Social Norms in Networks*

Philip Ushchev† Yves Zenou‡

October 28, 2019

Abstract

Although the linear-in-means model is the workhorse model in empirical work on peer effects, its theoretical properties are understudied. In this study, we develop a social-norm model that provides a microfoundation of the linear-in-means model and investigate its properties. We show that individual outcomes may increase, decrease, or vary non-monotonically with the taste for conformity. Equilibria are usually inefficient and, to restore the first best, the planner needs to subsidize (tax) agents whose neighbors make efforts above (below) the social norms. Thus, giving more subsidies to more central agents is not necessarily efficient. We also discuss the policy implications of our model in terms of education and crime.

Keywords: Social norms, conformism, local-average model, welfare, anti-conformism, network formation.

JEL Classification: D85, J15, Z13.

*We thank the editor, three anonymous referees, Arthur Campbell, Ben Golub, Matt O. Jackson, Michelle Rendall, and the participants of seminars at the University of Adelaide, Monash University, the University of Technology Sydney, and the Fifth Annual Network Science and Economics Conference, Bloomington, April 2019, for their helpful comments. Ushchev acknowledges that the article was prepared within the framework of the HSE University Basic Research Program and funded by the Russian Academic Excellence Project '5-100'.

†National Research University Higher School of Economics, Russia, and Monash University, Australia. Email: ph.ushchev@gmail.com.

‡Monash University, Australia, and CEPR. E-mail: yves.zenou@monash.edu.

1 Introduction

There is substantial empirical evidence showing that peer effects matter in education (Calvó-Armengol et al., 2009; Epple and Romano, 2011; Sacerdote, 2011), crime (Ludwig et al., 2001; Patacchini and Zenou, 2012; Damm and Dustmann, 2014), risky behavior (Clark and Loheac, 2007; Hsieh and Lin, 2017), performance in the workplace (Herbst and Mas, 2015), participation in extracurricular activities (Boucher, 2016), obesity (Christakis and Fowler, 2007), environmentally friendly behavior (Brekke et al., 2010; Czajkowski et al., 2017), and tax compliance and tax evasion (Fortin et al., 2007; Alm et al., 2017), among other outcomes. The standard model used in these studies is the so-called *linear-in-means model*, which can be written as

$$x_{ig} = z_{ig}\beta + y_g\gamma + \frac{\theta}{(N_g - 1)} \sum_{j=1, j\neq i}^{N_g} x_{jg} + \epsilon_{ig} \tag{1}$$

where x_{ig} is the outcome of individual i belonging to group g,[1] z_{ig} are the observable characteristics of individual i (e.g., age, race, and gender), y_g are the observed exogenous characteristics that are common to all individuals in the same group g,[2] N_g is the number of individuals in group g, and ϵ_{ig} is an error term. Parameter θ captures the "social interaction effect" of the average outcome of the reference group on an individual's own outcome; this is the key parameter of interest that is estimated to measure peer effects.[3]

As noted by Blume et al. (2015), Boucher and Fortin (2016), and Kline and Tamer (2018), it is useful to interpret the linear-in-means model as corresponding to a perfect information game in which (1) is the best-reply function of individual i choosing action (outcome) x_i. The corresponding utility function is such that individuals have a preference to conform to the *average* action of their neighbors in a social network. For this reason, this game is often referred to as the *local-average model*. Surprisingly, the theoretical properties of this model in terms of comparative

[1] For example, in relation to crime, x_{ig} is the criminal effort of individual i in neighborhood g and, in relation to education, it is the test score of student i in classroom g.

[2] For example, y_g are the average education or income level in a neighborhood g or the average education or income level of students' parents in a classroom g.

[3] If all agents belong to the same group g, this model is not identified, because it is difficult to distinguish between the endogenous effect θ and the exogenous effect γ. Manski (1993) referred to this as the *reflection problem,* because it is difficult to distinguish between an individual's behavior and the behavior being "reflected" back on the individual. The literature on peer effects has proposed different ways of causally interpreting θ, including field experiments that randomly allocate individuals to groups (see, e.g., Sacerdote, 2011, for an overview of peer effect studies in education).

statics, welfare, and policies have not been investigated. On the contrary, the literature on games on networks[4] (Ballester et al., 2006; Bramoullé et al., 2014; Jackson and Zenou, 2015; Bramoullé and Kranton, 2016)[5] studies the properties of another model, the *local-aggregate model*, in which the sum (not the average) of actions (or outcomes) of neighbors affects own action.[6]

Thus, there is a discrepancy between the theoretical analysis of the local-aggregate model and the empirical applications using the linear-in-means model or local-average model. In this study, we analyze the comparative statics, welfare properties, and policy implications of the local average model and show that these properties are very different from those of the local-aggregate model.[7] Indeed, we show that the differences between the local aggregate and the local average, although seemingly minor, lead to substantial divergence in both positive and normative prescriptions. In other words, the local-aggregate model fails to approximate the local-average model in each of the following key dimensions: comparative statics, welfare properties, and policy recommendations.

Our main findings are summarized as follows. First, we characterize the Nash equilibrium in the local-average model and show that individual efforts, social norms,[8] and aggregate effort are the weighted sums of productivity, whereby the weights are

[4]The economics of networks is a growing field. For overviews, see Jackson (2008), Ioannides (2012), and Jackson et al. (2017).

[5]One can interpret the group g in (1) in terms of networks so that group g captures all agents who individual i is connected to. In that case, the game underlying the linear-in-means model is a game on networks in which $N_g - 1$ is the number of agents who are directly connected (direct friends) to i.

[6]The key difference between the local-average and the local-aggregate model is that the former aims to capture the role of *social norms*, such as conformist behavior or peer pressure, on outcomes (Patacchini and Zenou, 2012; Liu et al., 2014; Blume et al., 2015; Topa and Zenou, 2015; Boucher, 2016), while the latter highlights the role of knowledge spillovers on outcomes (Ballester et al., 2006, 2010; Bramoullé et al., 2014; De Marti and Zenou, 2015). Bramoullé et al. (2009) provide conditions for identification in the local-average model while Liu et al. (2014) derive conditions for identification in the local-aggregate model.

[7]In this study, we are interested only in *positive* peer effects, which is why we compare the local-aggregate model with the local-average one—both are games with *strategic complementarities*; that is, an increase in the effort of a neighbor increases the marginal utility of own effort. Another well-studied model in network games is a game with strategic substitutability (Bramoullé and Kranton, 2007; Bramoullé et al., 2014; Allouch, 2015) in which there are *negative* peer effects, that is, an increase in the effort of an individual's neighbor decreases the marginal utility of making own effort. This is not the topic of analysis in this study, since we focus on the linear-in-means model in which peer effects are supposed to be positive.

[8]There are different definitions of social norms in the literature (see, e.g., Akerlof, 1997; Dutta et al., 2019). Here, we define the social norm of an agent as the average action of her neighbors.

non-linear functions of the taste for conformity. To understand these results, we compare two extreme cases: *pure individualism* and *total conformism*. Under pure individualism, each agent's equilibrium effort is equal to her intrinsic productivity and is independent of her own social norm. By contrast, under total conformism, all agents choose the same level of effort, which is equal to the weighted mean of individual productivity, whereby the weights are proportional to the degree (numbers of links) of the agents in the network. Whether total effort is higher under pure individualism or total conformism depends on the correlation between the productivity distribution across individuals and the degree distribution of the social network.

Second, we provide comparative statics of individual and aggregate efforts with respect to the key parameters of the model. We focus especially on the taste for conformity. Endogenous social norms give rise to general-equilibrium effects. A complex interplay between these effects may result in a non-monotonic relationship between the taste for conformity and individual efforts. Whether an individual is above or below her social norm is key for understanding the shape of this relationship. Interestingly, in regular networks, aggregate effort remains neutral to changes in the taste for conformity and is always equal to aggregate productivity.

We also study the impact of adding a link on the equilibrium efforts of all agents in the network. *All* agents in the network increase their effort if and only if a link between two agents with sufficiently high productivities is added in the network. This result is driven by the following *snowball effect*. When a link is formed between two very productive agents, their social norm increases, because the effort of the newly added agent is high. The best response for the agent for whom the social norm increases is to increase her effort. This, in turn, increases the effort of her neighbor, which increases her social norm, and so forth. Note that, when a link is created between a high-productive and a low-productive agent, then the low-productive agent increases her effort, because her social norm increases while the high-productive agent decreases her effort because her social norm is reduced. As a result, the impact of adding this link on the effort of all agents in the network is ambiguous. Using these results, we discuss the *key-link policy*, whose aim is to determine the link between two agents which, once removed, reduces total crime the most. We show that, irrespective of the network structure, the planner should remove the link between the two most productive agents in the network.

Third, we provide a complete welfare analysis of the local-average model. We derive a necessary and sufficient condition for the equilibrium to be socially optimal. However, this condition is not likely to hold in most networks. Indeed, each agent exerts externalities on her neighbors, which she does not take into account when making effort. In particular, when the effort of agent i's neighbor (say, agent j) is below her own social norm, then an increase in i's effort increases the social norm of j, which has a negative impact on j's conformist utility, because j's effort is now further away from her own social norm. In this case, agent i exerts a negative externality on her neighbor j. To restore the first best, the planner taxes agents who exert negative externalities on their neighbors. If the effort of agent i's neighbor (say, agent j) is *above* her own social norm, then the reasoning is the same in reverse, so that to restore the first best, the planner *subsidizes* agents who exert *positive* externalities on their neighbors. This is very different from the policy implications of the local-aggregate model, in which agents always exert positive externalities on their neighbors so that the planner always subsidizes agents and gives higher subsidies to more central agents. Here, if central agents have higher productivity, they are more likely to exert negative externalities on their neighbors, since the latter are more likely to have effort below their own social norms. For example, in a star-shaped network, if the central agent has, on average, higher productivity than that of the peripheral agents, in the local-aggregate model, to restore the first best, the planner gives the highest subsidy to the central agent. By contrast, in the local-average model, the planner taxes the central agent and subsidizes the peripheral agents.

We also consider different extensions of our benchmark model. First, we extend our utility function so that agents have different tastes for conformity. We show that all our results are robust to this extension. Second, we consider an *anti-conformist* model in which agents benefit from deviating from the social norm of their friends. We show that if agents are not too anti-conformist, then our results hold even if some agents provide zero effort in equilibrium. However, when agents become more anti-conformist, then either no equilibrium exists or multiple equilibria prevail. We also consider a model in which agents may want to make effort above the average effort of their friends. In this model, contrary to our benchmark model in which agents either overinvest or underinvest in efforts compared to the first best, we show that they tend to mostly overinvest, because they always want to exert efforts above the social norm of their neighbors. Finally, we extend our model to directed and weighted networks and show that all our results are robust to this extension.

Next, we study the implications of our model for network formation. Specifically, we consider a two-stage model in which, in the first stage, agents form links, and in the second stage, they exert effort. We show that, in the local-aggregate model, the unique pairwise Nash equilibrium is the complete network. On the contrary, in the local-average model, the unique pairwise Nash equilibrium is the complete *homophilous* network in which agents of the same type form a complete network but never create links with agents of the other type. In other words, the local-average model provides a simple explanation of homophilous behavior, whereas the local-aggregate model fails to do so.

Finally, we discuss the differences in policy implications of the local-average and the local-aggregate models. We show that, in the former model, group-based policies are more efficient while in the latter model, it is better to implement individual-based or key-player policies.

Contributions to the literature Other researchers have studied the local-average (conformist) model in network games.[9] Patacchini and Zenou (2012) and Liu et al. (2014) characterized Nash equilibrium and showed that it exists and is unique; Blume et al. (2015) and Golub and Morris (2017) introduced imperfect information;[10] Boucher (2016) embedded the local-average model into a network formation model, while Olcina et al. (2017) embedded it into a learning model.[11] To the best of our knowledge, ours is the first study analyze the comparative statics properties of the local-average model as well as its welfare and policy implications. Ours is also the first study to examine how adding or removing a link changes the effort of all agents in the network.

