2\mathbf{G^2})(\rho'\mathbf{I} + \zeta'\mathbf{G})$$

which can be rewritten as

$$(\rho-\rho')\mathbf{I} + (\zeta - \zeta' + \beta_1\rho' - \beta_1'\rho)\mathbf{G} + (\beta_1\zeta' - \beta_1'\zeta + \beta_2\rho' - \beta_2'\rho)\mathbf{G^2} + (\beta_2\zeta' - \beta_2'\zeta)\mathbf{G^3} = 0$$

if $\mathbf{I}, \mathbf{G}, \mathbf{G^2}, \mathbf{G^3}$ are linearly independent, then

$$\rho = \rho' \tag{57}$$

$$\zeta' - \beta_1\rho' = \zeta - \beta_1'\rho \tag{58}$$

$$\beta_1\zeta' - \beta_2\rho' = \beta_1'\zeta - \beta_2'\rho \tag{59}$$

$$\beta_2\zeta' = \beta_2'\zeta \tag{60}$$

Consider now $\beta_2'\zeta \neq 0$, thus $\zeta \neq 0$. From (60) define $\zeta' = \lambda\zeta$ and $\beta_2' = \lambda\beta_2$ so (59) becomes

$$\beta_1\lambda\zeta - \beta_2\rho = \beta_1'\zeta - \lambda\beta_2\rho \tag{61}$$

and (58) becomes

$$\lambda\zeta - \beta_1\rho = \zeta - \beta_1'\rho \tag{62}$$

from (62)

$$\beta_1' = \beta_1 - \frac{\lambda\zeta}{\rho} + \frac{\zeta}{\rho} \tag{63}$$

sustituting in (61) we get

$$\beta_1\lambda\zeta - \beta_2\rho = (\beta_1 - \frac{\lambda\zeta}{\rho} + \frac{\zeta}{\rho})\zeta - \lambda\beta_2\rho \tag{64}$$

$$\beta_1\lambda\zeta - \beta_2\rho = \beta_1\zeta - \frac{\lambda\zeta^2}{\rho} + \frac{\zeta^2}{\rho} - \lambda\beta_2\rho \tag{65}$$

$$\beta_1\lambda\zeta + \frac{\lambda\zeta^2}{\rho} + \lambda\beta_2\rho = \beta_1\zeta + \frac{\zeta^2}{\rho} + \beta_2\rho \tag{66}$$

If $\beta_1\zeta + \frac{\zeta^2}{\rho} + \beta_2\rho \neq 0$ and $\rho \neq 0$,

$$\lambda(\beta_1\zeta + \frac{\zeta^2}{\rho} + \beta_2\rho) = \beta_1\zeta + \frac{\zeta^2}{\rho} + \beta_2\rho \tag{67}$$

i.e. λ=1, so the two sets of parameters are the same.
Consider now $\beta_2\zeta = 0$. This can be due to either $\zeta = 0$, or $\beta_2 = 0$ (or both).
Consider first the case of $\zeta = 0$, then the coefficients associated to $\mathbf{G}$ and $\mathbf{G^2}$ become

$$\beta_1\rho' = \beta_1'\rho \tag{68}$$

$$\beta_2\rho' = \beta_2'\rho \tag{69}$$

So β_1 and β_2 are identified if $\rho \neq 0$, and thus identified from (57).
Consider now the case of $\beta_2 = 0$ so that the problem collapses to the case of Bramoullé et al. (2009) so that the coefficients are identified if $\zeta + \beta_1\rho \neq 0$. ■

Proof of Proposition 11. We can write (27) as:

$$x = (\mathbf{I} - \beta_1\mathbf{G})^{-1}(\rho\mathbf{I} + \zeta\mathbf{G}^*)z + (\mathbf{I} - \beta_1\mathbf{G})^{-1}\epsilon$$

