2014-02

Conformism and Self-Selection in Social Networks

Vincent Boucher

novembre / november 2014

Centre de recherche sur les risques les enjeux économiques et les politiques publiques

www.crrep.ca

Abstract

I present a model of conformism in social networks that incorporates both peer effects and self-selection. I find that equilibrium behaviors are linked through the Laplacian matrix of the equilibrium network. I show that conformism has positive social value and that social welfare can be bounded by network centrality and connectivity measures. I apply the model using empirical data on high school student participation in extracurricular activities. I find that the local effects of conformism (i.e. endogenous peer effect for a fixed network structure) range from 7.5% to 45%, depending on the number of peers that an individual has. Simulations show that the optimal policies of an inequality-averse policy-maker change in relation to the size of a school. Small schools should encourage shy students to integrate more with other students, while large schools should focus on promoting role models within the school.

Mots-clés : Conformism, peer effects, network formation

Classification JEL : D85, C31, C57.

Boucher : Department of Economics, Université Laval; and CIRPÉE vincent.boucher@ecn.ulaval.ca

Part of this work was done while I was visiting the Faculty of Economics at the University of Cambridge (UK); Iwould like to thank Sanjeev Goyal and the Faculty, as well as INET-Cambridge, for their time and hospitality. I would like to thank Luc Bissonnette for his precious help and suggestions. I also would like to thank Yann Bramoullé, Charles Bellemare, Bernard Fortin, Edoardo Gallo, David Karp, Guy Lacroix, Steeve Marchand, Angelo Mele, Youcef Msaid and Onur Özgür for their comments and suggestions. I would like to thank the participants at many conferences, including Les Journées du CIRPÉE (2013, QC), Coalition Theory Workshop (2014, Brussel), the Coalitions and Network workshop (Concordia U., Montreal, 2014) and the CEA annual conference (BC, 2014). I also gratefully acknowledge financial support from FRQSC.

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

1. Introduction

Do teenagers smoke because their friends smoke, or do they smoke in the hope of making new friends? Does peer pressure come from *influence* or *self-selection*?

The literature on peer effects in social networks has mainly focused on influence, while the literature on network formation has focused on self-selection. To date, these two sources of social interactions have mainly been studied separately.

In this paper, I present a model of conformism in social networks, where both peer effects and self-selection affect behavior. The magnitude of each effect can be clearly identified because changes to an individual's peer group induce *discontinuous* changes in the individual's behavior, while changes in how an individual's peers behave (holding constant who an individual chooses as his peers) induces *continuous* changes in the individual's behavior. This is important as I show that peer effects and self-selection have different policy implications. I characterize the set of all (Nash) equilibria and present an equilibrium refinement (perfect and robust) based on the potential function of the game. I estimate the model using student-level data on participation in high school activities.

I characterize the relationship between individual behavior and the network structure for all equilibria of the game. Behavior can be expressed as a function of the Laplacian matrix of the network. This specific mathematical structure implies that individuals are affected by the *entire* distribution of their peers' behavior, and not just the average behavior of their peers. As a result, the overall impact of conformism on the outcome variable increases with the number of peers an individual has, with each marginal peer having less impact than the previous one.

I show that, for a social planer with quadratic preferences, equilibria can be ranked according to the variance of the equilibrium behaviors and that conformism has positive social value. I present bounds for the equilibrium variance based on the structure of the equilibrium network. The variance is bounded below by a measure of network centrality and above by a function of the edge-connectivity of the network.

This has powerful policy implications. A social planer who wants to prevent the emergence of bad equilibria should promote integration and ensure that no group of individuals is isolated from the rest of the network. A social planer who wants to support the emergence of good equilibria should focus on promoting role models. A comprehensive public policy should thus consider both centrality and connectivity.

I apply the model empirically using data on the choice of extracurricular activities by high school students.[1] Although it is not feasible to estimate the true model, I provide bounds for the density of the equilibrium network; each bound can be interpreted as a latent space model (see, for example, Goldsmith-Pinkham & Imbens (2013)). In practice, the bounds lead to roughly the same estimated parameters. Depending on the number of peers that a student has, the local impact of conformism (that is, the endogenous peer effect for a given network structure) ranges from 7.5% to 45%. Using simulations, I show that the cost of increasing connectivity in small schools is relatively low and may lead to large welfare gains. The opposite is true for large schools where the cost is high and the benefit to welfare is low. This suggests that small schools should focus on connectivity and large schools should focus on centrality.

This paper contributes to the literature on conformism in social networks. I focus on quadratic preferences, as in Bisin et al. (2006), Bisin & Özgür (2012) and Patacchini & Zenou (2012). Bisin et al. (2006) and Bisin & Özgür (2012) present dynamic theoretical models for fixed network structures and provide identification conditions. Patacchini & Zenou (2012) also focus on quadratic preferences and present an empirical application, which assumes that the network is exogenous. I present a static model of conformism with quadratic preferences and self-selection.

[1] Such activities include chess clubs and sport teams, for example.

This paper also contributes to the theoretical literature featuring games in endogenous networks. Hojman & Szeidl (2006) present a model where individuals simultaneously choose their behavior and the agents that they link with within a network. They find that, with mostly homogeneous agents, the equilibrium network is minimally connected. Herman (2013) and Kinateder & Merlino (2014)) focus on the provision of local public goods. This paper complements the literature by focusing on conformism among heterogeneous agents and by showing that equilibrium behaviors are linked through the Laplacian matrix of the equilibrium network.

This paper also contributes to the recent empirical literature on peer effects in endogenous networks. Goldsmith-Pinkham & Imbens (2013) and Hsieh & Lee (2011) present models where there is endogeneity, which is due to the presence of an unobserved variable. They find that the estimated endogenous effects, controlling for the endogeneity of the network, are similar to those estimated when the network is exogenous. Using a similar approach, Patacchini & Rainone (2014) also find that the bias due to the potential endogeneity of a network's structure is small. However, as noted by Badev (2013) and by Boucher & Fortin (2014), this does not imply that the network formation process can be ignored, since it may still affect the efficiency of public policies. For instance, Badev (2013) finds that ignoring the network formation process leads to biases (from 10% to 15%) on the predicted impact of public policies.

In this paper, as in Badev (2013), the endogeneity is created by the fact that the endogenous variable (behavior) directly affects the value of links, assuming that all the relevant variables are observed. I also find a small bias on the estimated parameters, although the endogenous structure of the network may still strongly impact the effectiveness of public policies. I provide additional insights as of why the network formation process does not substantially bias the estimated impact of peer effect for the AddHealth database.

The remainder of the paper is as follows. In section 2, I present a microeconomic model where individuals simultaneously choose their behavior and their peer groups. In section 3, I present an empirical application using data on student participation in extracurricular activities. I conclude in section 4.

2. The Microeconomic Model

The economy is composed of n individuals. Individuals simultaneously make two decisions: their behavior ($y_i \in \mathbb{R}$), and their peers ($g_i \in \{0,1\}^{n-1}$). I assume that individuals have preferences for conformism:[2]

$$u_i(y,g) = \sum_{j\neq i}\left[\mathbf{z}_{ij}\delta - \frac{\lambda}{2}(y_i - y_j)^2 + \eta_{ij}\right] g_{ij}g_{ji} - \frac{1}{2}(y_i - \mathbf{x}_i\boldsymbol{\beta} - \varepsilon_i)^2$$

where $\lambda \geq 0$, $\mathbf{z}_{ij} = \mathbf{z}_{ji}$ is a vector of pair-specific characteristics,[3] and $\mathbf{x}_i$ is a vector of individual characteristics. The distributions of the unobserved shocks ε_i and $\eta_{ij} = \eta_{ji}$ are left free for the moment, but distributional assumptions will be made in section 3. Individuals incur a cost if they choose a behavior y_i that is different than their type (i.e. $\mathbf{x}_i\boldsymbol{\beta}+\varepsilon_i$), and different than their peers' behavior, y_j. However, individuals gain from forming connections with peers that provide a net positive value. Since the value of a link is multiplied by $g_{ij}g_{ji}$, a link is only created under mutual consent, i.e. $g_{ij} = g_{ji} = 1$.

Note that if there are no social interactions, i.e. $\lambda = 0$, individuals' utilities reduce to $u_i(y,g) = \sum_{j\neq i}\left[\mathbf{z}_{ij}\boldsymbol{\delta} + \eta_{ij}\right] g_{ij}g_{ji} - \frac{1}{2}(y_i - \mathbf{x}_i\boldsymbol{\beta} - \varepsilon_i)^2$ and the optimal behavior for each individual is given by $y_i = \mathbf{x}_i\boldsymbol{\beta} - \varepsilon_i$. Similarly, a link is created only if $\mathbf{z}_{ij}\boldsymbol{\delta} + \eta_{ij} \geq 0$. However, if $\lambda > 0$, the utility function is not separable in (y_i, g_i), so the optimal decision for y_i is a function of g_i, and vice-versa.