One may argue that many peer-effect empirical studies cannot distinguish between the local-average and the local-aggregate model because, in the usual case, the size of the reference group is constant in the sample. For example, the neighborhood is the class, or co-workers, and the network is the same for everyone, namely, a complete graph in which all the students in a class, residents of a neighborhood, or employees of a firm are interlinked. Fortunately, because of network data availability, many recent studies have precisely described the network of agents (see, e.g., Christakis and Fowler, 2007; Bramoullé et al., 2009; Calvó-Armengol et al., 2009; Banerjee et al., 2013; for overviews, see Breza, 2016; Jackson et al., 2017) and therefore, can easily distinguish between the two models. Thus, the results of the present study can be used to derive adequate policy recommendations for each model.[12]

[9]Some studies have introduced conformity in the utility function without an explicit network analysis but the social norm is usually assumed to be exogenous. See, among others, Akerlof (1980, 1997), Kandel and Lazear (1992), Bernheim (1994), and Fershtman and Weiss (1998).

[10]See Ghiglino and Goyal (2010), Bloch and Quérou (2013), and Chen et al. (2018), who also developed theoretical network models with average effects but focusing on different issues.

[11]Olcina et al. (2017) forms part of the wide literature on learning on networks using the DeGroot model, whereby the utility function is implicitly assumed to be equivalent to the local-average model. For an overview of this literature, see Golub and Sadler (2016).

[12]For example, Carrell et al. (2013) assigned students to peer groups so that the academic performance of the least able students was maximized. The authors showed that using average peer effects to "optimally" design these groups without taking into account the network relationships between these students could backfire, since they found a negative and significant treatment effect

The rest of the paper unfolds as follows. In Section 2, we develop the local-average model and characterize the best response functions. In Section 3, we study the comparative statics properties of the model. In Section 4, we investigate the welfare properties of the local-average model. Section 5 considers different extensions of our model. In Section 6, we examine the policy implications of our results. Finally, Section 7 concludes. All proofs are in Online Appendix A. Online Appendix B provides a comparison between the local-average and the local-aggregate model. In Online Appendix C, we provide a probabilistic interpretation of our model. In Online Appendix D, we provide a simple example that shows how a mean-preserving spread of the productivity impacts own and aggregate outcome. In Online Appendix E, we provide additional results and examples on the comparative statics of the taste for conformity while in Online Appendix F, we compare equilibrium and first-best outcomes for specific networks. In Online Appendix G, we consider different extensions of our model.

2 The local-average model

2.1 Definitions and notation

Consider $n \geq 2$ individuals (or agents) who are embedded in a *network* $\mathbf{g}$. The *adjacency matrix* $\mathbf{G} = [g_{ij}]$ is an $(n \times n)$-matrix with $\{0, 1\}$ entries, which keeps track of the *direct connections* in the network. By definition, agents i and j are *directly connected* if and only if $g_{ij} = 1$; otherwise, $g_{ij} = 0$. We assume that the network is *undirected*, that is, $g_{ij} = g_{ji}$, and has *no self-loops*, that is, $g_{ii} = 0$.

Denote by $\widehat{\mathbf{G}} = [\widehat{g}_{ij}]$ the $(n \times n)$ row-normalized adjacency matrix defined by $\widehat{g}_{ij} := g_{ij}/d_i$, where d_i is individual i's *degree*, or the number of her direct neighbors, that is, $d_i := \sum_{j=1}^{n} g_{ij}$.

Each agent $i = 1, 2, \ldots, n$ is described by: (i) her *productivity* $\alpha_i \in \mathbb{R}_+$, which is an exogenous characteristic; (ii) her *effort* $x_i \in \mathbb{R}_+$, which is agent i's choice variable; and (iii) her position in the network $\mathbf{g}$, which defines her social norm. Following the standard notation, we set

$$\boldsymbol{\alpha} := (\alpha_1, \alpha_2, \ldots, \alpha_n)^T \in \mathbb{R}_+^n, \qquad \mathbf{x} := (x_1, x_2, \ldots, x_n)^T \in \mathbb{R}_+^n,$$

while the subscript $(-i)$ means dropping a vector's ith coordinate:

$$\mathbf{x}_{-i} := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)^T \in \mathbb{R}_+^{n-1}.$$

for the least able students.

Finally, agent i's *social norm*, $\overline{x}_i$, is defined as the average effort across her neighbors, namely,

$$\overline{x}_i := \sum_{j=1}^{n} \widehat{g}_{ij} x_j \tag{2}$$

In equilibrium, each agent's effort x_i is represented[13] as a convex combination of her own exogenous productivity α_i and her endogenous social norm $\overline{x}_i$. This is very much in the spirit of the linear-in-means model (1).

2.2 Preferences

Agent i's utility function has a standard linear-quadratic structure and is given by

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{1}{2} x_i^2 - \frac{\theta}{2} \left(x_i - \overline{x}_i\right)^2, \tag{3}$$

where $\alpha_i > 0$ stands for agent i's *individual productivity*, while $\theta > 0$ is the *taste for conformity*.[14]

The utility function (3) has two terms. The first term, $\alpha_i x_i - x_i^2/2$, is the utility of exerting x_i units of effort when there is *no interaction* with other individuals. The second term, $-\theta \left(x_i - \overline{x}_i\right)^2/2$, captures the *peer-group pressure* faced by agent i, who seeks to minimize her social distance from her reference group, and suffers a utility reduction equal to $\theta \left(x_i - \overline{x}_i\right)^2/2$ from failing to conform to others.[15]

For the sake of analytical convenience, we reparametrize the taste for conformity by setting

$$\lambda := \frac{\theta}{1+\theta}, \qquad 0 \leq \lambda < 1. \tag{5}$$

[13]See equation (11) below.

[14] Note the difference between (3) and the local-aggregate model (Ballester et al., 2006), where the utility of agent $i = 1, 2, \ldots, n$ is given by

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{1}{2} x_i^2 + \theta \sum_{j=1}^{n} g_{ij} x_i x_j, \tag{4}$$

that is, it is the aggregate effort of peers, $\sum_{j=1}^{n} g_{ij} x_j$, which positively affects own utility. In Appendix B, we compare the local-average and the local-aggregate model.

[15]This is the standard way in which economists have modeled conformity (see, among others, Akerlof, 1980, 1997; Kandel and Lazear, 1992; Bernheim, 1994; Fershtman and Weiss, 1998; Patacchini and Zenou, 2012; Boucher, 2016).

By plugging (5) into (3), we obtain

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{1}{2}x_i^2 - \frac{1}{2}\left(\frac{\lambda}{1-\lambda}\right)(x_i - \overline{x}_i)^2 \tag{6}$$

The two parameterizations, (3) and (6), are clearly equivalent. Indeed, as observed from (5), λ is a monotone transformation of θ.

We now point out some important properties of the utility function (6), which provides useful intuition about our main results. First, if i and j are neighbors, we have

$$\frac{\partial U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})}{\partial x_j} \gtreqless 0 \iff x_i \gtreqless \overline{x}_i. \tag{7}$$

In other words, when agent j makes effort x_j, she exerts a positive (negative) *externality* on her neighbor i if and only if the effort of i is above (below) i's social norm. This is important in the welfare section, since we observe that the equilibrium effort differs from the first-best, because agents fail to internalize externalities when choosing their effort levels. These externalities are positive or negative depending on whether the effort is above or below the social norm. This highlights the importance of having endogenous social norms.

Second, efforts are *strategic complements*. Indeed, for $\widehat{g}_{ij} > 0$,

$$\frac{\partial^2 U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})}{\partial x_i x_j} > 0, \tag{8}$$

which means that the higher is the effort of an individual's peer, the higher is the individual's marginal utility of exerting effort.

Third, the cross-effect of individual i's effort x_i and the taste for conformity λ is given by:

$$\frac{\partial^2 U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})}{\partial x_i \partial \lambda} \lesseqgtr 0 \iff x_i \gtreqless \overline{x}_i. \tag{9}$$

In other words, if $x_i > \overline{x}_i$ ($x_i < \overline{x}_i$), then, when agents become more conformist, an increase in x_i increases (reduces) the gap between x_i and $\overline{x}_i$, which leads to a decrease (increase) in the utility level. In other words, an increase in λ decreases (increases) the marginal utility of exerting effort for individual i if $x_i > \overline{x}_i$ ($x_i < \overline{x}_i$). We refer to this assumption when discussing the comparative statics of λ.

Finally, the cross-effects of effort and productivity are positive, as for any $i, j, k = 1, 2, \ldots, n$ we have

$$\frac{\partial^2 U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})}{\partial x_j \partial \alpha_k} \geq 0. \tag{10}$$

Hence, productivities $\boldsymbol{\alpha}$ and efforts $\mathbf{x}$ satisfy the standard Milgrom–Shannon conditions, which guarantee monotone comparative statics in supermodular games (see Proposition 2 below). However, this is not the case for the comparative statics in terms of the taste for conformity λ (see (9)).

To summarize, the utility function (3)—equivalently, (6)—is the standard way economists have modeled conformity. However, the social norm $\overline{x}_i$ is usually assumed to be exogenous (see, e.g., Akerlof, 1980, 1997), which makes the problem less interesting, because it abstracts from general equilibrium effects (Dutta et al., 2019). Here, we endogenize the social norm by making it dependent on the network structure. In that case, agents create externalities for each other through the social norm that they do not take into account when exerting their effort. This leads to new policy implications that we explore in Sections 4 and 6.

2.3 Nash equilibrium

Each individual i chooses x_i to maximize (6) taking the network structure $\mathbf{g}$ and the effort choices $\mathbf{x}_{-i}$ of other agents as given. By computing agent i's first-order condition (FOC) with respect to x_i, we obtain the following best-reply function for each i:

$$x_i = (1-\lambda)\alpha_i + \lambda \overline{x}_i. \tag{11}$$

After some normalizations, it should be clear that (11) is equivalent to the standard linear-in means model (1) in which individual effort is a function of individual observable characteristics α_i, which can also depend on the characteristics of neighbors, and on the endogenous peer effect $\overline{x}_i$.

Combining (11) with the definition (2) of agent i's social norm, we find that the vector $\mathbf{x}^* := (x_1^*, x_2^*..., x_n^*)^T$ of equilibrium efforts must be a solution to

$$\mathbf{x} = (1-\lambda)\boldsymbol{\alpha} + \lambda \widehat{\mathbf{G}}\mathbf{x}, \tag{12}$$

where $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_n)^T$ is the productivity vector.[16]

[16] Observe that the linear-in-means model (1) is closely related to the *spatial-autoregressive (SAR) model* in the spatial econometrics literature (LeSage and Pace, 2009) and is usually written in matrix form as

$$\mathbf{x} = \boldsymbol{\beta} + \lambda \widehat{\mathbf{G}}\mathbf{x} + \boldsymbol{\epsilon},$$

where, as in our model, $\widehat{\mathbf{G}}$ is a row-normalized matrix that captures the distance or proximity in the geographical space (or any other space, e.g., the social space) between different agents or entities, such as geographical areas. In this literature, the main reason for the matrix $\widehat{\mathbf{G}}$ to be row-normalized is to obtain an intuitive interpretation of λ as the weighted average impact of neighbors but also to avoid explosive spatial multipliers implied by λ (by analogy to time-series econometrics,

Proposition 1 (Equilibrium efforts, norms, and utilities)

(i) There exists a unique interior Nash equilibrium $\mathbf{x}^*$*, which is given by*

$$\mathbf{x}^* = \widehat{\mathbf{M}}\boldsymbol{\alpha}, \tag{13}$$

where $\widehat{\mathbf{M}} = [\widehat{m}_{ij}]$ *is an* $(n \times n)$*-matrix of marginal effects defined as follows:*[17]

$$\widehat{\mathbf{M}} := (1-\lambda)\left(\mathbf{I} - \lambda\widehat{\mathbf{G}}\right)^{-1} = \; (1-\lambda)\sum_{k=0}^{\infty} \lambda^k \widehat{\mathbf{G}}^k. \tag{14}$$