Consider two sets of parameters (β_1, ρ, ζ) and $(\beta_1', \rho', \zeta')$ that provide the same estimates. Then

$$(\mathbf{I} - \beta_1\mathbf{G})^{-1}(\rho\mathbf{I} + \zeta\mathbf{G}^*) = (\mathbf{I} - \beta_1'\mathbf{G})^{-1}(\rho'\mathbf{I} + \zeta'\mathbf{G}^*)$$

Multiplying both sides by $(\mathbf{I} - \beta_1\mathbf{G})(\mathbf{I} - \beta_1'\mathbf{G})$ we obtain

$$(\mathbf{I} - \beta_1'\mathbf{G})(\rho\mathbf{I} + \zeta\mathbf{G}^*) = (\mathbf{I} - \beta_1\mathbf{G})(\rho'\mathbf{I} + \zeta'\mathbf{G}^*)$$

This can be rewritten as

$$(\rho - \rho')\mathbf{I} + (\zeta - \zeta')\mathbf{G}^* - (\rho\beta_1' - \rho'\beta_1)\mathbf{G} - (\zeta\beta_1' - \zeta'\beta_1)\mathbf{G}\mathbf{G}^* = 0$$

Suppose $\mathbf{I}, \mathbf{G}, \mathbf{G}^*, \mathbf{G}\mathbf{G}^*$ to be linearly independent. Then $\rho = \rho'$ and $\zeta = \zeta'$. If $\rho \neq 0$ or $\zeta \neq 0$ then it immediately follows that $\beta_1 = \beta_1'$. ■

Proof of Proposition 12. We can write (26) as

$$x = (\mathbf{I} - \beta_1\mathbf{G} - \beta_2\mathbf{G^2})^{-1}(\rho\mathbf{I} + \zeta\mathbf{G}^*)z + (\mathbf{I} - \beta_1\mathbf{G} - \beta_2\mathbf{G^2})^{-1}\epsilon$$

Consider two sets of parameters $(\beta_1, \beta_2, \rho, \zeta)$ and $(\beta_1', \beta_2', \rho', \zeta')$ that provide the same estimates. Then

$$(\mathbf{I} - \beta_1\mathbf{G} - \beta_2\mathbf{G^2})^{-1}(\rho\mathbf{I} + \zeta\mathbf{G}^*) = (\mathbf{I} - \beta_1'\mathbf{G} - \beta_2'\mathbf{G^2})^{-1}(\rho'\mathbf{I} + \zeta'\mathbf{G}^*)$$

Premultiplying both sides by $(\mathbf{I} - \beta_1\mathbf{G} - \beta_2\mathbf{G^2})(\mathbf{I} - \beta_1'\mathbf{G} - \beta_2'\mathbf{G^2})$ we obtain

$$(\mathbf{I} - \beta_1'\mathbf{G} - \beta_2'\mathbf{G^2})(\rho\mathbf{I} + \zeta\mathbf{G}^*) = (\mathbf{I} - \beta_1\mathbf{G} - \beta_2\mathbf{G^2})(\rho'\mathbf{I} + \zeta'\mathbf{G}^*)$$

This can be rewritten as

$$(\rho-\rho')\mathbf{I}+(\zeta-\zeta')\mathbf{G}^*-(\rho\beta_1'-\rho'\beta_1)\mathbf{G}-(\zeta\beta_1'-\zeta'\beta_1)\mathbf{G}\mathbf{G}^*-(\rho\beta_2'-\rho'\beta_2)\mathbf{G^2}-(\zeta\beta_2'-\zeta'\beta_2)\mathbf{G^2}\mathbf{G}^* = 0$$

Suppose $\mathbf{I}, \mathbf{G}, \mathbf{G^2}, \mathbf{G}^*, \mathbf{G}\mathbf{G}^*, \mathbf{G^2}\mathbf{G}^*$ to be linearly independent. Then $\rho = \rho'$ and $\zeta = \zeta'$. If $\rho \neq 0$ or $\zeta \neq 0$ then it immediately follows that $\beta_1 = \beta_1'$ and $\beta_2 = \beta_2'$.