The model induces a strategic form game $\Gamma = \langle N, \{S_i\}_{i\in N}, \{u_i\}_{i\in N}\rangle$, where $S_i = \mathbb{R}\times\{0,1\}^{n-1}$. Let $S = S_1 \times ... \times S_n$. A *Nash equilibrium* (NE) of Γ is an allocation $(y^*, g^*) \in S$ such that for all $i \in N$,

$$(y_i^*, g_i^*) \in \arg\max_{(y_i,g_i)\in S_i} u_i(y_i, y_{-i}^*, g_i, g_{-i}^*)$$

[2]The preferences are similar to the instantaneous utility of Bisin & Özgür (2012).
[3]The model can be extended to non-symmetric $\mathbf{z}_{ij}$.

Let us first consider the optimal choice of g, given y. Intuitively, individuals should choose g_i such that they keep links with a positive value and discard links with a negative value. If this holds true, the aggregate utility function reduces to:

$$w_i(y) = \sum_{j \neq i} \max\{\mathbf{z}_{ij}\boldsymbol{\delta} - \frac{\lambda}{2}(y_i - y_j)^2 + \eta_{ij}, 0\} - \frac{1}{2}(y_i - \mathbf{x}_i\boldsymbol{\beta} - \varepsilon_i)^2 \qquad (1)$$

This assumes that individuals play *as if* every link with positive value is created. Define $\tilde{\Gamma} = \langle N, \{\tilde{S}_i\}_{i \in N}, \{w_i\}_{i \in N} \rangle$, where $\tilde{S}_i = \mathbb{R}$. The following proposition shows that the set of equilibria of $\tilde{\Gamma}$ is a subset of the equilibria of Γ.

Proposition 1. *If y^* is a NE of $\tilde{\Gamma}$, then there exists g^* such that (y^*, g^*) is a NE of Γ.*

Proposition 1 is convenient as it reduces the dimensionality of the problem. Note, however, that proposition 1 does *not* imply that any equilibrium of Γ can be found from the resolution of $\tilde{\Gamma}$. This is because the specification in (1) implicitly solves for the bilateral coordination problem: if i does not invest in a link with j (i.e. $g_{ij} = 0$), then j has no incentive to invest in the link, even if the value of the link is positive. This observation has motivated the introduction of equilibrium refinements allowing for bilateral deviations such as pairwise stability (Jackson & Wolinsky, 1996) and bilateral equilibria (Goyal & Vega-Redondo, 2007).

In what follows, I concentrate on the analysis of $\tilde{\Gamma}$, and hence implicitly assume that every link with positive value is created.[4] Note that from proposition 1, any equilibrium of the modified game $\tilde{\Gamma}$ is an equilibrium of the original game Γ.

A small technical issue with the function in (1) is that it is not differentiable everywhere. However, the next lemma shows that $w_i(y)$ is locally differentiable

[4]This is a standard assumption in the empirical literature on network formation. See, for example, Christakis et al. (2010) and Goldsmith-Pinkham & Imbens (2013).

around its maxima, and that there exists a generically unique maximum.[5]

Lemma 2. *Let* $y_i^* \in \arg\max_{y_i} w_i(y_i, y_{-i})$*. Then, for all* $j \neq i$*, we generically have that* $\mathbf{z}_{ij}\boldsymbol{\delta} - \frac{\lambda}{2}(y_i^* - y_j)^2 + \eta_{ij} \neq 0$*. Moreover,* y_i^* *exists and is generically unique.*

Lemma 2 says that individuals never maximize at a "kink" in the utility function. The intuition is as follows: if i and j are linked, the choice of y_i is affected by y_j. Hence, i chooses a different behavior when he is linked with j than when he is not. The value of their link *must* outweigh the cost of conforming with j. If the link has no value, i would be better off removing the link with j and re-optimizing without j as a peer.

Lemma 2 implies that the utility functions are differentiable around their maximum, so that it is possible to write the first-order conditions of the optimization problem. This property is at the source of the identification strategy, as it allows for the capture of the effect of marginally changing the value of y, keeping g constant. This allows for a local analysis of the model that looks at peer effects within a fixed network structure.

2.1. Local Analysis

From lemma 2, we can write the first-order conditions for the maximum of $w_i(y)$, evaluated at the equilibrium (y^*, g^*). Doing so reveals a closed-form relation between y^* and the Laplacian matrix of the equilibrium graph. The Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{G}$ is obtained from the adjacency matrix $\mathbf{G}$ and the diagonal degree matrix $\mathbf{D}$, where $\mathbf{D}_{ii} = \sum_j \mathbf{G}_{ij}$. Let $E(a_i)$ and $Var(a_i)$ be the mean and variance across individuals for any variable $\mathbf{a}$. This gives the following:

Proposition 3 (Structure). *Let* $(y^*, g^*) \in S$ *be a NE of the original game* Γ*. Then* $\mathbf{y}^* = (\mathbf{I} + \lambda \mathbf{L}^*)^{-1} [\mathbf{X}\boldsymbol{\beta} + \varepsilon]$*. Moreover,* y^* *is such that:*

[5] In this paper, "generically" refers to results that are robust to small changes in η, or equivalently to the fact that the set of values of η for which the proposition does not hold has a null Lebesgue measure.

1. $E(y_i^*) = E(\mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i)$
2. $Var(y_i^*) \leq Var(\mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i)$.

Note that proposition 3 holds for *any equilibrium* of the original game Γ. Note also that the closed-form expression is always well defined, since the Laplacian matrix is positive semi-definite, so $(\mathbf{I} + \lambda\mathbf{L}^*)$ is positive definite for any positive λ and hence invertible. Since $\mathbf{L}^*$ represents the choice of g^*, this expression gives an equilibrium relationship between the optimal choice of y and g. Points 1 and 2 of proposition 3 are direct implications of the closed-form expression. They characterize the effect of conformism on the distribution of behavior. The average behavior is equal to the average individual's type in the population, while the variance of behaviors is always smaller. Note that those features are implied by the conformism game for a given network, and are *not* a result of the network formation process. Also note that the first point of proposition 3 implies the absence of a network multiplier for uniform policy shocks. This feature of the conformism game is also found in Patacchini & Zenou (2012) and Liu et al. (2011), and is discussed in Boucher & Fortin (2014). Finally, note that it implies that the model can be estimated using within-group deviations (see section 3).

Proposition 3 has important implications. The fact that y^* depends on the Laplacian matrix of the network implies that individuals are affected by the full distribution of their peers' behavior, as opposed to their peers' average behaviour, which is the approach taken by Patacchini & Zenou (2012) and most of the empirical literature on peer effects. In order to discuss the additional information provided by the Laplacian matrix, I express the optimal choice for y_i as a function of y_{-i}:

$$y_i = \frac{\lambda}{1+\lambda n_i}\sum_{j\in N_i} y_j + \frac{1}{1+\lambda n_i}[x_i\boldsymbol{\beta} + \varepsilon_i] \tag{2}$$

$$\hat{y}_i = \frac{\lambda}{1+\lambda}\frac{1}{n_i}\sum_{j\in N_i} \hat{y}_j + \frac{1}{1+\lambda}[x_i\boldsymbol{\beta} + \varepsilon_i] \tag{3}$$

where (2) follows from proposition 3 and (3) is the model of Patacchini & Zenou (2012), rewriten in tems of this paper's notation.[6]

For both models, the optimal choice represents a convex combination of the individual's type and the behavior of his peers. Model (2) differs from model (3) mainly by the fact that the number of peers has an impact on the tradeoff between the individual's type and the behavior of his peers. As his number of peers grows, an individual puts less weight on his type and more weight on the behavior of his peers. Also note that the marginal impact of an additional peer decreases in the number of peers that an individual has.

Returning to points 1 and 2 of proposition 3, consider the following social welfare function:

$$W(y) = \sum_i ay_i + \frac{b}{2}y_i^2 \tag{4}$$

where $b < 0$, so the social planer is inequality averse with respect to behavior.[7] Since preferences are quadratic and all the equilibria have the same mean (see proposition 3), they can be ranked according to their variance. Point 2 of proposition 3 thus implies that any equilibrium of the game is preferred to a case where there would be no social interactions (i.e. when $\lambda = 0$). Put differently, conformism has social value.

Unfortunately, there is no general comparative static result between the equilibria, as the equilibrium distribution of y strongly depends on the distribution of individuals' types. For example, increasing the number of links in a network does not necessarily decrease the variance of equilibrium behaviors.[8] However, the variance can be bounded by functions of the network structure.