(ii) The equilibrium social norms $\overline{\mathbf{x}}^*$ *are given by*

$$\overline{\mathbf{x}}^* = \widehat{\mathbf{G}}\widehat{\mathbf{M}}\boldsymbol{\alpha} = (1-\lambda)\sum_{k=0}^{\infty} \lambda^k \, \widehat{\mathbf{G}}^{k+1}\boldsymbol{\alpha}. \tag{15}$$

(iii) For each $i = 1, 2, \ldots, n$*, agent* i*'s equilibrium utility level is given by*

$$U_i^*(\boldsymbol{\alpha}, \lambda, \mathbf{g}) = \frac{1}{2}\left[\alpha_i^2 - \frac{1}{\lambda}\left(\alpha_i - \sum_{j=1}^{n} \widehat{m}_{ij}\alpha_j\right)^2\right]. \tag{16}$$

Several comments are in order. First, taking a closer look at the structure of the marginal effect $\widehat{m}_{ij}$ of agent i's productivity on agent j's effort, we obtain

$$\widehat{m}_{ij} = \sum_{k=0}^{\infty} \underbrace{(1-\lambda)\lambda^k}_{\substack{\text{geometric} \\ \text{distribution}}} \; \widehat{g}_{ij}^{[k]}. \tag{17}$$

As seen from (17), $\widehat{m}_{ij}$ is decomposed into a series whose kth term is proportional to $\widehat{g}_{ij}^{[k]}$, that is, the normalized number of paths from i to j of length k in the

in which the autoregression parameter λ is expected to be strictly less than 1 in modulus; see Hamilton, 1994). Equation (12) is clearly equivalent to the spatial-autoregressive model and it gives a microfoundation of the SAR model via the utility function (3) or (6).

[17]Because $\widehat{\mathbf{G}}$ is row-normalized and $0 \leq \lambda < 1$, the matrix $\widehat{\mathbf{M}}$ of marginal effects is well defined and can be represented by the Neumann series. This follows from Corollary 5.6.16 in Horn and Johnson (1985, Ch. 5, p. 301), in which the suitable matrix norm is the maximum row sum norm.

social network. Surprisingly, the coefficients of the series are given by the standard geometric distribution with the odds ratio equal to $\theta \equiv \lambda/(1-\lambda)$. Therefore, although the game under study is fully deterministic, one may inquire whether the marginal effects $\widehat{m}_{ij}$ have some probabilistic origin. In Online Appendix C, we demonstrate that the local average model is observationally equivalent to an average outcome of a naive social learning model.[18]

Second, there is *no need to impose any conditions on* $\theta \equiv \lambda/(1-\lambda)$ (except that $\theta > 0$) to guarantee the existence of a unique and interior Nash equilibrium. This is not the case in the local aggregate model.[19]

Third, it is readily verified that, if agents are *ex ante homogeneous*, that is, if $\alpha_i = \alpha_j$ for any $i, j = 1, 2, \ldots, n$, then, regardless of the network structure, *the equilibrium effort levels are the same across agents*: $x_i^* = x_j^*$ for any $i, j = 1, 2, \ldots, n$. This result displays another significant difference with the local aggregate model, in which the outcome is represented by the Katz–Bonacich centralities of the agents. Here, the impact of the network structure on equilibrium is mediated by the *correlation* between the productivity distribution $\boldsymbol{\alpha}$ and the degree distribution of the network $\mathbf{g}$. We return to this property in Sections 3 and 4.

Fourth, instead of assuming (3) or (6), the following utility function can be assumed:

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{1}{2} x_i^2 - \frac{\theta}{2} \sum_{j=1}^{n} \widehat{g}_{ij} (x_i - x_j)^2, \tag{18}$$

and exactly the same first-order condition (11) can still be obtained and thus, the same equilibrium effort x_i^*. The interpretation of the utility function (18) is still in terms of conformism but now, each individual pays some cost from deviating from the action of *each* of her neighbors instead of the *average* action of her neighbors.

Even if the equilibrium effort is the same and equals x_i^*, the equilibrium utility is

[18] A similar result was obtained by Golub and Morris (2017).

[19] Indeed, in the local-aggregate model for which the utility function is given by (4), one needs a condition on θ (i.e., $\theta < 1/\mu(\mathbf{G}$, where $\mu(\mathbf{G})$ is the largest eigenvalue of $\mathbf{G}$), to prove the uniqueness of equilibrium. In the local-average model, one does not need such a condition, because the matrix to be inverted is $\left(\mathbf{I} - \lambda \widehat{\mathbf{G}}\right)$, where $\widehat{\mathbf{G}}$ is the row-normalized matrix of $\mathbf{G}$. The largest eigenvalue of $\widehat{\mathbf{G}}$ equals one and thus, the condition for invertibility of $\left(\mathbf{I} - \lambda \widehat{\mathbf{G}}\right)$ is $\lambda := \theta/(1+\theta) < 1$, which is always true.

different.[20] As a result, the equilibrium effort x_i^* and its comparative statics results are the same but the welfare analysis and its comparative statics may differ, because the equilibrium utilities and thus, welfare are different.[21]

Finally, in part (*iii*) of Proposition 1, we calculate the equilibrium utility level of each agent in the network as a function of the parameters of the model. An important aspect of this model is whether individual i's effort is above or below her own social norm. The following result clarifies this relationship.

Lemma 1 *For each $i = 1, 2, \ldots, n$, we have*

$$x_i^* \gtreqqless \overline{x}_i^* \iff \alpha_i \gtreqqless \sum_{j=1, j\neq i}^{n} \frac{\widehat{m}_{ij}}{(1-\widehat{m}_{ii})}\alpha_j. \tag{19}$$

This lemma shows that agent i's own effort is above (below) her social norm if and only if her productivity is higher (smaller) than the weighted average of the other productivities in the network. For example, in a star network, if the central agent is more productive than the others, then her effort is always above the social norm of her neighbors (the peripheral agents), who, in turn, exert effort below that of their social norm, since the latter is the effort of the central agent. This is a useful insight that helps us to understand the main results of Sections 3 and 4.

[20]It is easily verified that, in our model, the equilibrium utility is given by

$$U_i(x_i^*, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i^* - \frac{1}{2}(x_i^*)^2 - \frac{\theta}{2}(x_i^*)^2 + \theta x_i^* \bar{x}_i^* - \frac{\theta}{2}\left(\sum_{j=1}^{n} \widehat{g}_{ij} x_j^*\right)^2$$

while, in this new model with preferences given by (18), we have:

$$U_i(x_i^*, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i^* - \frac{1}{2}(x_i^*)^2 - \frac{\theta}{2}(x_i^*)^2 + \theta x_i^* \bar{x}_i^* - \frac{\theta}{2}\sum_{j=1}^{n} \widehat{g}_{ij} (x_j^*)^2$$

The only difference between these two utility functions is the last term which is clearly different, since $\left(\sum_{j=1}^{n} \widehat{g}_{ij} x_j^*\right)^2 = (\bar{x}_i^*)^2 \neq \sum_{j=1}^{n} \widehat{g}_{ij}(x_j^*)^2$.

[21]As noted by Boucher and Fortin (2016), another utility function could have generated the same first-order conditions (11) and thus, the same equilibrium effort x_i^*. It is given by

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{(1+\theta)}{2}\, x_i^2 + \theta x_i \sum_{j=1}^{n} \widehat{g}_{ij} x_j.$$

However, in this case, the properties of the model are very different, since it is no longer a conformist model.

2.4 Linear-in-means model and heterogeneity: An example

In the Introduction, we discuss how peer effects are estimated in the literature using the linear-in-means model (see (1)), which captures an *average* effect. In reality, the same average effect can have a very different impact on outcomes, depending on other moments of the distribution, in particular, the variance.[22] Contrary to the linear-in-means model, the local-average model can address this issue, since it encompasses a network approach whereby the group each individual belongs to is determined by her direct neighbors. In that case, the whole distribution matters in evaluating the impact of peers on outcomes.

To illustrate this, in Online Appendix D, we provide a simple example that shows how a mean-preserving spread of the productivity impacts own and aggregate outcome. This example shows that estimating a linear-in-means model may be misleading, because it focuses only on the average effect and does not take into account other characteristics of the distribution of efforts in the population. In this example, we show that the local-average model can have a very different prediction than the linear-in-means model, depending on the value of λ, the taste for conformity, and the value of t. Indeed, with exactly the same average characteristic (here, productivity) in the group (here, network), the individual effort level may vary a lot. In this example, these changes are driven by t, which is proportional to the standard deviation of the productivity distribution. As a result, when studying the impact of the social norm on individual effort, one should not only take into account the average social norm of the reference group but also its variance.

3 Comparative statics

We aim to understand the properties of our model by performing some comparative statics exercises of the Nash equilibrium with respect to the key parameters of the model: (*i*) the productivity vector $\boldsymbol{\alpha}$; (*ii*) the taste for conformity λ; and (*iii*) the density/sparsity of social network $\mathbf{g}$.

[22]For example, in a classroom of 30 students, the impact of an average test score of 50/100 is very different if all students have a test score of around 50/100 (i.e., low variance with a very homogeneous distribution of test scores) than when some students have very high test scores and others have very low test scores (i.e., high variance with a very heterogeneous distribution of test scores).

3.1 Effect of productivity

Let us start with the productivity $\boldsymbol{\alpha}$ of all agents. We have the following result:

Proposition 2 (Comparative statics for productivity)

(*i*) *For all* $i, j = 1, 2, \ldots, n$, *the marginal effects of a change in individual* i*'s productivity* α_i *on individual* j*'s equilibrium effort* x_j^* *and individual* j*'s social norm* $\overline{x}_j^*$ *are positive and do not exceed* 1*:*

$$0 < \frac{\partial x_j^*}{\partial \alpha_i} < 1, \qquad 0 < \frac{\partial \overline{x}_j^*}{\partial \alpha_i} < 1.$$

(*ii*) *The equilibrium utility of each individual* $i = 1, 2, \ldots, n$ *is increasing with her own productivity*:

$$\frac{\partial U_i^*(\boldsymbol{\alpha}, \lambda, \mathbf{g})}{\partial \alpha_i} > 0.$$

(*iii*) *For any* $j \neq i$, *agent* i*'s equilibrium utility* $U_i^*(\boldsymbol{\alpha}, \lambda, \mathbf{g})$ *increases (decreases) in response to a small change in* α_j, *if and only if agent* i*'s equilibrium effort* x_i^* *is above (below) her equilibrium social norm* $\overline{x}_i^*$*; that is,* $\operatorname{sign}\left[\frac{\partial U_i^*}{\partial \alpha_j}\right] = \operatorname{sign}\left(x_i^* - \overline{x}_i^*\right)$, *or equivalently, using Lemma 1,*

$$\frac{\partial U_i^*}{\partial \alpha_j} \gtreqless 0 \iff \alpha_i \gtreqless \sum_{l=1, l\neq i}^{n} \frac{\widehat{m}_{il}}{(1-\widehat{m}_{ii})}\alpha_l$$

The first result is straightforward because, as implied by (13), each x_i^* is a convex combination of productivity and social norms. The second result, although intuitive, is relatively difficult to show. Indeed, when own productivity α_i increases, own effort x_i^* increases, which raises U_i^*, the equilibrium utility of i, but the social norm $\overline{x}_i^*$ also increases, which can increase or decrease U_i^* depending on whether x_i^* is higher or lower than $\overline{x}_i^*$. We show in the proof that the first direct effect is stronger than the second indirect effect, so that an increase in α_i always increases U_i^*. When we analyze the effect of α_j on U_i^* for $j \neq i$, we find a similar result, that is, the impact depends on whether x_i^* is above or below $\overline{x}_i^*$.

3.2 Effect of conformity

We now look at the impact of taste for conformity λ on individual and social outcomes.

3.2.1 Pure individualism versus total conformism

To obtain some intuition, we begin by contrasting two extreme cases: *pure individualism* ($\lambda = 0$), where i's utility depends only on own productivity α_i; and *total conformism* ($\lambda \to 1$), where i's utility depends only on others' behavior. To obtain these results, we use the observational equivalence between our models and that of the Markov chain developed in Online Appendix C to compare the outcomes generated by perfect individualism ($\lambda = 0$) and total conformism ($\lambda = 1$).

It is straightforward to observe that, under pure individualism ($\lambda = 0$), we have $x_i^* = \alpha_i$. In this case, norms play no role, and there are incentives for an individual to exert neither higher nor lower effort than her intrinsically desirable level, α_i. However, the outcome when $\lambda \to 1$ is less obvious.

Proposition 3 (Totally conformist agents) *For any network structure, individual efforts in a totally conformist society are given by*

$$\lim_{\lambda \to 1} x_i^*(\lambda) = \boldsymbol{\pi}\boldsymbol{\alpha} = \sum_{j=1}^{n} \pi_j \alpha_j, \qquad \text{for all} \qquad i = 1, \ldots, n, \tag{20}$$

where $\boldsymbol{\pi} \equiv (\pi_1, \pi_2, \ldots, \pi_n)$ *are normalized degrees of agents:*

$$\pi_i := \frac{d_i}{\sum_{j=1}^{n} d_j}, \quad \text{for all } i = 1, 2, \ldots n, \tag{21}$$

Proposition 3 shows that, for any network structure, when agents are perfectly conformist, the equilibrium effort depends only on the weighted productivity in the network, where the weights depend on the network structure. This implies, in particular, that π_j is the probability that a *perfectly conformist* individual i exerts a level α_j of effort. This means that, when $\lambda \to 1$, the *effort of all agents in the network is the same* and that the level of these efforts depends on the network structure captured by $\boldsymbol{\pi}$ and on the productivity distribution captured by $\boldsymbol{\alpha}$. Thus, the probabilistic interpretation of the model helps us to understand the totally conformist society, which is otherwise difficult to characterize.[23]

We are now equipped to compare the purely individualist society ($\lambda \to 0$) and the totally conformist society ($\lambda \to 1$).

Proposition 4 (Individualist versus conformist society)

[23] A similar result in terms of conformity limits was shown by Golub and Morris (2017), but in the context of imperfect information.

(*i*) *Individual effort:*

$$\lim_{\lambda\to 0} x_i^*(\lambda) \gtreqless \lim_{\lambda\to 1} x_i^*(\lambda) \iff \alpha_i \gtreqless \sum_{j=1}^{n} \pi_j \alpha_j$$

(*ii*) *Aggregate effort:*

$$\lim_{\lambda\to 0} \sum_i x_i^*(\lambda) \gtreqless \lim_{\lambda\to 1} \sum_i x_i^*(\lambda) \iff \sum_{j=1}^{n} \alpha_j \gtreqless n \sum_{j=1}^{n} \pi_j \alpha_j$$

Part (*i*) of Proposition 4 shows that the effort exerted by each agent i can be higher or lower in a pure individualist society than in a completely conformist one if the productivity of i is above or below the weighted average productivity in the network. This result depends on both own productivity and the network structure. Part (*ii*) of Proposition 4 shows that conformity is not necessarily good for aggregate effort. However, when α_i and π_i are positively (negatively) correlated, that is, agents with higher productivity have (less) more central positions in the network,[24] then perfect conformity increases aggregate effort.

How do individual and aggregate efforts change when the taste for conformity varies? To answer this question, we study the comparative statics with respect to the conformity parameter λ.

3.2.2 The impact of the taste for conformity on outcomes

Let us totally differentiate (11) with respect to λ. We obtain

$$dx_i^* = \underbrace{-\alpha_i \, d\lambda}_{\substack{\text{productivity}\\ \text{effect}}} + \underbrace{\overline{x}_i^* \, d\lambda}_{\substack{\text{direct}\\ \text{norm effect}}} + \underbrace{\lambda \left(\partial \overline{x}_i^* / \partial \lambda\right) d\lambda}_{\substack{\text{indirect}\\ \text{norm effect}}} \tag{22}$$

Indeed, when λ increases, the individual effort of individual i, x_i^*, is affected in three different ways. First, there is a negative *productivity effect*, according to which, when

[24] Indeed, it is straightforward to show that:

$$\sum_{j=1}^{n} \pi_j \alpha_j \gtreqless \frac{1}{n} \sum_{j=1}^{n} \alpha_j \iff \text{Corr}(\boldsymbol{\pi}, \boldsymbol{\alpha}) \gtreqless 0,$$

where $\text{Corr}(\boldsymbol{\pi}, \boldsymbol{\alpha})$ is the correlation between $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$. If $\text{Corr}(\boldsymbol{\pi}, \boldsymbol{\alpha}) > 0$ (< 0), then more productive agents are also more (less) central (in terms of degree centrality) in the network.

conformity increases, the impact of own productivity on effort decreases. Second, there is a positive *direct social-norm effect*, indicating that, when λ increases, the impact of the social norm on own effort increases. These are straightforward direct effects due to the fact that, when λ increases, agents pay more attention to their neighbors than to themselves. There is a third, more subtle effect, the *indirect social-norm effect*, which can be positive or negative. This effect shows that, when λ increases, the social norm itself changes as i changes her effort and her peers become more conformist. The effect is ambiguous as i's friends may increase or decrease their effort following an increase in λ. As a result, the total effect of λ on x_i^* is ambiguous. To understand this better, using (11), (22) can be written as

$$dx_i^* = -\left(x_i^* - \overline{x}_i^*\right) \frac{d\lambda}{1-\lambda} + \lambda \frac{\partial \overline{x}_i^*}{\partial \lambda} d\lambda$$

We now see that the total impact of a change of λ crucially depends on whether the individual effort of i is above or below her own social norm. As observed from (9), this is because the effect of λ on the marginal utility of effort is ambiguous and depends on the gap, $x_i - \overline{x}_i$, between the individual's effort and her social norm. In particular, when λ increases, agents become more conformist, and the gap between x_i and $\overline{x}_i$ matters more.

Recall that $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)$ is the normalized degree distribution of the network $\mathbf{g}$ (see (21)). We obtain the following result.

Proposition 5 (Non-monotonicity of individual efforts in conformism)

(*i*) *For any $\lambda \in (0,1)$, if $\partial x_i^*/\partial \lambda > 0$ for some i, then it has to be that $\partial x_j^*/\partial \lambda < 0$ for some $j \neq i$.*

(*ii*) *When λ is small, we have*

$$\frac{\partial x_i^*}{\partial \lambda} \gtrless 0 \iff \alpha_i \lesseqgtr \sum_{j=1}^{n} \widehat{g}_{ij} \alpha_j \tag{23}$$

(*iii*) *Assume that the following conditions hold:*

$$\sum_{j=1}^{n} \pi_j \alpha_j \leq \alpha_i < \sum_{j=1}^{n} \widehat{g}_{ij} \alpha_j . \tag{24}$$

Then, agent i's individual effort $x_i^*(\lambda)$ has an interior global maximum in λ.

(iv) Assume that the following conditions hold:

$$\sum_{j=1}^{n} \pi_j \alpha_j \geq \alpha_i > \sum_{j=1}^{n} \widehat{g}_{ij} \alpha_j. \tag{25}$$

Then, agent i's individual effort $x_i^*(\lambda)$ has an interior global minimum in λ.

Part (*i*) of Proposition 5 provides an expression of the impact of conformity on individual i's effort. We show that it crucially depends on whether both individual i and all other agents in the network (since all agents are path-connected to each other) make efforts above or below the social norm of their friends. In particular, if we order agents by their productivity in descending order, so that $\alpha_{\max} := \alpha_1$ and $\alpha_{\min} := \alpha_n$ are the highest and lowest values of productivity among the n agents in the network, respectively, then, by Lemma 1, it has to be that $x_1^* > \overline{x}_1^*$ and $x_n^* < \overline{x}_n^*$. As a result, because some individuals exert effort above the norm and some below the norm, the total impact of λ on an individual is ambiguous, and has to increase for some individuals and decrease for others. Equation (23) shows that, for small λ, the sign of this derivative depends only on whether i's productivity is above or below that of her peers.[25]

Observe that this comparative statics result is very different to that obtained in the local-aggregate model in which an increase in λ or θ (social multiplier or social interaction effect in the local-aggregate model; see (4)) always leads to an increase in effort x_i^*. This is important for policy purposes, because, as noted by Boucher and Fortin (2016), if there is a positive policy shock on λ, and we observe that individual effort either decreases or the effect is non-monotonic, then we know that the underlying utility function is defined by the local-average model (see (3) or (6)) and not by the local-aggregate model. To know which utility function each agent has when choosing her effort is important for policy implications, as discussed in Section 6 below.

Parts (*ii*) and (*iii*) of Proposition 5 provide sufficient (but not necessary) conditions for x_i^* to vary *non-monotonically* with λ.[26] Based on these conditions, which depend only on the productivity parameters and the structure of the network, α_i cannot be neither too high nor too low for the relationship between x_i^* and λ to be non-monotonic. Clearly, if λ_i is very high (low), which implies that x_i^* is very likely to be above (below) $\overline{x}_i^*$, then $\frac{\partial x_i^*}{\partial \lambda}$ is negative (positive). Conditions (24) and (25)

[25] Proposition G4 in Online Appendix G.2 generalizes Proposition 5 when the taste for conformity is individual specific and equal to λ_i for each agent i.

[26] We give sharper conditions for some specific types of network structures in Section E.3 in Online Appendix E.

also guarantee a global interior maximum or minimum in λ. In particular, if α_i is above (below) the productivity in the network, there is a global interior maximum (minimum), which means that an increase in λ first has a positive (negative) impact on x_i^* and then a negative (positive) one.

In fact, the non-monotonicity expressed in parts (*ii*) and (*iii*) of Proposition 5 can be complex and not necessarily U shaped or bell shaped. In Figure 1, we provide an example for a chain network with 13 nodes in which increasing λ yields an S shape. In this chain network, node 0 is in the middle, nodes $1, 2, 3, 4, 5$, and 6 are on the right side of node 0, while nodes $-1, -2, -3, -4, -5$, and -6 are on the left side of node 0.[27]

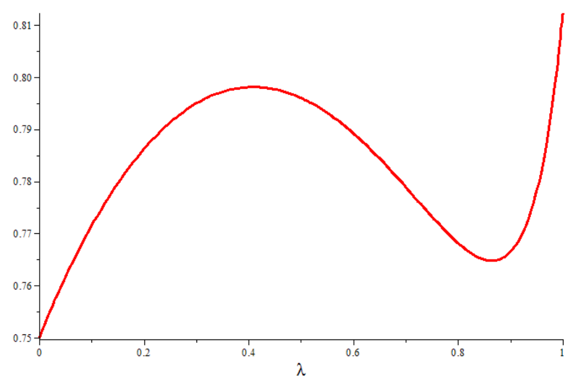

Figure 1: Non-monotonic effect of λ on individual effort for a chain network with $n = 13$

In Proposition 5, we show that the impact of λ on *individual effort* is very complex and difficult to sign. In Corollary E.1 in Online Appendix E, we show that the same non-monotonicity results hold for the *aggregate effort*, which is an important aspect of this model.[28] Also, in Proposition E1 in Online Appendix E, we demonstrate that, in regular networks, the aggregate effort does not vary with λ. This is because, in a regular network, there is perfect compensation between the positive impact of λ on low-productive agents and the negative impact of λ on high-productive agents. As a result, neither the average nor aggregate effort in a regular network are affected by a change in λ. In Section E.4 in Online Appendix E, we illustrate this result by means

[27]The values of productivity are assumed to be: $\alpha_0 = 0.75$, $\alpha_1 = 1 = \alpha_{-1}$, $\alpha_2 = 0.5 = \alpha_{-2}$, $\alpha_3 = \alpha_{-3} = 0.25$, $\alpha_4 = 0.5 = \alpha_{-4} =$, $\alpha_5 = 2\alpha_{-5}$, and $\alpha_6 = 0.5 = \alpha_{-6}$.