Let $e(g)$ be the edge connectivity of g, i.e. the minimal number of links in g that need to be removed in order to disconnect the network. Additionally,

[6]This is in fact a special case of their model, see equation (6) on p. 10 of Patacchini & Zenou (2012).

[7]For an inequality averse social planner, social welfare is increasing with Pigou-Dalton transfers.

[8]This follows from the Grone-Merris conjecture (see Bai (2011)) and the fact that $Var(y_i)$ is not Schur-convex in the eigenvalues of the Laplacian matrix.

let $C_{ij}(g) = 1 + \frac{\sum_{k\in N_i} n_k + \sum_{k\in N_j} n_k}{n_i+n_j}$. The quantity $C^* = \max_{j\neq i} C_{ij}(g^*)$ can be interpreted as a measure of network centrality (or polarisation) for a given network density. I provide an illustration in the Appendix (see Figure 6). We have the following:

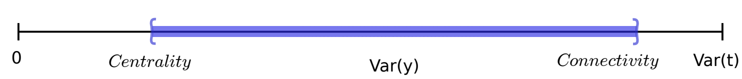
Proposition 4 (Dispersion). *Let $(y^*, g^*) \in S$ be a NE of Γ, and define $e^* = e(g^*)$, $C^* = \max_{j\neq i} C_{ij}(g^*)$ and $t_i = \mathbf{x_i}\boldsymbol{\beta} + \varepsilon_i$, we have:*

$$Var(y_i^*) \geq Var(t_i) - \frac{\mathbf{t}'\mathbf{t}}{n}\frac{\lambda C^*[2+\lambda C^*]}{[1+\lambda C^*]^2} \tag{5}$$

$$Var(y_i^*) \leq Var(t_i) - \frac{\mathbf{t}'\mathbf{t}}{n} + \frac{\mathbf{t}'\mathbf{t}}{n[1+2\lambda e^*(1-cos(\pi/n)]^2} \tag{6}$$

Figure 1 provides a visual representation of proposition 4. The lower bound (5) is strictly decreasing in the network centrality, and the upper bound (6) is strictly decreasing in the edge connectivity. Those bounds have important implications for policy making.

Figure 1: Equilibrium Variance of Behavior ($t_i = \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i$)



Consider a social planner (for example, a high school principal) who wants to reduce the variance of the equilibrium behavior. In order to decrease the lower bound, the school's principal should promote the centrality of the network. This is a generalization of the *key player* argument (see Ballester et al. (2006)) when the network is endogenous. Promoting strong role models improves access to equilibrium outcomes with low variance. In practice, policies such as organizing sports teams or tournaments is likely to generate such role models. However, this does not guarantee that the equilibrium variance will be low, as the upper bound can be quite large.

In order to prevent bad equilibria, a school principal should increase the edge connectivity of the graph. Notice that if the network is disconnected, the edge connectivity is equal to zero and the upper-bound is not binding. In order to decrease this upper bound, a principal should proceed by linking network components (e.g. the principal could encourage shy students to participate in a team activity where they can make new friends and become better integrated with the rest of the student body). Put differently, a society composed of many segregated groups may result in equilibria with high variances.

Proposition 4 clearly exposes the tradeoff between promoting centrality and connectivity, as well as the implications that they have for the variance of equilibrium behavior. Note also that, even if C^* is not necessarily increasing in the number of links (see Figure 6) in Appendix for an example), the expected effect of adding links has a positive impact (i.e. decreases the expected equilibrium variance).[9]

Recall that this section's analysis is conditional on the equilibrium network structure. I now study the full model.

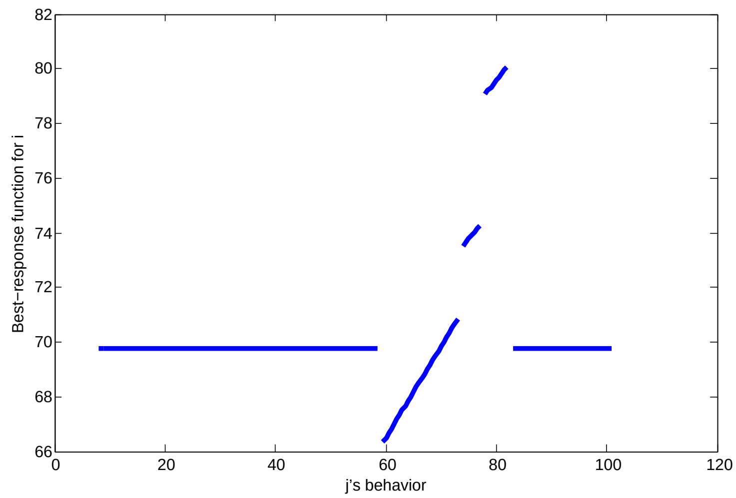
2.2. Global Analysis

In order to develop an intuition of the shape of the best-response functions in $\tilde{\Gamma}$, I start by describing an example of a best-response function in a two-dimensional space.

Figure 2 displays an example of a best-response function for $i \in N$ in the space (y_i, y_j) (that is, keeping y_k constant for all $k \in N \setminus \{i, j\}$). When y_j is too small, i and j are not linked and small changes in y_j do not affect the value of y_i. However, as y_j increases, there is a point where it becomes profitable for i and j to form a link. The link creation has a downward discontinuous effect on y_i (since $y_j < y_0$). When y_i and y_j are such that i and j are linked, y_i reacts linearly to a change in y_j (with a slope of $\frac{\lambda}{\lambda n_i + 1} < 1$, where n_i is the number of peers that i has; see proposition 3). Note that the changes in the slope and the jumps are due to the creation or removal of links with the $n-2$ other individuals. As y_j increases further, there is a point where a link between

[9] This follows from the Grone-Merris conjecture (see Bai (2011)) and the exptression for $Var(y_i)$ in the proof of proposition 4.

Figure 2: Best-Response Functions



i and j is no longer profitable, such that y_i goes back to y_0 and is no longer affected by small changes in y_j.

The n-dimensional best-response function follows the same intuition. This is problematic for existence (of a Nash equilibrium), as none of the standard fixed-point theorems apply. However, $\tilde{\Gamma}$ is a "potential game" (Monderer & Shapley, 1996).[10] A potential game is a game that admits a function $\psi(y)$ such that for all $i \in N$, and any y_i, y_i', we have $\psi(y_i, y_{-i}) - \psi(y_i',, y_{-i}) = w_i(y_i, y_{-i}) - w_i(y_i', y_{-i})$. In the case of $\tilde{\Gamma}$, a valid *potential function* is the following:

$$\psi(y) = \sum_{i \in N} \sum_{j:j<i} \max\left\{ \mathbf{z}_{ij}\boldsymbol{\delta} - \frac{\lambda}{2}(y_i - y_j)^2 + \eta_{ij}, 0 \right\} - \frac{1}{2}\sum_{i \in N}(y_i - \mathbf{x}_i\boldsymbol{\beta} - \varepsilon_i)^2 \quad (7)$$

The maximum of a potential function is a NE (Monderer & Shapley, 1996). Hence, the existence of a NE for $\tilde{\Gamma}$ (and for Γ) follows directly from the following proposition:

[10] The crucial assumption on the initial game Γ that allows for the existence of a potential function is the mutual consent assumption, i.e. the fact that the value of the link is multiplied by $g_{ij}g_{ji}$.

Proposition 5 (Potential's maximum). *The potential function $\psi(y)$ admits a generically unique maximum y^*.*

Since the maximum of the potential is (generically) unique, it also exhibits robustness properties. I will discuss two of those properties: trembling-hand perfection (Carbonell-Nicolau & McLean (2011): Theorem 1), and robustness to canonical elaborations (Ui (2001): Theorem 3). Precise definitions of these concepts can be found in the referenced papers. I focus on the implications of these robustness results, namely the robustness to the presence of small, uncorrelated mistakes on the part of the individuals (trembling-hand perfection), and robustness to the presence of a small amount of imperfect information (canonical elaborations).

Network games are complex in terms of the requirements for individual rationality. Here, individuals have to simultaneously choose the network structure *and* their action. Due to the high dimensionality of the strategy space and the potentially large number of players, the likelihood that individuals make small, uncorrelated, mistakes while playing the game is high. Since the maximum of the potential function is trembling-hand perfect, it is robust to such small mistakes.

Also note that Γ is a game with complete information. This contrasts with part of the literature on social interactions (e.g. Blume et al. (2011)). In an economy where individuals are not ex-ante connected in a network, it seems unlikely that all of the individuals' characteristics are common knowledge. However, since the maximum of the potential is robust to canonical elaborations, it is robust to the incorporation of a small amount of imperfect information, i.e. the maximum of the potential function would still be a (Bayesian-Nash) equilibrium if, with high probability, the individuals' (private) types are such that their preferences are given by (1).

These robustness properties will justify, in the next section, the assumption that the data is generated by the potential's maximum.