[28]For example, in crime, we would be interested in analyzing how conformity affects individual crime effort but also the total crime level in the network.

of a circular network. When we rewire the links in this network without changing the network topology, we show that the convergence of agents' efforts to the average effort can be faster or slower than in the original network depending on the rewiring.

To summarize, in this section, we show that the impact of the taste for conformity λ on i's effort depends on the productivity of each individual and the network topology, which determines the links between all agents and, thus, the peer pressure (via the social norm) that neighbors exert on own effort. Therefore, the effect of a higher taste for conformity on own effort is complex and determined by whether the individual is an "underdog" or someone who has high productivity. If we consider crime, this determination is important, since it shows how delinquents influence each other and how an individual's crime effort is affected by the degree of conformism in the peer group she belongs to.

3.3 Do agents exert more effort in denser networks?

We now consider the consequences of a change in the network structure by asking the following question: how does adding a new link to the existing network affect the equilibrium efforts? In the local-aggregate model, the answer is straightforward: because of strategic complementarities, regardless of the productivities αs, all agents always exert more effort in denser networks. However, as the following proposition shows, this is not always true in the local-average model.

Proposition 6 *Assume that agents i and j are not connected to each other ($g_{ij} = 0$). Then, adding a link between i and j leads to:*

(i) an increase in everyone's effort, if the following two conditions hold simultaneously:

$$\alpha_i > \frac{\sum_{l \neq i} \left(\widehat{m}_{il} - \lambda \widehat{m}_{jl}\right) \alpha_l}{1 - \lambda - \widehat{m}_{ii} + \lambda \widehat{m}_{jj}}, \tag{26}$$

$$\alpha_j > \frac{\sum_{l \neq j} \left(\widehat{m}_{jl} - \lambda \widehat{m}_{il}\right) \alpha_l}{1 - \lambda - \widehat{m}_{jj} + \lambda \widehat{m}_{ii}}; \text{and} \tag{27}$$

(ii) a reduction of everyone's effort, if the inequalities are opposite in (26) *and* (27)*.*

Otherwise, there is an ambiguous outcome.

This proposition shows that, in any network, adding a link between two agents who have high (low) productivities not only increases (decreases) the effort of these two agents but also increases (reduces) the effort of all the other agents in the network. Indeed, if we connect agent i to a high-productivity agent j, then i's norm increases and the best response for i is to increase her effort (see (11)). This implies that the norm of i's neighbors increases, which, in turn, increases their effort, and so forth. Similarly, if we connect j to a high-productivity agent i, then j's norm increases and the best response for j is to increase her effort. We have again the same snow-ball effect. The same reasoning applies in the opposite direction if we connect two low-productive agents. Indeed, if agent i connects to low-productivity agent j, then i's norm decreases, which reduces i's effort. This, in turn, decreases the norm of i's neighbors, which reduces their effort, and so forth.

To illustrate this result, consider a star network with three agents in which agent 1 is in the center. The row-normalized adjacency matrix is then given by

$$\widehat{\mathbf{G}}^S = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Let us first assume that $\alpha_1 = 2$, $\alpha_2 = 1$ and $\alpha_3 = 0.5$, so that the star is more productive than the peripheral agents are. It is easily verified that

$$\mathbf{x}^{*S} = \frac{1}{4(1+\lambda)} \begin{pmatrix} 3\lambda + 8 \\ -\lambda^2 + 8\lambda + 4 \\ \lambda^2 + 8\lambda + 2 \end{pmatrix}, \quad \overline{\mathbf{x}}^{*S} = \frac{1}{4(1+\lambda)} \begin{pmatrix} 8\lambda + 3 \\ 3\lambda + 8 \\ 3\lambda + 8 \end{pmatrix}.$$

According to part (ii) of Proposition 6, adding the link $2 - -3$ between the two less productive agents should decrease the efforts of all agents in the network. Let us verify this. By adding the link between agents 2 and 3, the network becomes complete and the row-normalized adjacency matrix is now given by

$$\widehat{\mathbf{G}}^C = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{pmatrix}$$

In that case, we obtain

$$\mathbf{x}^{*C} = \frac{1}{2(2+\lambda)} \begin{pmatrix} 8 - \lambda \\ 4 + 3\lambda \\ 2 + 5\lambda \end{pmatrix}, \quad \overline{\mathbf{x}}^{*C} = \frac{1}{2(2+\lambda)} \begin{pmatrix} 3 + 4\lambda \\ 5 + 2\lambda \\ 6 + \lambda \end{pmatrix}.$$

It is easily verified that $x_i^{*S} > x_i^{*C}$, for all $i = 1, 2, 3$, so that adding the link $2 - 3$, indeed, *decreases* the effort of all agents in the network. Consider first agent 2. By adding the link 2−3, her social norm decreases, that is, $\overline{x}_2^{*S} > \overline{x}_2^{*C}$, since, before adding the link 2−−3, the social norm of agent 2 was equal to the effort of agent 1, a very productive agent, while, after adding the link 2−−3, it becomes the average of the efforts of 1 and that of 3, a low-productive agent. Since agent 2's norm decreases, her best response is to decrease her effort. The same reasoning applies for agent 3, whose norm changes from the effort of agent 1 to the average effort of agents 2 and 3. Thus, agent 3's norm decreases and her best response is to decrease her effort. Since both agents 2 and 3 reduce their effort, the social norm of agent 1, which is the average effort of agents 2 and 3, decreases and her best response is to decrease her effort. As a result, by adding the link 2−−3, all agents reduce their effort.

Assume now that $\alpha_1 = 0.5$, $\alpha_2 = 1$, and $\alpha_3 = 2$ so that the peripheral agents are now the most productive ones in the network. Then, it is easily verified that adding the link 2−−3 *increases* the effort of all agents in the network, as predicted by part (i) of Proposition 6. This is because, when the link 2−−3 is added, the social norm of agent 2 increases, as it changes from being equal to the effort of agent 1, a low-productive agent, to the average effort of agents 1 and 3, where 3 is a high-productive agent. Her best response is to increase her effort. The same applies to agent 3. Since both agents 2 and 3 increase their effort, agent 1 also increases her effort, because her social norm increases.

Finally, if we assume that $\alpha_1 = 1$, $\alpha_2 = 0.5$, and $\alpha_3 = 2$, then adding the link 2−3 has no clear monotonic effect on the effort of all agents in the network. Indeed, on the one hand, it increases the social norm of agent 2, who increases her effort, but reduces the social norm of agent 3, who decreases her effort. This implies that the effect on the social norm of agent 1 (and her effort), which is the average effort of agents 2 and 3, would be ambiguous.

Observe that the results obtained in Proposition 6 can easily be extended to removing links.

Remark 1 *Assume $g_{ij} = 1$. If both* (26) *and* (27) *hold, then removing the link between agents i and j decreases the effort of all agents in the network. If the inequalities are opposite in* (26) *and* (27)*, then removing the link between agents i and j increases the effort of all agents in the network. Otherwise, the effect of removing the link i-j is ambiguous.*

This is an important result that has interesting policy implications. Consider crime. The usual objective of the planner is to reduce total crime, which, here, amounts to reducing aggregate effort. Thus, Remark 1 helps us answer the following question: if the planner wants to reduce total crime, which *link* should she remove from the network? This is referred to as the *key-link* policy. As Remark 1 shows, the planner needs to remove the link between the two most productive agents in the network and this is independent of the network structure. Ballester et al. (2010) determined the key link in the local-aggregate model and showed that it strongly depends on the network structure, in particular, the Katz–Bonacich centrality of the two agents involved in the key link. The main advantage of our result in Remark 1 is that the planner does not need to know the network but only the crime productivity of all agents in the network, which can be determined in the data by their crime records.

What does a key-link policy mean in the real-word? A link removal would lead to a disruption of the communication between two criminals. For example, when a police officer keeps watch over a street, she disrupts the possible communication between criminals from the same neighborhood. Another example of a key-link policy is to move a delinquent teenager to another residential location where there are less delinquents.[29] By doing so, this delinquent stops her activities and communication with other delinquents in the older residential area.[30]

4 Welfare and first best

We now analyze socially optimal outcomes. For that, let us first calculate the first-best outcome of this economy and then determine the taxes/subsidies that can restore the first best.

4.1 First best

Define the social welfare $\mathcal{W}$ as

$$\mathcal{W} := \sum_{i=1,2,\ldots,n} U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}). \tag{28}$$

The following proposition characterizes the first best and establishes a necessary and sufficient condition for the Nash equilibrium in efforts to be socially optimal.

Proposition 7 (First best)

[29]See, for example, Ludwig et al. (2001) and Kling et al. (2005, 2007), who study the moving to opportunity experiment that relocates families from high- to low-poverty neighborhoods. The authors found that this policy reduces juvenile arrests by 30 to 50% of the arrest rate for control groups.

[30]For a general discussion of removing links and disrupting the network in criminal activities, see Lindquist and Zenou (2019).

(i) *For each $i = 1, 2, \ldots, n$, the first-best effort $\mathbf{x}^O$ is a solution to*

$$x_i = (1 - \lambda)\,\alpha_i + \lambda\,\overline{x}_i + \lambda\sum_{j=1}^{n}\widehat{g}_{ji}\,(x_j - \overline{x}_j)\,, \tag{29}$$

or, in matrix form,

$$\mathbf{x} = (1 - \lambda)\,\boldsymbol{\alpha} + \lambda\widehat{\mathbf{G}}\mathbf{x} + \lambda\widehat{\mathbf{G}}^T\left(\mathbf{I} - \widehat{\mathbf{G}}\right)\mathbf{x}. \tag{30}$$

(ii) *For the Nash equilibrium to be the first best ($\mathbf{x}^* = \mathbf{x}^O$), it is necessary and sufficient that the vector $\boldsymbol{\alpha}$ of productivity satisfies the following system of linear constraints:*

$$\widehat{\mathbf{G}}^T(\mathbf{I} - \widehat{\mathbf{G}})\widehat{\mathbf{M}}\boldsymbol{\alpha} = \mathbf{0}. \tag{31}$$

(iii) *Moreover, for any network,*

$$\sum_{i=1}^{n} x_i^O = \sum_{i=1}^{n}\alpha_i. \tag{32}$$

Part (i) of Proposition 7 clearly shows the difference in effort between the Nash equilibrium (see (11)) and the first best (see (29)). In particular, compared to the Nash equilibrium, the first best has an extra term, $\lambda\sum_{j=1}^{n}\widehat{g}_{ji}\,(x_j - \overline{x}_j)$, which could be positive or negative. In fact, this extra term is the result of the following derivation: $\sum_{j\neq i}\frac{\partial U_j}{\partial \overline{x}_j}\frac{\partial \overline{x}_j}{\partial x_i}$ (see the proof of Proposition 7), where $\frac{\partial \overline{x}_j}{\partial x_i} > 0$, that is, an increase in i's effort increases the average effort of j's friends if i and j are friends, and $\frac{\partial U_j}{\partial \overline{x}_j} = (\frac{\lambda}{1-\lambda})\,(x_j - \overline{x}_j) \gtreqless 0$. This last result implies that, if $x_j > \overline{x}_j$ ($x_j < \overline{x}_j$), then an increase in $\overline{x}_j$ reduces (increases) the difference between x_j and $\overline{x}_j$, which, because of conformism, increases (decreases) utility. Thus, at the Nash equilibrium, when deciding their individual effort, agents do not take into account the effect of their effort of the social norm of their peers, which creates an externality that can be positive or negative. Indeed, if individual i has friends for whom $x_j > \overline{x}_j$ ($x_j < \overline{x}_j$), then when she exerts her effort, she does not take into account the fact that she positively affects $\overline{x}_j$, the norm of her friends, which increases (decreases) the utility of their neighbors. In that case, compared to the first best, individual i underinvests (overinvests) in effort, because she exerts positive (negative) externalities on her friends.