Table 1: Summary Statistics

Variable	**Mean**	**Std**	**Min**	**Max**
Number of activities	2.456	2.656	0	33
Age	14.255	1.748	10	19
Gender	0.522	0.500	0	1
White	0.625	0.484	0	1
Hispanic	0.127	0.333	0	1
Black	0.213	0.410	0	1
Asian	0.057	0.231	0	1
Mother works	0.691	0.462	0	1
Mother completed HS	0.771	0.420	0	1
Mother completed college	0.337	0.473	0	1

Number of schools: 100
Number of individuals: 30,241
Number of pairs of individuals: 9,962,501

3. Peer Effects and Extracurricular Activities

In this section, I apply the model to conformism in the choice of extracurricular activities among high school students. This specific context is well adapted to the model as teenagers are (1) more likely to participate an activity if their friends do too, and (2) create new friendships while participating in these activities.

I use the AddHealth database, which is widely used in empirical models of peer effects (e.g. Badev (2013), Christakis et al. (2010), Goldsmith-Pinkham & Imbens (2013) and Hsieh & Lee (2011)). The dependent variable represents the number of extracurricular activities in which a teenager participates. Explanatory variables include age, gender, race, and the mother's education and employment status. In order to reduce the computing time, I focus on the 100 smallest schools in the database. This sample comprises 30,241 individuals for a total of 9,962,501 pairs of individuals. Table 1 presents the summary statistics.

Following proposition 3, and since $\mathbf{1}$ is an eigenvector of $\mathbf{L}$, the model can be estimated using within-group deviations. Let $\tilde{\mathbf{y}}_r = (\mathbf{I} - \frac{1}{n_r}\mathbf{1}\mathbf{1}')\mathbf{y}_r$ and $\tilde{\mathbf{X}}_r = (\mathbf{I} - \frac{1}{n_r}\mathbf{1}\mathbf{1}')\mathbf{X}_r$ for each school $r = 1, ..., R$ of size n_r, and define $\mathbf{M}_r(\lambda) = (\mathbf{I} + \lambda\mathbf{L}_r)$.

I start by describing the choice of $\tilde{\mathbf{y}}$ for a *fixed* network structure. Following proposition 3 and in the spirit of Lee et al. (2010), I define the following quasi-maximum likelihood estimator (QMLE) for $\tilde{\mathbf{y}}$, given a fixed network structure and normally distributed errors:

$$\begin{aligned}\ln P(\tilde{\mathbf{y}}|\mathbf{G},\tilde{\mathbf{X}};\boldsymbol{\beta},\lambda,\sigma_\varepsilon,\sigma_\eta) &= \frac{1}{R}\sum_{r=1}^{R} c - \frac{1}{2}\ln(|\sigma_\varepsilon^2\mathbf{M}_r^{-2}(\lambda)|) \\ &- \frac{1}{2\sigma_\varepsilon^2}\left[\mathbf{M}_r(\lambda)\tilde{\mathbf{y}}_r - \tilde{\mathbf{X}}_r\boldsymbol{\beta}\right]'\left[\mathbf{M}_r(\lambda)\tilde{\mathbf{y}}_r - \tilde{\mathbf{X}}_r\boldsymbol{\beta}\right] \quad (8)\end{aligned}$$

The identification of $\boldsymbol{\beta}$, σ_ε and λ, given $\mathbf{G}$, follows from standard results for linear models (see Blume et al. (2011) or Lee et al. (2010)).

In order to estimate the full model, i.e. $P(\tilde{\mathbf{y}},\mathbf{G}|\tilde{\mathbf{X}},\mathbf{Z})$, one needs to use the density of the network, i.e. $P(\mathbf{G}|\tilde{\mathbf{X}},\mathbf{Z})$. However, as discussed in section 2, the full model typically features many equilibria. As noted by many authors (e.g. Tamer (2003)), games with multiple equilibria may result in incoherent estimators. Some innovative techniques have been recently proposed and applied to address this issue (e.g. Galichon & Henry (2011)), but they unfortunately cannot be applied here. In order to define a coherent estimator, I assume that the data is generated by the (unique) equilibrium that maximizes the potential function. Specifically:

Assumption 1 (Equilibrium Selection). *Any observed equilibrium* (y^*, g^*) *is such that:*
(2.1) $g_{ij} = 1$ *iff* $\mathbf{z}_{ij}\boldsymbol{\delta} - \frac{\lambda}{2}(y_i^* - y_j^*)^2 + \eta_{ij} > 0$ *for all* $i,j \in N$ *and*
(2.2) $y^* = \arg\max \psi(y)$.

The first condition assumes that the coordination problem between i and j has been solved so that every link with positive value is created. The second condition says that the observed equilibrium maximizes the potential function of the game.

Assumption 1 allows for the specification of a coherent estimator, but relies on the fact that one is able to compute $y^* = \arg\max_{y\in Y}\psi(y)$. As noted in the previous sections, the potential function is not globally concave, which makes its maximization infeasible in practice. The estimation strategy thus consists of using approximations of the potential function. I use the following result:

Proposition 6. *Let ϕ^* be the potential function's maximum. Also let:*
6.1 - $\mathbf{G}_0$ be such that $g_{0,ij} = 1$ iff $\mathbf{z}_{ij}\boldsymbol{\delta} + \eta_{ij} \geq 0$, and
6.2 - $\mathbf{G}_1$ be such that $g_{1,ij} = 1$ iff $\mathbf{z}_{ij}\boldsymbol{\delta} + \eta_{ij} - \frac{\lambda}{2}(\mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i - \mathbf{x}_j\boldsymbol{\beta} - \varepsilon_j)^2 \geq 0$
Then, $\phi(\mathbf{G}_0) \geq \phi^* \geq \phi(\mathbf{G}_1)$.

I then use $P(\mathbf{G}_0|\tilde{\mathbf{X}}, \mathbf{Z})$ and $P(\mathbf{G}_1|\tilde{\mathbf{X}}, \mathbf{Z})$ as approximations for $P(\mathbf{G}|\tilde{\mathbf{X}}, \mathbf{Z})$. Note that for $\mathbf{G}_0$, the QMLE in (8) can be estimated directly since the network structure is exogenous. Then, the model with an exogenous network structure represents one of the bounds of the true model. This model *underestimates* the cost of creating a link by abstracting from the cost imposed by y on the value of the links. The opposite is true for the specification for $\mathbf{G}_1$, where the approximation *overestimates* the cost of creating a link by only considering first-order effects.

I estimate the full approximated models using both specifications. I assume that $\mathbf{z}_{ij} = -|\mathbf{x}_i - \mathbf{x}_j|$ and that $\eta_{ij} \sim N(0, \sigma_\eta)$. Note that the specification for $P(\mathbf{G}_1|\tilde{\mathbf{X}}, \mathbf{Z})$ has to be simulated, since it does not allow for a closed-form density (see Train (2009)). Also note that $P(\mathbf{G}_1|\tilde{\mathbf{X}}, \mathbf{Z})$ allows for the identification of $\boldsymbol{\delta}/\sigma_\varepsilon$ and λ/σ_η (see proposition 6.2) so, together with (8), all the parameters of the model are identified. Finally, $P(\mathbf{G}_0|\tilde{\mathbf{X}}, \mathbf{Z})$ is a simple (scale-identified) probit model.

3.1. Results

As an initial benchmark, Table 2 presents a simple (misspecified) ordinary least squares (OLS) estimation. Participation in extracurricular activities decreases with age, but increases with the socio-economic status of the mother (level of education and employment status). Blacks and Asians participate more than whites and Hispanics. Table 3 presents the probit estimation for the network formation process (specification 6.1). The social network features homophily with respect to all variables, which is consistent with the literature.

Tables 4 and 5 present the quasi-maximum likelihood estimation for both bounds of the model. Table 4 presents results for specification (6.1), while Table 5 present results for specification (6.2). The estimated coefficients for both specifications are very similar, as are the log-likelihood values. As noted in the introduction, the small bias due to the endogeneity of the network is consistent with the literature. This should not be surprising looking at the marginal effects in Table 3. Although statistically significant, the contribution of the individual characteristics on the probability of a link are extremely small! Then, in the context of this particular database, the bias created by the endogeneity of the network is also small.

Remark also that for both specifications, the coefficients associated with individual effects are quite close to the OLS estimates. Also note that, in Table 5, none of the network formation parameters (i.e. $\boldsymbol{\delta}$) are statistically significant. A possible explanation is that $\mathbf{z}_{ij} = -|\mathbf{x}_i - \mathbf{x}_j|$ is highly correlated with $(\mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i - \mathbf{x}_j\boldsymbol{\beta} - \varepsilon_j)^2$ so the additional contribution of $\mathbf{z}_i$ is not significant.