This result contrasts with that obtained in the local-aggregate model in which agents always underinvest in effort, because they always exert positive externalities on their neighbors. Here, even though the efforts are *strategic complements* (see (8)), agents can exert positive or negative externalities on their neighbors. This is why, in the local-aggregate model, the planner always wants to subsidize agents (Helsley and Zenou, 2014) while, in the local-average model, the planner subsidizes agents who underinvest in effort and taxes agents who overinvest in effort. We investigate these issues in detail in Section 4.3 below.

Part (*ii*) of Proposition 7 gives an exact condition on the productivity vector $\boldsymbol{\alpha}$ that ensures that the Nash equilibrium in efforts is always optimal. Unfortunately, this condition is very unlikely to hold in most networks, as shown in Online Appendix E.

Finally, in part (*iii*), we demonstrate that, for any network, the aggregate first-best effort is independent of λ, the taste for conformity, and is equal to the aggregate productivity in the network. In particular, this implies (see Proposition E1 in Online Appendix E) that, for *regular networks*, we have:

$$\sum_{i=1}^{n} x_i^* = \sum_{i=1}^{n} x_i^O = \sum_{i=1}^{n} \alpha_i.$$

In other words, for regular networks, even if the individual effort is generally not optimal, the aggregate effort in a network is always optimal. This is because, in regular networks, the positive and negative externalities imposed by agents on their neighbors exactly cancel out, so that the aggregate effect is optimal. Consequently, when the network is regular, some agents overinvest while others underinvest, and it is not possible that all agents underinvest. This result stands in sharp contrast to the local aggregate model, in which all agents exert too little effort in equilibrium, regardless of whether the network is regular or not.

Remark 2 *If agents are ex ante homogeneous in productivity, that is, $\alpha_i = \alpha_j$ for all $i, j = 1, 2, \ldots, n$, then the Nash equilibrium in effort is always optimal. Furthermore, if $\det(\widehat{\mathbf{G}}) \neq 0$, the converse is also true.*

Indeed, if agents are ex ante homogeneous, we know that, in equilibrium, the position in the network does not matter and all agents exert the same effort level, which is equal to the common social norm in the network. As a result, there are no more social interactions, since $x_i = \overline{x}_i$, for all i, and each utility depends only on own productivity. Thus, the equilibrium is always optimal.

In Online Appendix F, we illustrate condition (31) for specific networks. We show that for the equilibrium efforts to be optimal, there needs to be some compensation for the externalities that agents exert on others. In particular, for bipartite networks, such as the star and circular network, the average productivity of the different agents has to be the same, which is very unlikely to be the case.

4.2 Equilibrium versus first-best outcomes in a sufficiently conformist society

Let $\text{Corr}(\boldsymbol{\pi}, \boldsymbol{\alpha})$ be the correlation between the productivity distribution $\boldsymbol{\alpha}$ and the degree distribution $\boldsymbol{\pi}$.

Proposition 8 (First best in a sufficiently conformist society)
If $\text{Corr}(\boldsymbol{\pi}, \boldsymbol{\alpha}) < 0$ $(\text{Corr}(\boldsymbol{\pi}, \boldsymbol{\alpha}) > 0)$*, then there exists* $\underline{\lambda} \in (0,1)$ *such that, in equilibrium, for any* $\lambda > \underline{\lambda}$*, all agents underinvest (overinvest) in effort compared to the first best.*

This result implies that more central agents make higher effort and exert stronger externalities on their neighbors. As a result, all agents overprovide effort. This result is true when λ is sufficiently high since, in that case, externalities to neighbors become very important. For example, in a star network with three agents, we show below that λ does not need to be very high ($\lambda > \underline{\lambda} = 1/2$) for the result in Proposition 8 to hold (see footnote 31).

Remark 3 *In a perfectly conformist society,*

$$\lim_{\lambda \to 1} x_i^O = \frac{1}{n} \sum_{j=1}^{n} \alpha_j, \qquad \text{for all } i = 1, 2, \dots, n \tag{33}$$

This result shows that, when the society becomes perfectly conformist, the first-best effort is the same for all agents and does not depend on the position of each agent in the network. All agents should make an effort equal to the average productivity in the network. This implies that, unless the network is regular, the equilibrium in effort is never optimal when λ is sufficiently close to 1.

4.3 Restoring the first best

Let us return to the general case in which λ can take any value and assume that condition (31) does not hold. Then, to restore the first best, the planner can either subsidize or tax efforts. Let S_i^O denote the optimal per-effort subsidy for each agent i, where

$$S_i^O = \frac{\lambda}{(1-\lambda)} \sum_{j \neq i} \widehat{g}_{ji} \left(x_j^O - \overline{x}_j^O \right),$$

If we add one stage before the effort game is played, the planner announces the optimal per-effort subsidy S_i^O for each agent i such that,

$$U_i^{S_i^O} = \left(\alpha_i + S_i^O\right) x_i - \frac{1}{2} x_i^2 - \frac{1}{2} \left(\frac{\lambda}{1-\lambda} \right) \left(x_i - \overline{x}_i \right)^2 \tag{34}$$

Observe that, when each agent i chooses x_i that maximizes (34), she takes S_i^O as given, in particular, x_j^O and $\overline{x}_j^O$. In that case, the solution of this maximization problem for each agent i is the first-best.

Proposition 9 (Subsidies) *The first best is restored if the social planner gives to each agent i the following tax/subsidy per unit of effort:*

$$S_i^O = \frac{\lambda}{(1-\lambda)} \sum_{j \neq i} \widehat{g}_{ji} \left(x_j^O - \overline{x}_j^O \right) \tag{35}$$

or, in matrix form:

$$\mathbf{S}^O = \frac{\lambda}{(1-\lambda)} \, \widehat{\mathbf{G}}^T \left(\mathbf{I} - \widehat{\mathbf{G}} \right) \mathbf{x}^O.$$

By doing so, the planner restores the first best and subsidizes (taxes) agents *whose neighbors make efforts above (below) their social norms*. In other words, it is necessary to subsidize agents who exert effort below that of their neighbors and to tax those who exert effort above that of their neighbors.

Let us illustrate this result with an example. Assume a star network in which $n = 3$, and agent $i = 1$ is the star. Set $\alpha_1 = 2$, $\alpha_2 = \alpha_3 = 1$, so that the star is more productive than the peripheral agents are. Since $\alpha_1 = 2 > 1 = (\alpha_2 + \alpha_3)/2$, condition (31) is not satisfied, and hence, the Nash equilibrium is not optimal. We have

$$\mathbf{x}^* = \frac{1}{(1+\lambda)} \begin{pmatrix} 2+\lambda \\ 1+2\lambda \\ 1+2\lambda \end{pmatrix}, \qquad \mathbf{x}^O = \frac{1}{(1+4\lambda)} \begin{pmatrix} 2+5\lambda \\ 1+6\lambda \\ 1+6\lambda \end{pmatrix}.$$

The star agent *overinvests* compared to the first best $\left(x_1^* > x_1^O\right)$. Indeed, since $x_2^* = x_3^* < \overline{x}_2^* = \overline{x}_3^* = x_1^*$, the externality term $\lambda \sum_{j=1}^n \widehat{g}_{ji} \left(x_j - \overline{x}_j \right)$ (see (29)) is *negative* and the star, when deciding her effort level, does not take into account the *negative externalities* she exerts on agents 2 and 3. For the peripheral agents 2 or 3, we obtain $x_2^* = x_3^* \gtreqless x_3^O = x_2^O \iff \lambda \gtreqless 1/2$, so that they may overinvest or underinvest in effort, depending on the value of λ.[31] However, the externality term is always *positive*, since $x_1^* > \overline{x}_1^*$ and thus, agents 2 and 3 always exert *positive externalities* on agent 1. As a result, the planner should tax agent 1 and subsidize agents 2 and 3. Since $x_2 = x_3$, it is easily verified that the subsidies per unit of effort are equal to $S_1^O = \frac{2\lambda}{(1-\lambda)}(x_2^O - x_1^O) < 0$ and $S_2^O = S_3^O = \frac{\lambda}{(1-\lambda)}(x_1^O - x_2^O) > 0$. The subsidies or taxes exactly correct for the externalities exerted by the agents. We obtain:

$$\mathbf{S}^O = \frac{\lambda}{(1+4\lambda)} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \tag{36}$$

Clearly, this result strongly depends on the productivity values. For example, if $\alpha_1 = 0.5$ and $\alpha_2 = \alpha_3 = 1$ so that the productivity of the central agent is the lowest, then, to restore the first best, the planner now needs to subsidize agent 1 (the star) and to tax agents 2 and 3 (the peripheral agents) since, now, the former exerts *positive externalities* on agents 2 and 3 while the latter exert *negative externalities* on agent 1.

5 Extensions

In this section, we develop several extensions of the baseline local-average model. We consider weighted networks, heterogeneous tastes for conformity, anti-conformist attitudes, ambitious behavior, and network formation. These extensions show how various features of individual behavior affect our main results and how our model can be applied to a wide range of different contexts.

[31] Observe that, for the star network with $n = 3$ and $\alpha_1 = 2$, $\alpha_2 = \alpha_3 = 1$, we have

$$\sum_{j=1}^{3} \pi_j \alpha_j = \frac{\alpha_1}{2} + \frac{\alpha_2 + \alpha_3}{4} = \frac{3}{2} > \frac{4}{3} = \frac{1}{3}\sum_{j=1}^{3} \alpha_j,$$

which means that $Corr(\boldsymbol{\pi}, \boldsymbol{\alpha}) > 0$, since the star has both a higher productivity and a higher degree than the peripheral agents. Thus, our results confirm Proposition 8. When $\lambda > \underline{\lambda} = 1/2$, all three agents in the star network *overinvest* in effort compared to the first best. It is also easily verified that, if we now assume for the same network that $\alpha_1 = 1$, $\alpha_2 = \alpha_3 = 2$, then, $\sum_{j=1}^{3} \pi_j \alpha_j = \frac{3}{2} < \frac{5}{3} = \frac{1}{3}\sum_{j=1}^{3} \alpha_j$, and thus, $Corr(\boldsymbol{\pi}, \boldsymbol{\alpha}) < 0$. In this case, if $\lambda > \underline{\lambda} = 1/2$, all three agents in the star network *underinvest* in effort compared to the first best.

5.1 Weighted networks

Consider, first, an extension of the baseline local-average model in which the network is *directed*, *weighted*, and may have *self-loops*, as in the standard DeGroot model (Golub and Jackson, 2010). Let $\mathbf{W} = [w_{ij}]$ be an arbitrary $(n \times n)$ row-normalized irreducible matrix with non-negative entries. Each cell w_{ij}, $i, j = 1, 2, \ldots, n$, gives the relative impact (weight) of agent j's effort on agent i's social norm $\overline{x}_i$, defined as follows: $\overline{x}_i \equiv \sum_{j=1}^{n} w_{ij} x_j$. In particular, we do not rule out self loops, that is, we allow for the possibility that $w_{ii} > 0$ for some i. Otherwise, agent i's utility function is the same as in the baseline model and given by (3).

In Online Appendix G.1, we study this more general model with the adjacency matrix $\mathbf{W}$ and show that most of our results (total conformism, comparative statics, and welfare) remain qualitatively the same.

In Proposition G3, we show that a slight change in the network $\mathbf{W}$ may increase everyone's effort if highly productive agents have more impact on everyone's social norms. This echoes our result established in Proposition 6 in which we demonstrate that adding a link in the network may increase everyone's effort if this link is between two highly productive agents. However, in the weighted network model, this result is much easier to prove, because the standard calculus technique can be used to study the consequences of small changes in the weights on outcomes.

5.2 Heterogeneous tastes for conformity

In Online Appendix G.2, we relax the assumption that λ is the same across all agents by allowing each agent i to have a specific taste for conformity λ_i. We first show that our existence, uniqueness, and interiority results when λ is the same for all agents (Proposition 1) are robust to this extension.

Then, Proposition G4 provides additional intuition about the non-monotonicity results of Proposition 5. We show that higher conformity of some agents—namely, those who exert efforts below their social norms—increases everyone's effort because of strategic complementarities, while higher conformity of the others has the opposite effect. Therefore, it is not surprising that the total effect is ambiguous, as Proposition 5 states.

5.3 Anti-conformism

We now consider what happens if agents are *anti-conformist*,[32] that is if the taste for conformity θ is negative. In this case, the magnitude of $|\theta|$ can be viewed as the degree to which an agent wants to be different from the others (although not necessarily better than the others). In other words, each individual obtains a benefit of $\frac{\theta}{2}\left(x_i - \overline{x}_i\right)^2$ if she does *not* conform to the norm of her neighbors. This model can still be considered a local-average model but it is now a game with strategic *substitutes* ($\theta < 0$) instead of strategic *complements* ($\theta > 0$).

In Proposition G5 of Online Appendix G.3, we derive our main results for the anti-conformist model. We show that our model can be extended to the case of anti-conformity if agents are not too non-conformist ($|\theta| < 1/2$), although it loses a good deal of tractability. In particular, because we have a game with strategic substitutes, even if the equilibrium is unique, it is not always interior.

For example, in the case of a dyad ($n = 2$), in Online Appendix G.3, we show that the two agents exert strictly positive effort only if they are not too heterogeneous in terms of productivities, not too anti-conformist, or both. When this is not the case, then some agent may exert zero effort (see (G.25), which totally characterizes the Nash equilibrium for the dyad network). Indeed, in the anti-conformist model with a dyad network, when the difference in productivity between the two agents is too high, then it becomes optimal for the low-productivity agent to exert zero effort, because she wants to differentiate herself as much as possible from the high-productivity agent (whose effort is her social norm). On the contrary, when the productivity difference is not too large, then the low-productivity agent can still differentiate herself from the high-productivity agent and exert positive effort. This never happens in the conformist model, because agents always want to be as close as possible to each other.[33]

Furthermore, we show that, in contrast to Proposition 2, the impact of α_j on x_i^* is a priori ambiguous. Finally, we demonstrate that if agents are very anti-conformist, there are either multiple equilibria or an equilibrium fails to exist.

[32]See Bramoullé et al. (2004), Bramoullé (2007), Grabisch and Rusinowska (2010a,b), and Grabisch et al. (2017) for network models with anti-conformist agents but in very different settings.

[33]Observe that, when all agents have the same productivity α and $|\theta|$ is low enough, then the conformist and anti-conformist models lead to the same outcome, that is, all agents make an effort equal to α. When $|\theta|$ becomes larger, then even with ex ante identical agents and a regular network, in the anti-conformist model, there may be multiple equilibria in which one agent makes a higher effort than the other. See Figure G5 in Online Appendix G.3. This never occurs in the conformist model since, with identical productivities, all agents always make the same effort for any possible network, including the regular one.

5.4 Ambition and social norms

It seems realistic to assume that agents may benefit from choosing an effort that is higher than the average effort of their neighbors. To address this issue, let us extend our utility function (3) so that, for each individual i, it is now given by

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{x_i^2}{2} - \frac{\theta}{2}\left(x_i - \beta_i \overline{x}_i\right)^2,$$

where $\beta_i \geq 1$ is agent i's *ambition factor*. Since $\beta_i \geq 1$, the "reference effort" of each individual i is now higher than the social norm $\overline{x}_i$ of her neighbors.[34] Denote $\beta_{\max} \equiv \max\{\beta_1, \ldots, \beta_n\}$. Then, if $\lambda\beta_{\max} < 1$, there exists a unique interior equilibrium.

To investigate the welfare properties of this model, assume that all agents are ex ante identical, that is, $\alpha_1 = \ldots = \alpha_n = \alpha > 0$, and that they have the same ambition factor, that is, $\beta_1 = \ldots = \beta_n = \beta > 1$. In that case, we show in Appendix G.4, that, in equilibrium, each individual exerts an effort above the social norm (average effort) of their direct friends.

Moreover, for *regular networks*, we show that all agents *overinvest* in equilibrium compared to the social optimum. For *non-regular networks*, we demonstrate (see Proposition G6) that if the agents are either sufficiently conformist (high λ) or sufficiently non-conformist (low λ), they all overinvest in equilibrium compared to the first best. These results contrast with Proposition 7 for the benchmark model, in which the equilibrium is socially optimal when productivities do not vary across agents. This is because ambitious behavior creates an additional positive externality, which cannot be fully internalized by individuals even in the absence of any ex ante heterogeneity.

5.5 Network formation

Thus far, we have assumed that the network is fixed and taken as given when agents decide their effort level. Consider now a two-stage game in which, in the first stage, agents create links (endogenous network formation) while, in the second stage, they exert effort.

Assume, for simplicity, that there are two types of agents: high-productivity agents for which $\alpha_i = \alpha^H$ and low-productivity agents for which $\alpha_i = \alpha^L$, with $\alpha^H > \alpha^L > 0$. Assume also that creating or severing a link is costless.

[34] See Ghiglino and Goyal (2010), who also develop a model in which agents want to consume more than the average consumption (social norm) of their neighbors.

In Proposition G7 in Online Appendix G.5, we show that, in the *local-aggregate model*, the only pairwise Nash equilibrium[35] is the *complete network* in which all agents of any type are linked to each other. On the contrary, in the *local-average model*, the only pairwise Nash equilibrium is a network of two disconnected components; in each component, all agents of the same type form a complete network. This network is called the *completely homophilous network*.

This means that, in the local-aggregate model, there is complete "integration" of the two types of agents while, in the local-average model, there is complete "segregation" of the two types of agents so that *extreme homophily behavior* prevails in equilibrium. In other words, in the local-aggregate model, even if agents are heterogeneous in terms of productivities, complete homophily cannot emerge because, independently of the type, there is always a benefit of forming new links due to strong positive spillovers.

On the contrary, in the local-average model, an agent of one type never wants to form a link with an agent of the other type. Indeed, when agents have the same productivity α, independently of their position in the network, they all exert the same effort level and have the same social norm, both equal to α. As a result, they no longer bear the cost of not conforming to their social norm and their equilibrium utility equals $\alpha^2/2$. However, if an agent forms a link with someone of a different type, she suffers an extra *loss*, because a gap between her effort and her social norm emerges. For this reason, in the local-average model, agents of one type are better off not having links with agents of the other type. Using the same reasoning, one can show that, if we introduce a cost of forming and severing links, we still have the same pattern, that is, complete homophily or segregation in the local-average model, and integration and heterophily in the local-aggregate model, but there may be more than one equilibrium. Furthermore, we can easily generalize our results to more than two types of agents.

6 Policy implications: local-average versus local-aggregate model

As stated in the Introduction, there are two main models of games on networks with positive peer effects (strategic complementarities): the local-average and the local-aggregate model. In the local-average model, deviating from the average effort of one's peers negatively affects the utility of an individual (see (3)). The closer each individual's effort is to the average of her friends' efforts, the higher is her utility.

[35] For a precise definition of pairwise Nash equilibrium, see Bloch and Jackson (2006).

By contrast, in the local aggregate model, the sum of the efforts of an individual's peers positively affects the utility of each individual (see (4)). When peers exert more effort, the utility derived from own effort increases.

We believe that it is important to be able to disentangle different behavioral peer-effect models because, even if they look very similar, they have different policy implications. To highlight these differences between the models, we consider in the next subsection education and crime and observe how these two models yield different policy implications.

6.1 Policy implications: Education

In terms of education, if the local-aggregate model describes well the preferences of students (Calvó-Armengol et al., 2009), then any individual-based policy, such as *vouchers*, would be efficient, because if one or more "key" students (e.g., the disruptive ones) are positively affected by the policy, because of peer effects (social multiplier), many other students are also positively affected. If, on the contrary, we believe that the local-average model describes students' preferences more adequately, then we should change the social norm in the school or classroom (group-based policy) and attempt to implement the idea that it is "cool" to work hard at school. Affecting a few students will not change anything if it does not change the social norm in the school.

An example of an educational policy that has attempted to change the social norm of students is the *charter-school policy*. Charter schools are very good at screening teachers and selecting the best ones. In particular, the "No Excuses policy" (Angrist et al., 2010, 2012) is a highly standardized and widely replicated charter model that features a long school day, an extended school year, selective teacher hiring, and strict behavioral norms, while it emphasizes traditional reading and math skills. The main objective is to change the social norms of disadvantaged children by being very strict on discipline. This is a typical policy that is in accordance with the local-average model, since its aim is to change the social norms of students in terms of education. Angrist et al. (2012) focus on special needs students who may be underserved. The study's results show average achievement gains of 0.36 standard deviations in math and 0.12 standard deviations in reading for each year spent at a charter school called the Knowledge is Power Program (KIPP) Lynn, with the largest gains coming from the Limited English Proficient (LEP), Special Education (SPED), and low-achievement groups. The authors show that the average reading gains were driven almost entirely by SPED and LEP students, whose reading scores rose by roughly 0.35 standard deviations for each year spent at KIPP Lynn.[36]

In summary, an effective policy for the local-average model would be to change people's perceptions of "normal" behavior (i.e., their social norm) so that a *school-based policy* could be implemented. Meanwhile, for the local-aggregate model, this would not be necessary and an *individual-based policy* should instead be implemented.

6.2 Policy implications: Crime

It is well documented that crime is, to a large extent, a group phenomenon, and the source of crime is located in the intimate social networks of individuals (see, e.g., Warr, 2002; Bayer et al., 2009; Damm and Dustmann, 2014).

In the local-aggregate model, a *key-player policy* (Ballester et al., 2006; Zenou, 2016; Lee et al., 2018), whose aim is to remove the criminal that reduces total crime in a network the most, would be the most effective way of reducing total crime.[37] In other words, the removal of the key player can have large effects on crime because of the feedback effects or "social multipliers" at work. Indeed, as the proportion of individuals participating in criminal behavior increases, the impact on others is multiplied through social networks. Thus, criminal behavior can be magnified, and interventions can become more effective.

On the contrary, a key-player policy would have nearly no effect in the local-average model, since it would not affect the social norm that committing crime is morally wrong. To be effective, one would have to change the norm for each of the criminals, which is clearly a more difficult objective. In that case, it is necessary to target a group or gang of criminals to reduce crime drastically. This illustrates the fact that, for the local-aggregate model, *individual-based policies* are more appropriate while, for the local-average model, *group-based policies* are more effective.[38]

[36]See also Curto and Fryer (2014), who study the SEED schools, which are boarding schools serving disadvantaged students located in Washington DC and Maryland. The SEED schools, which combine a "No Excuses" charter model with a 5-day-a-week boarding program, are the United States' only urban public boarding schools for the poor for students in grades 6–12. Using admission lotteries, Curto and Fryer (2014) show that attending a SEED school increases achievement by 0.211 standard deviation in reading and by 0.229 standard deviation in math per year.

[37]In Section 3.3, we also discuss the difference between the local-average and local-aggregate models in terms of the *key-link* policy, whose aim is to choose how to optimally remove a link between two criminals in order to minimize the total crime level in a network. See, in particular, Remark 1 and the discussion that follows.

[38]For recent overviews on place-based policies, see Kline and Moretti (2014) and Neumark and Simpson (2015).

6.3 Which model is the most empirically relevant?

Which model is relevant is clearly an empirical question. To statistically identify whether the average model or the aggregate model is more appropriate for a particular outcome, Liu et al. (2014) proposed the following methodology. It is necessary to estimate an augmented model, which includes both average and aggregate peer effects, and to determine which one is statistically significant. Using data for the National Longitudinal Study of Adolescent to Adult Health (Add Health), Liu et al. (2014) showed that, for study effort in education, the endogenous peer effect is mostly captured by a social-conformity (local average) effect rather than a social-multiplier (local aggregate) effect. This implies that a charter-school policy that aims to change the social norms of students (as in Angrist et al., 2010, 2012) would be the most effective policy to improve education in schools. On the other hand, for sport activities, Liu et al. (2014) found that both social-conformity and social-multiplier effects contribute to the endogenous peer effect. Moreover, Lee et al. (2018), who studied juvenile delinquency, showed that the local-aggregate model is at work for the AddHealth data. This implies that a key-player policy would be the most effective policy to reduce crime for adolescents in the United States.