The estimated values for the main parameter of interest, $\ln(\lambda)$, are -2.5204 and -2.5149 for $\mathbf{G}_0$ and $\mathbf{G}_1$, respectively. This leads to values of λ of 8.04% and 8.08%. However, as discussed above, the magnitude of peer effects depends on the numbers of friends an individual has. Recall that:

$$y_i = \frac{\lambda}{1+\lambda n_i}\sum_{j\in N_i} y_j + \frac{1}{1+\lambda n_i}[x_i\boldsymbol{\beta} + \varepsilon_i]$$

Figure 3 displays the distribution of the implied peer effects using the observed distribution of the number of peers (i.e. $\frac{\lambda}{1+\lambda n_i}$, using $\lambda = 8.08\%$). The impact of a change in *one* peer's behavior ranges from 4.5% (for an individual with 10 peers) to 7.5% (for an individual with only one peer). Put differently, when choosing the number of activities he participates in, a student with only one friend puts a weight of 92.5% on his individual characteristics (i.e. his type), while a student with 10 friends puts a weight of only 55% on his individual characteristics (i.e. $\frac{1}{1+\lambda n_i}$, using $\lambda = 8.08\%$).

Finally, Figure 4 presents, for each school, the variance in behavior, the (expected) lower-bound on that variance (see proposition 4) and the variance in individuals' types. The average value of the lower-bound is 3.23 (based on centrality), the average variance in behavior is 6.70 and the average variance in individuals' types is 8.07. As the observed networks in every schools are disconnected, the upper-bound is never binding.

I now discuss the policy implications of the model.

3.2. Implications for Public Policy

The model features both peer effects and self-selection. This suggests that a policymaker can affect an equilibrium through two different channels: by *using* the shape of the network, or by *affecting* the shape of the network. I now discuss these two approaches.

Suppose that the policymaker wants to use the shape of network to influence the equilibrium outcome. Specifically, suppose that the policymaker has the ability to increase the type $x_i\boldsymbol{\beta}+\varepsilon_i$ of any individual by a small amount. Which individual should he affect? The typical answer to that question comes from Ballester et al. (2006): the most central individual in the network. In the context of this paper, the aggregate impact on the *sum* of behaviors is the same, irrespective on the individual selected. This follows from the fact that $\mathbf{1}$ is the eigenvector of $\mathbf{L}$ associated with the eigenvalue 0. The impact on the *variance* of behaviors, however, depends non-trivially on the network structure, as well as on individuals' types, and cannot directly be linked to a centrality measure (see proposition 4).

In any case, the precise answer to the question of which individual the policymaker should affect is likely to be irrelevant, since (1) the measured impact of peer effect is relatively small (the value of $\frac{\lambda}{1+\lambda n_i}$ ranges from 4.5% to 7.5%), and (2) the analysis would only hold locally since large policy shocks will also affect the structure of the network. The policymaker should then focus on affecting the network formation process.

According to proposition 4, a policymaker can affect the network structure by increasing the network centrality, or by increasing the edge-connectivity. As noted above, every school in the sample is disconnected, so the edge-connectivity is equal to 0. In order to increase the edge-connectivity, a policymaker would have to connect the network.[11] Figure 5 shows how many links would have to

[11] That is, adding enough links so that the network has only one component.

Table 2: OLS - Within-Group Deviations

Variable	**Estimate**	**S.E.**
Age	-0.0387**	(0.0132)
Gender	0.2288**	(0.0300)
Hispanic	0.1119*	(0.0551)
White	0.2677**	(0.0517)
Black	0.4233**	(0.0609)
Asian	0.5514**	(0.0749)
Mother works	0.1426**	(0.0356)
Mother completed HS	0.1337**	(0.0421)
Mother completed college	0.3010**	(0.0356)

be added to the network, as a function of the school size. We see that for large schools, this number can be extremely high. Figure 6 shows the relative gain of connecting the network. We see that for large schools, the impact is quite small. This suggests that for small schools, policymakers should focus on promoting integration, while for larger schools, policymakers should focus on promoting role models.

4. Conclusion

I presented a model of conformism featuring both peer effects and self-selection. I found that the network structure and the equilibrium behaviors are linked through the Laplacian matrix of the network. This particular relationship implies that the number of peers influences the magnitude of peer effects. Moreover, social welfare can be bounded by the centrality and edge-connectivity of the equilibrium network.

Using the theoretical predictions of the model, I measure the impact of conformism on student participation in extracurricular activities. I find that the magnitude of the endogenous peer effects ranges from 7.5% to 45% depending on an individual's number of peers. Results indicate that small schools should focus on promoting integration, while large schools should focus on creating role models.

Table 3: Probit (see proposition 6.1) - Marginal Effects †

Variable	**Estimate**	**S.E.**
Age	3.2727**	(0.0393)
Gender	3.522**	(0.0579)
Hispanic	1.1593**	(0.0806)
White	1.9104**	(0.0761)
Black	4.0803**	(0.1108)
Asian	1.2784**	(0.1071)
Mother works	0.4781**	(0.0540)
Mother completed HS	1.0078**	(0.0591)
Mother completed college	0.3334**	(0.0519)
Log-likelihood: -141 381.53		

† For each $z_{ij} = -|x_i - x_j|$, coefficients are multiplied by a factor of 1000 for ease of interpretation.

Table 4: QMLE - Exogenous Network

Variable	**Estimate**	**S.E.**
$\ln(\lambda)$	-2.5204**	(0.1317)
Age	-0.0389	(0.0223)
Gender	0.2496**	(0.0434)
Hispanic	0.1284	(0.0978)
White	0.2657**	(0.0813)
Black	0.4364**	(0.0955)
Asian	0.5653**	(0.1206)
Mother works	0.1495**	(0.0431)
Mother completed HS	0.128*	(0.0503)
Mother completed college	0.3262**	(0.0459)
$\ln(\sigma_\varepsilon)$	1.0382**	(0.0236)
Log-likelihood: -432.7814		

Table 5: QMLE - Full Model (Specification 6.2)

Variable	**Estimate**	**S.E.**
$\ln(\lambda)$	-2.5149**	(0.1257)
Age	-0.0475*	(0.0220)
Gender	0.2502**	(0.0435)
Hispanic	0.2027*	(0.0982)
White	0.3031**	(0.0809)
Black	0.5824**	(0.0969)
Asian	0.6289**	(0.1213)
Mother works	0.0867	(0.0448)
Mother completed HS	0.2098**	(0.0514)
Mother completed college	0.3259**	(0.0462)
$\ln(\sigma_\varepsilon)$	1.0380**	(0.0232)
Network Formation		
Intercept	0.0155	(0.0144)
Age	0.0048	(0.005)
Gender	-0.0137	(0.0156)
Hispanic	0.0050	(0.0089)
White	-0.0081	(0.0126)
Black	0.0062	(0.0127)
Asian	-0.0031	(0.0179)
Mother works	0.0013	(0.0035)
Mother completed HS	0.0065	(0.0079)
Mother completed college	0.0022	(0.0019)
$\ln(\sigma_\varepsilon)$	-2.9911**	(0.6309)
Log-likelihood: -432.9868		

Figure 3: Local Marginal Effect for *One* Peer ($\frac{\lambda}{1+\lambda n_i}$)

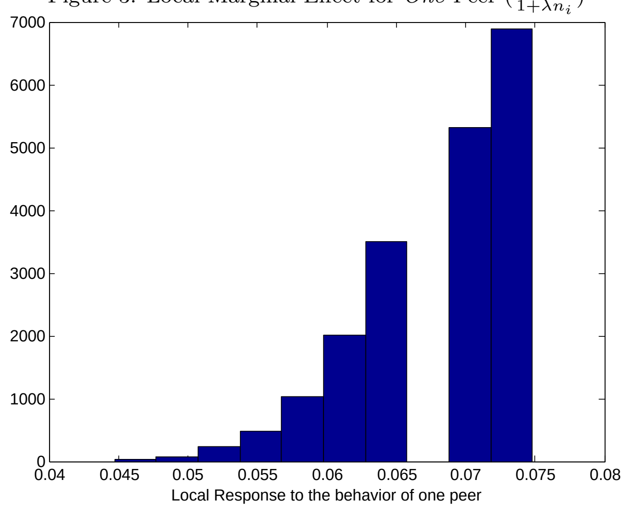
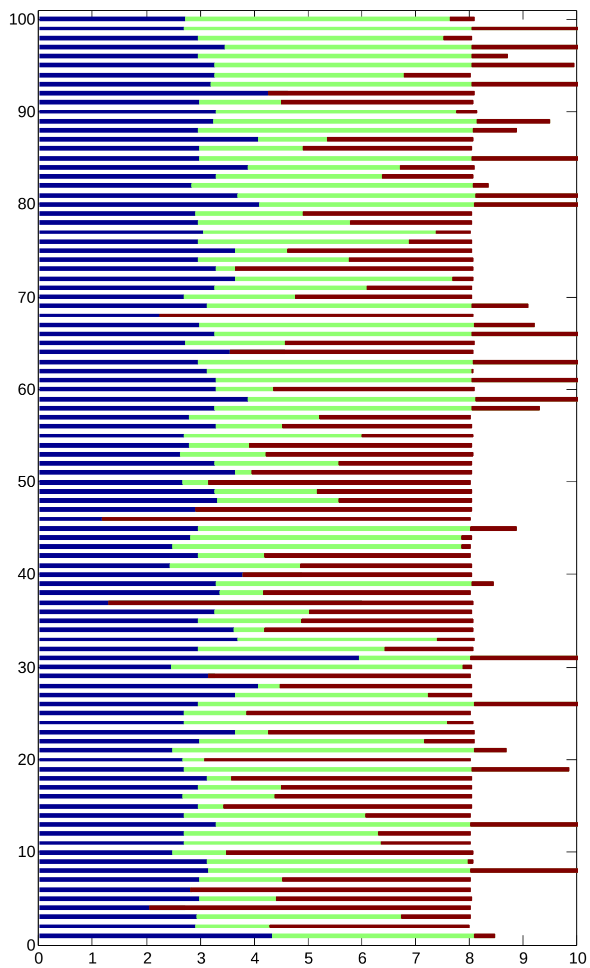


Figure 4: Variance of Types and Behaviors (All Schools)†



† For expected variance of the types, using 1000 draws. Variances are truncated at 10 to clarify the presentation. The upper-bound is never binding as all schools are disconnected (the value of the upper-bound is thus always equal to the variance of the types). The average value of the lower-bound is 3.23. The average variance of behavior is 6.70. The average variance of the types is 8.07.

Figure 5: Number of Links Needed to Connect the Network as a Function of School Size

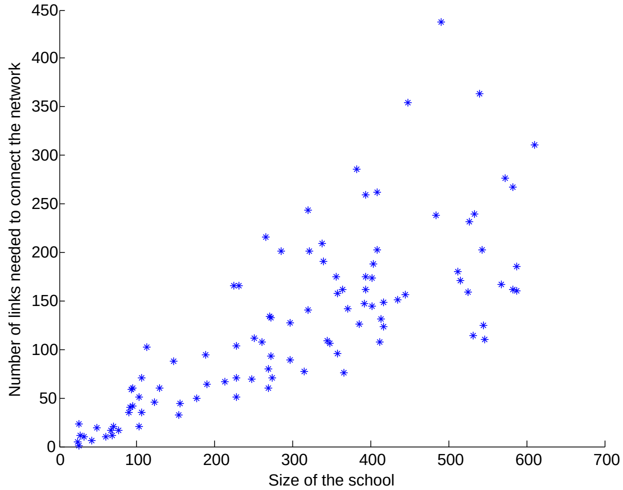
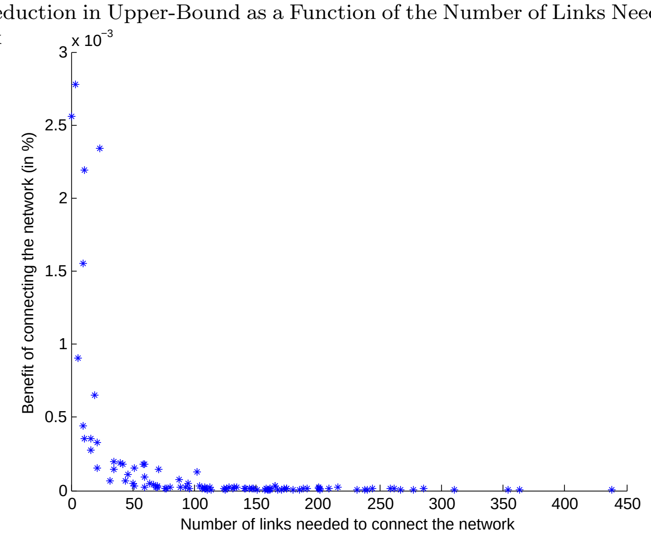


Figure 6: Reduction in Upper-Bound as a Function of the Number of Links Needed to Connect the Network



5. References

Badev, A. (2013). Discrete games in endogenous networks: Theory and policy.

Bai, H. (2011). The grone-merris conjecture. *Transactions of the American Mathematical Society*, *363*(8), 4463–4474.

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

Bisin, A., Horst, U., & Özgür, O. (2006). Rational expectations equilibria of economies with local interactions. *Journal of Economic Theory*, *127*(1), 74–116.

Bisin, A., & Özgür, O. (2012). Dynamic linear economies with social interactions. *Working Paper*.

Blume, L. E., Vienna, I., Brock, W. A., Durlauf, S. N., Jayaraman, R., & ESMT, B. (2011). Linear social network models. Tech. rep., Working paper.

Boucher, V., & Fortin, B. (2014). Recent developments in the empirics of network effects. *Forthcoming in The Oxford Handbook on the Economics of Networks, Editors: Yann Bramoullé, Andrea Galeotti, Brian W. Rogers*.

Carbonell-Nicolau, O., & McLean, R. P. (2011). Refinements of nash equilibrium in potential games. Tech. rep., Working Papers, Department of Economics, Rutgers, the State University of New Jersey.

Christakis, N. A., Fowler, J. H., Imbens, G. W., & Kalyanaraman, K. (2010). An empirical model for strategic network formation. Tech. rep., National Bureau of Economic Research.

Galichon, A., & Henry, M. (2011). Set identification in models with multiple equilibria. *The Review of Economic Studies*, (p. rdr008).

Goldsmith-Pinkham, P., & Imbens, G. W. (2013). Social networks and the identification of peer effects. *Journal of Business & Economic Statistics*, *31*(3), 253–264.

Goyal, S., & Vega-Redondo, F. (2007). Structural holes in social networks. *Journal of Economic Theory*, *137*(1), 460–492.

Herman, P. (2013). Endogenous network formation and the provision of a public good.

Hojman, D. A., & Szeidl, A. (2006). Endogenous networks, social games, and evolution. *Games and Economic Behavior*, *55*(1), 112–130.

Hsieh, C.-S., & Lee, L.-F. (2011). A social interactions model with endogenous friendship formation and selectivity. Tech. rep., Working paper.

Jackson, M. O., & Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of economic theory*, *71*(1), 44–74.

Kinateder, M., & Merlino, L. P. (2014). Public goods in endogenous networks.

Lee, L.-f., Liu, X., & Lin, X. (2010). Specification and estimation of social interaction models with network structures. *The Econometrics Journal*, *13*(2), 145–176.

Liu, X., Patacchini, E., & Zenou, Y. (2011). *Peer effects in education, sport, and screen activities: local aggregate or local average?*. Centre for Economic Policy Research.

Monderer, D., & Shapley, L. S. (1996). Potential games. *Games and economic behavior*, *14*(1), 124–143.

Newman, M. W., Libraty, N., On, O., On, K. A., On, O., On, K. A., & Newman, M. W. (2000). The laplacian spectrum of graphs.

Patacchini, E., & Rainone, E. (2014). The word on banking-social ties, trust, and the adoption of financial products. Tech. rep., Einaudi Institute for Economics and Finance (EIEF).

Patacchini, E., & Zenou, Y. (2012). Juvenile delinquency and conformism. *Journal of Law, Economics, and Organization*, *28*(1), 1–31.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies*, *70*(1), 147–165.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Ui, T. (2001). Robust equilibria of potential games. *Econometrica*, *69*(5), 1373–1380.

6. Appendix: Proofs

Proof (of Proposition 1). To simplify the notation, let $\upsilon_{ij} = \mathbf{z}_{ij}\boldsymbol{\delta} + \eta_{ij}$. Then, simply let g^* be such that [$g^*_{ij} = 1$ iff $\upsilon_{ij} \geq \frac{\lambda}{2}(y_i - y_j)^2$] for all $i, j \in N$. $\square$

Proof (f Lemma 2).

Step 1: $\mathbf{z}_{ij}\boldsymbol{\delta} - \frac{\lambda}{2}(y_i - y_j)^2 + \eta_{ij} \neq 0$

Consider y_i that maximizes the utility w_i. Note that for all the ij such that $\mathbf{z}_{ij}\boldsymbol{\delta} + \eta_{ij} \leq 0$, any y_i leads to the same value for the link, as the max is equal to 0. Then, a small perturbation of η_{ij} will lead to the same optimal choice of y_i. Now, consider $\upsilon_{ij} = \mathbf{z}_{ij}\boldsymbol{\delta} + \eta_{ij} > 0$ and define $t_i = \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i$ and the following sets:

$$
\begin{aligned}
A_i(y) &= \left\{ j \neq i | y_i \in \left[y_j - \sqrt{\frac{2\upsilon_{ij}}{\lambda}}, y_j + \sqrt{\frac{2\upsilon_{ij}}{\lambda}} \right) \right\} \\
B_i(y) &= \left\{ j \neq i | y_i \in \left(y_j - \sqrt{\frac{2\upsilon_{ij}}{\lambda}}, y_j + \sqrt{\frac{2\upsilon_{ij}}{\lambda}} \right] \right\} \\
C_i(y) &= \left\{ j \neq i | y_i \in \left(y_j - \sqrt{\frac{2\upsilon_{ij}}{\lambda}}, y_j + \sqrt{\frac{2\upsilon_{ij}}{\lambda}} \right) \right\} \\
\tilde{A}_i(y) &= A_i(y) \setminus C_i(y) \\
\tilde{B}_i(y) &= B_i(y) \setminus C_i(y)
\end{aligned}
$$

Now, define the following directional derivatives:

$$
\begin{aligned}
\partial_+(y_i) &= -(y_i - t_i) - \lambda \sum_{j \in A_i(y)} (y_i - y_j) \\
\partial_-(y_i) &= -(y_i - t_i) - \lambda \sum_{j \in B_i(y)} (y_i - y_j)
\end{aligned}
$$

It is easy to see that y_i^* maximizes $w_i(y_i, y_{-i})$ only if $\partial_+(y_i) \leq 0$ and $\partial_-(y_i) \geq 0$, or equivalently if:

$$
\begin{aligned}
-(y_i - t_i) &\leq \lambda \sum_{j \in A_i(y)} (y_i - y_j) \\
-(y_i - t_i) &\geq \lambda \sum_{j \in B_i(y)} (y_i - y_j)
\end{aligned}
$$

These two conditions can be met simultaneously only if

$$
\sum_{j \in A_i(y)} (y_i - y_j) \geq \sum_{j \in B_i(y)} (y_i - y_j)
$$

or equivalently if

$$\sum_{j\in\tilde{A}_i(y)}(y_i-y_j)\geq\sum_{j\in\tilde{B}_i(y)}(y_i-y_j)$$

Now, note that any $j\in\tilde{A}_i(y)$ is such that $y_i-y_j=-\sqrt{\frac{2\upsilon_{ij}}{\lambda}}$, and any $j\in\tilde{B}_i(y)$ is such that $y_i-y_j=\sqrt{\frac{2\upsilon_{ij}}{\lambda}}$. Thus, it is required that

$$-\sum_{j\in\tilde{A}_i(y)}\sqrt{\frac{2\upsilon_{ij}}{\lambda}}\geq\sum_{j\in\tilde{B}_i(y)}\sqrt{\frac{2\upsilon_{ij}}{\lambda}}$$

which is only possible if $|\tilde{A}_i(y)|=|\tilde{B}_i(y)|=0$.

Step 2: The maximum exists and is generically unique.

Existence follows from the fact that any $y_i<\min\{y_j,t_i\}$ or $y_i>\max\{y_j,t_i\}$ is dominated by some $y_i'\in[\min\{y_j,t_i\},\max\{y_j,t_i\}]$. Now, suppose that there exists two maxima: y_i,y_i'. From step 1, this implies that y_i and y_i' do not induce the same network structure, i.e. there exists j such that $\upsilon_{ij}-\frac{\lambda}{2}(y_i-y_j)^2<0$ while $\upsilon_{ij}-\frac{\lambda}{2}(y_i'-y_j)^2>0$. Then, taking a small perturbation of η_{ij}, the first-order conditions still hold for y_i and y_i', but they no longer yield the same utility. □

Proof (of Proposition 3).
Step 1: $\mathbf{y}=[\mathbf{I}+\lambda\mathbf{L}]^{-1}[\mathbf{X}\boldsymbol{\beta}+\varepsilon]$

Let (g^*,y^*) be a NE of Γ. Since the strategy space for y_i is unbounded and from lemma 2, any NE is interior for y_i, the first-order conditions need to apply. At g^*, for all $i\in N$, the following holds: $0=\sum_{j:g^*_{ij}=g^*_{ji}=1}-\lambda(y_i-y_j)-y_i+\mathbf{x}_i\boldsymbol{\beta}+\varepsilon_i$, which is equivalent to $y_i=-\lambda n_i^*y_i+\sum_{j:g^*_{ij}=g^*_{ji}=1}\lambda y_j+\mathbf{x}_i\boldsymbol{\beta}+\varepsilon_i$. In matrix form, we have:

$$\mathbf{y}=-\lambda(\mathbf{D}-\mathbf{G})\mathbf{y}+\mathbf{X}\boldsymbol{\beta}+\varepsilon$$

Defining $\mathbf{L}=\mathbf{D}-\mathbf{G}$, which is positive semi-definite (Godsil and Royle, 2001, p.280), we have: $\mathbf{y}=[\mathbf{I}+\lambda\mathbf{L}]^{-1}[\mathbf{X}\boldsymbol{\beta}+\varepsilon]$.

Step 2: $E(y_i^*)=E(\mathbf{x}_i\boldsymbol{\beta}+\varepsilon_i)$

Remark that $\mathbf{1}$ is the eigenvector of $\mathbf{L}$ associated with eigenvalue 0, so

$$\frac{1}{n}\mathbf{1}'[\mathbf{I}+\lambda\mathbf{L}]\mathbf{y}=\frac{1}{n}\mathbf{1}'\mathbf{y}=\frac{1}{n}\mathbf{1}'[\mathbf{X}\boldsymbol{\beta}+\varepsilon]$$

Step 3: $Var(y_i^*)\leq Var(\mathbf{x}_i\boldsymbol{\beta}+\varepsilon_i)$.

We have:

$$Var(y_i)=\frac{1}{n}\mathbf{y}'\mathbf{y}-E(y_i)^2+Var(\mathbf{x}_i\boldsymbol{\beta}+\varepsilon_i)-\frac{1}{n}[\mathbf{X}\boldsymbol{\beta}+\varepsilon]'[\mathbf{X}\boldsymbol{\beta}+\varepsilon]+E(\mathbf{x}_i\boldsymbol{\beta}+\varepsilon_i)^2$$

which, from step 2 above, is equivalent to:

$$Var(y_i) = \frac{1}{n}\mathbf{y}'\mathbf{y} + Var(\mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i) - \frac{1}{n}[\mathbf{X}\boldsymbol{\beta} + \varepsilon]'[\mathbf{X}\boldsymbol{\beta} + \varepsilon]$$

Rewriting, we have:

$$Var(y_i) = Var(\mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i) - \frac{\lambda}{n}[\mathbf{X}\boldsymbol{\beta} + \varepsilon]'(\mathbf{I} + \lambda\mathbf{L})^{-2}(2\mathbf{I} + \lambda\mathbf{L})\mathbf{L}[\mathbf{X}\boldsymbol{\beta} + \varepsilon]$$

Since the product of jointly diagonalizable positive semi-definite matrices is positive definite, $(\mathbf{I} + \lambda\mathbf{L})^{-2}(2\mathbf{I} + \lambda\mathbf{L})\mathbf{L}$ is positive semi-definite (this can be seen by applying the Spectral Theorem), with completes the proof. $\square$

PROOF (OF PROPOSITION 4). Let $0 = \xi_1 \leq ... \leq \xi_n$ be the eigenvalues of $\mathbf{L}$. Using the expression in step 3 of the proof of proposition 3 and applying the Spectral Theorem, we have:

$$Var(y_i) = Var(\mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i) - \frac{\lambda}{n}[\mathbf{X}\boldsymbol{\beta} + \varepsilon]'\mathbf{T}'\mathbf{U}\left(\frac{2\xi_i + \lambda\xi_i^2}{(1 + \lambda\xi_i)^2}\right)\mathbf{T}[\mathbf{X}\boldsymbol{\beta} + \varepsilon]$$

where $\mathbf{T}$ is orthonormal and $\mathbf{U}(a_i)$ is a diagonal matrix with entries $a_1, ..., a_n$. The bounds are then simply obtained by using the following bounds (see Newman et al. (2000)):

$$\begin{aligned} \xi_n &\leq \max_{ij} C_{ij} \\ \xi_1 &\geq 2e(1 - cos(\pi/n)) \end{aligned}$$

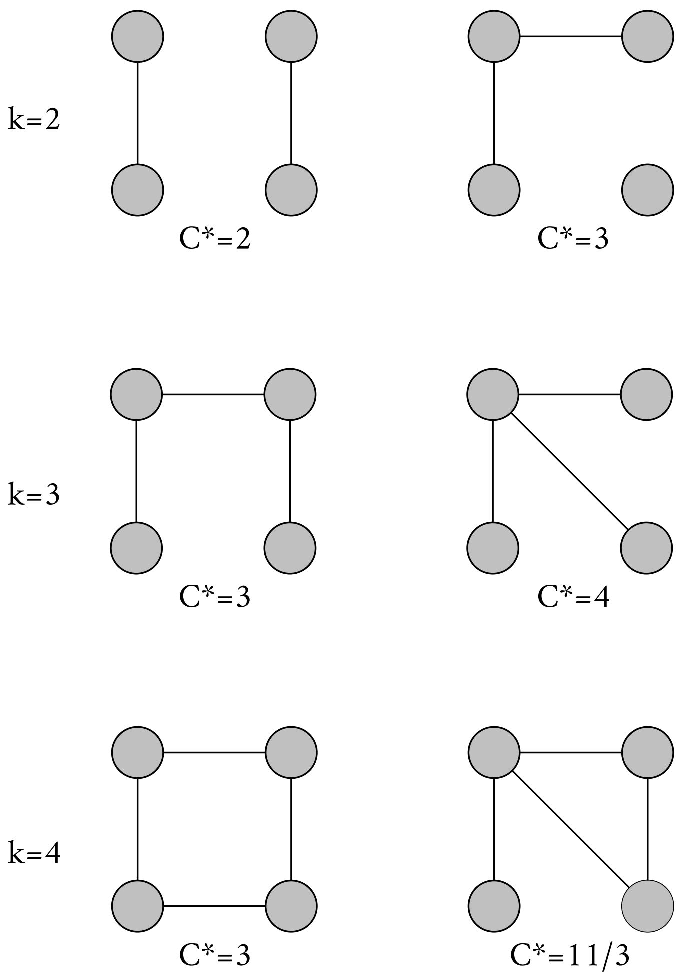
$\square$

PROOF (OF PROPOSITION 5). I now show that the potential function admits a generically unique maximum.

Let $t_i = \mathbf{x}_i\boldsymbol{\beta} + \varepsilon_i$. Without a loss of generality, assume that $t_1 \leq ... \leq t_n$. It is sufficient to show that any y is (weakly) dominated by some y^* such that $y_i^* \in [t_1, t_n]$ for all $i \in N$.

Consider y and a non-empty $M \subseteq N$ such that for all $i \in M$, $y_i \notin [t_1, t_n]$, while $y_i \in [t_1, t_n]$ for all $i \in N \setminus M$. Also define $M_1 = \{i \in N | y_i < t_1\}$ and $M_n = \{i \in N | y_i > t_n\}$ so that $M = M_1 \cup M_n$. For any y, we can write $\psi(y)$ as

Figure 7: Example for $n = 4$.



(setting empty sums to 0):

$$
\begin{aligned}
\psi(y) &= \sum_{i,j\in M_1} \max\{\upsilon_{ij} - \frac{\lambda}{2}(y_i - y_j)^2, 0\} + \sum_{i,j\in M_n} \max\{\upsilon_{ij} - \frac{\lambda}{2}(y_i - y_j)^2, 0\} \\
&+ \sum_{i\in M_1, j\in M_n} \max\{\upsilon_{ij} - \frac{\lambda}{2}(y_i - y_j)^2, 0\} \\
&+ \sum_{i\in M_1, j\in N\setminus M} \max\{\upsilon_{ij} - \frac{\lambda}{2}(y_i - y_j)^2, 0\} \\
&+ \sum_{i\in M_n, j\in N\setminus M} \max\{\upsilon_{ij} - \frac{\lambda}{2}(y_i - y_j)^2, 0\} \\
&+ \sum_{i,j\in N\setminus M} \max\{\upsilon_{ij} - \frac{\lambda}{2}(y_i - y_j)^2, 0\} \\
&- \frac{1}{2}\sum_{i\in M_1}(y_i - t_i)^2 - \frac{1}{2}\sum_{i\in M_n}(y_i - t_i)^2 - \frac{1}{2}\sum_{i\in N\setminus M}(y_i - t_i)^2
\end{aligned}
$$

Now generate the allocation y^* such that: $y_i^* = y_i \ \forall i \in N \setminus M$, $y_i^* = t_1$ $\forall i \in M_1$ and $y_i^* = t_n \ \forall i \in M_n$. We have:

$$
\begin{aligned}
\psi(y^*) &= \sum_{i,j\in M_1} \max\{\upsilon_{ij}, 0\} + \sum_{i,j\in M_n} \max\{\upsilon_{ij}, 0\} \\
&+ \sum_{i\in M_1, j\in M_n} \max\{\upsilon_{ij} - \frac{\lambda}{2}(z_1 - z_n)^2, 0\} \\
&+ \sum_{i\in M_1, j\in N\setminus M} \max\{\upsilon_{ij} - \frac{\lambda}{2}(z_1 - y_j)^2, 0\} \\
&+ \sum_{i\in M_n, j\in N\setminus M} \max\{\upsilon_{ij} - \frac{\lambda}{2}(z_n - y_j)^2, 0\} \\
&+ \sum_{i,j\in N\setminus M} \max\{\upsilon_{ij} - \frac{\lambda}{2}(y_i - y_j)^2, 0\} - \frac{1}{2}\sum_{i\in M_1}(t_1 - t_i)^2 \\
&- \frac{1}{2}\sum_{i\in M_n}(t_n - t_i)^2 - \frac{1}{2}\sum_{i\in N\setminus M}(y_i - t_i)^2
\end{aligned}
$$

One can easily verify that $\psi(y^*) > \psi(y)$, which leads to a contradiction.

Now, suppose that the equilibrium is not unique. There exists $\tilde{y}, y^* \in \arg\max_y \psi(y)$. Note that, by lemma 2, it implies that they do not induce the same network structure. Then, there exists $i, j \in N$ such that $\upsilon_{ij} - \frac{\lambda}{2}(\tilde{y}_i - \tilde{y}_j)^2 > 0$, while $\upsilon_{ij} - \frac{\lambda}{2}(y_i^* - y_j^*)^2 < 0$ (also by lemma 2). Hence, by taking a small perturbation of η_{ij} such that $\tilde{\upsilon}_{ij} > \upsilon_{ij}$, we have $\tilde{y} \in \arg\max_y \tilde{\psi}(y)$ (by lemma 2) and $\tilde{\psi}(\tilde{y}) > \tilde{\psi}(y^*)$. $\square$

Proof (of Proposition 6). Remark that the potential function can be written as follows:

$$\sum_{ij\in g^*} \mathbf{z}_{ij}\boldsymbol{\delta} + \eta_{ij} - \frac{\lambda}{2}\mathbf{y}^{*\prime}\mathbf{L}\mathbf{y}^* - \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \varepsilon)'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \varepsilon)$$

Using proposition 3, this is equivalent to

$$\sum_{ij\in g^*} \mathbf{z}_{ij}\boldsymbol{\delta} + \eta_{ij} + \frac{1}{2}(\mathbf{X}\boldsymbol{\beta} + \varepsilon)' \left[\mathbf{I} + \lambda\mathbf{L}\right]^{-1} (\mathbf{X}\boldsymbol{\beta} + \varepsilon)$$

We have $\mathbf{I} \leq \mathbf{I} + \lambda\mathbf{L}$ (in the sense of definiteness), which implies that $\mathbf{I} \geq [\mathbf{I} + \lambda\mathbf{L}]^{-1}$ so $(\mathbf{X}\boldsymbol{\beta} + \varepsilon)'(\mathbf{X}\boldsymbol{\beta} + \varepsilon) \geq (\mathbf{X}\boldsymbol{\beta} + \varepsilon)' \left[\mathbf{I} + \lambda\mathbf{L}\right]^{-1} (\mathbf{X}\boldsymbol{\beta} + \varepsilon)$. Similarly, we have $\mathbf{I} \geq \mathbf{I} - \lambda^2\mathbf{L}^2$, which is equivalent to $\mathbf{I} \geq [\mathbf{I} - \lambda\mathbf{L}][\mathbf{I} + \lambda\mathbf{L}]$ and to $[\mathbf{I} + \lambda\mathbf{L}]^{-1} \geq [\mathbf{I} - \lambda\mathbf{L}]$. Finally, it implies that $(\mathbf{X}\boldsymbol{\beta} + \varepsilon)'[\mathbf{I} + \lambda\mathbf{L}]^{-1}(\mathbf{X}\boldsymbol{\beta} + \varepsilon) \geq (\mathbf{X}\boldsymbol{\beta} + \varepsilon)'(\mathbf{X}\boldsymbol{\beta} + \varepsilon) - \lambda(\mathbf{X}\boldsymbol{\beta} + \varepsilon)'\mathbf{L}(\mathbf{X}\boldsymbol{\beta} + \varepsilon)$. $\square$