6.4 An illustrative example

Let us illustrate the above discussion about individual versus group-based policy with a simple example. Consider the network $\mathbf{g}$ in Figure 2 with $n = 11$ players. This network was considered by Ballester et al. (2006) to illustrate their formula of the key player. In this network, player 1 bridges together two fully intra-connected groups with five players each.

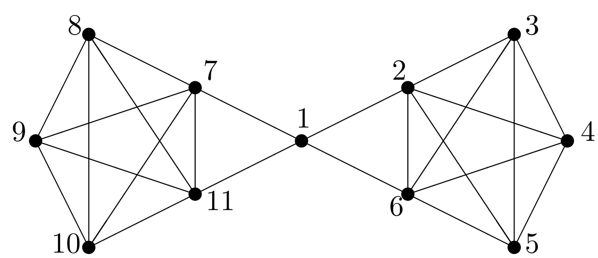

Figure 2: A bridge network

6.4.1 An individual-based policy: Key player

Consider a network-crime model in which agents choose crime effort that maximizes their utility, which can be based on either the local-aggregate or the local-average model. As an illustration of an individual-based policy, we consider the key player policy, which consists of determining the player who, once removed from the network, reduces total crime effort the most. To make the comparison between the two models easier, we assume that agents are ex ante identical, that is, $\alpha_i = 1$, for all $i = 1, \ldots n$. We also assume that $\theta = 0.2$.

The local-aggregate model Consider the local-aggregate model whereby the utility function is given by (4). Then, if $\theta < 1/\mu(\mathbf{g})$ (where $\mu(\mathbf{g})$ is the largest eigenvalue of $\mathbf{g}$),[39] then a unique Nash equilibrium in efforts exists, which is equal to:

$$\mathbf{x}^* = (\mathbf{I} - \theta\mathbf{G})^{-1}\,\mathbf{1}$$

It is easily verified that the key player is agent 1 (Ballester et al., 2006). In particular, the total crime effort in equilibrium is equal to 91.67 while, after the removal of individual 1, it is 50. Thus, the removal of player 1 leads to a decrease of total crime activity by 45.46%. This is because removing player 1 disrupts the network and leads to two different networks that are no longer connected. The change in efforts after the removal of agent 1 varies a lot depending on the position in the network. For example, agent 2, who was directly linked to 1, reduces her effort from 9.17 to 5 (45.47% reduction) while agent 3, who was two links away from 1, decreases her effort from 7.78 to 5 (35.73% reduction).

The local-average model Consider now the local-average model in which the utility function is given by (3). We have shown that the Nash equilibrium is given by

$$\mathbf{x}^* = \widehat{\mathbf{M}}\boldsymbol{\alpha} = \frac{1}{(1+\theta)}\left(\mathbf{I} - \frac{\theta}{(1+\theta)}\mathbf{G}\right)^{-1}\mathbf{1}$$

It is easily verified that all agents make the same effort level equal to 1 (which is the social norm) so that total crime effort is 11. Let us remove player 1 (or in fact any other player) and renormalize the resulting adjacency matrix. It is easily checked that nothing changes since each player still makes an effort of 1 and the social norm is exactly the same and equal to 1. Because there is one less player in the network, the total effort is now given by 10 and the reduction in total crime is then equal to 9.09%.

In summary, an individual policy, such as the key player, has a big impact on total crime when the preferences of agents are based on the local-aggregate model while

[39] This condition is verified for the network displayed in Figure 2, since $\theta = 0.2 < 0.227 = 1/\mu(\mathbf{g})$.

it has nearly no impact when the preferences are based on the local-average model. As a result, if the planner believes that the agents have preferences according to the local-aggregate model and implements a key-player policy while, in fact, agents have local-average preferences, then this example shows that this policy will fail to reduce crime, as agents will not change their criminal behavior.

6.4.2 A group-based policy: Changing the norm

Consider again the network **g** in Figure 2 and implement a group-based policy, which is common to everybody. For example, consider a reduction of α from 1 to 0.7. All agents in the network are affected in the same way.

The local-aggregate model By implementing such a policy, it is easily verified that total crime effort decreases from 91.67 (before the policy) to 64.17 (after the policy), giving a reduction in total crime of 30%.

The local-average model In this model, the effort and social norm change for all agents in the network. It is easily verified that all agents now reduce their crime effort to 0.7 and the social norm is now given by 0.7. As a result, we switch from a total crime effort of 11 (before the policy) to 7.7 (after the policy), that is, a reduction in total crime of 30%. In other words, changing the social norm from 1 to 0.7 now has a large impact on total crime in the network.

In summary, a group-based policy, such as changing the social norm by reducing the productivity of all agents in the network, has a much bigger impact on total crime when the preferences of agents are based on the local-average model. However, a group-based policy is less efficient when the agent's preferences are based on the local-aggregate model. Again, if the planner has the wrong beliefs about agents' preferences, then the impact of a group-based policy on reducing crime may be limited.

7 Concluding remarks

In this study, we analyze the linear-in-means model (also known as the local-average model in the network literature), which is the workhorse model in empirical work on peer effects. Apart from their position in the network, agents are heterogeneous in terms of productivity. We characterize the Nash equilibrium in efforts of this game in which each agent minimizes the social distance between her own effort and that of her peers (her own social norm). While individual productivity positively affects equilibrium effort, the impact of taste for conformity is non-monotone. Both the sign and the magnitude of this conformity effect depend on whether an individual is above or below her own social norm. We also study how adding or removing a link affects the aggregate effort in the network and show that it depends on the productivity of the agents involved in the link. Equilibria are usually inefficient and we provide a condition on the productivity distribution and the network structure that guarantees the efficiency of equilibrium. Because this condition often fails to hold, we show how to restore the first best. Unexpectedly, the optimal taxation/subsidy scheme is to subsidize agents whose peers would exert efforts above their social norms while taxing agents whose peers would exert efforts below their social norms. Hence, the planner does not necessarily subsidize central agents, as is the case in the local-aggregate model.

More generally, we consider our framework to be rich enough to encompass many real-world situations in which people are conformist and dislike to deviate from the social norms of their friends. We also believe that our results lead to important policy implications that can be tested empirically. In particular, we shed light on the debate on whether individual-based policies are more effective in maximizing welfare or minimizing total activity (in the case of crime) than group- or place-based policies.

References

[1] Akerlof, G.A. (1980), "A theory of social custom of which unemployment may be one consequence," *Quarterly Journal of Economics* 94, 749–775.

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

[3] Allouch, N. (2015), "On the private provision of public goods on networks," *Journal of Economic Theory* 157, 527–552.

[4] Alm, J., Bloomquist, K.M., and McKee, M. (2017), "When you know your neighbour pays taxes: Information, peer effects, and tax compliance," *Fiscal Studies* 38, 587–613.

[5] Angrist, J.D., Dynarski, S.M., Kane, T.J., Pathak, P.A., and Walters, C.R. (2010), "Inputs and impacts in charter schools: KIPP Lynn," *American Economic Review Papers and Proceedings* 100, 239–243.

[6] Angrist, J.D., Dynarski, S.M., Kane, T.J., Pathak, P.A. and Walters, C.R. (2012), "Who benefits from KIPP?" *Journal of Policy Analysis and Management* 31, 837–860.

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

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

[9] Banerjee, A.V., Chandrasekhar, A.G., Duflo, E., and Jackson, M.O. (2013),"The diffusion of microfinance,"*Science* 341(6144).

[10] Bayer, P., Hjalmarsson, R., and Pozen. D. (2009), "Building criminal capital behind bars: Peer effects in juvenile corrections," *Quarterly Journal of Economics* 124, 105–147.

[11] Bernheim, B.D. (1994), "A theory of conformity," *Journal of Political Economy* 102, 841–877.

[12] Bloch, F., and Jackson, M.O. (2006), "Definitions of equilibrium in network formation games," *International Journal of Game Theory* 34, 305–318.

[13] Bloch, F., and Quérou, N. (2013), "Pricing in social networks," *Games and Economic Behavior* 80, 243–261.

[14] Blume, L.E., Brock, W.A., Durlauf, S.N., and Jayaraman, R. (2015), "Linear social interactions models," *Journal of Political Economy* 123, 444–496.

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

[16] Boucher, V. (2016), "Conformism and self-selection in social networks," *Journal of Public Economics* 136, 30–44.

[17] Boucher, V. and Fortin, B. (2016), "Some challenges in the empirics of the effects of networks," In: Bramoullé, Y., Galeotti, A., and Rogers, B.W. (Eds.), *The Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, pp. 277–302.

[18] Bramoullé, Y. (2007), "Anti-coordination and social interactions," *Games and Economic Behavior* 58, 30–49.

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

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

[21] Bramoullé, Y. and Kranton, R. (2016), "Games played on networks," In: Bramoullé, Y., Rogers, B.W., and Galeotti, A. (Eds.), *The Oxford Handbook of the Economics of Networks*, Oxford: Oxford University Press.

[22] Bramoullé, Y., Kranton, R., and D'Amours, M. (2014), "Strategic interaction and networks," *American Economic Review* 104, 898–930.

[23] Bramoullé, Y. López-Pintado, D., Goyal, S., and Vega-Redondo, F. (2004), "Network formation and anti-coordination games," *International Journal of Game Theory* 33, 1–19.

[24] Brekke, K.A., Kipperberg, G., and Nyborg, K. (2010), "Social interaction in responsibility ascription: The case of household recycling," *Land Economics* 86, 766–784.

[25] Breza, E.L. (2016), "Field experiments, social networks, and development," In: Bramoullé, Y., Rogers, B.W., and Galeotti, A. (Eds.), *The Oxford Handbook of the Economics of Networks*, Oxford: Oxford University Press, pp. 412–439.

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

[27] Carrell, S.E., Sacerdote, B.I., and West, J.E. (2013), "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica* 81, 855–882.

[28] Chen, Y.J., Zenou, Y., and Zhou, J. (2018), "Competitive pricing strategies in social networks," *RAND Journal of Economics* 49, 672–705.

[29] Christakis, N.A., and Fowler, J.H. (2007), "The spread of obesity in a large social network over 32 years," *New England Journal of Medicine* 357, 370–379.

[30] Clark, A.E., and Loheac, Y. (2007), "It wasn't me, it was them! Social influence in risky behaviour by adolescents," *Journal of Health Economics* 26, 763–784.

[31] Curto, V., and Fryer, R. (2014), "The potential of urban boarding schools for the poor: Evidence from SEED," *Journal of Labor Economics* 32, 65–93.

[32] Czajkowski, M., Hanley, N., and Nyborg, K. (2017), "Social norms, morals and self-interest as determinants of pro-environment behaviours: The case of household recycling," *Environment Resource Economics* 66, 647–670.

[33] Damm, A.P., and Dustmann, C. (2014), "Does growing up in a high crime neighborhood affect youth criminal behavior?" *American Economic Review* 104, 1806–1832.

[34] DeGroot, M.H. (1974), "Reaching a consensus," *Journal of the American Statistical Association* 69, 118–121.

[35] De Marti, J., and Zenou, Y. (2015), "Networks games under incomplete information," *Journal of Mathematical Economics* 61, 221–240.

[36] DeMarzo, P.M., Vayanos, D., and Zwiebel, J. (2003), "Persuasion bias, social influence, and unidimensional opinions," *Quarterly Journal of Economics* 118, 909–968.

[37] Dutta, R., Levine, D.K., and Modica, S. (2019), "Interventions with sticky social norms: A critique," Unpublished manuscript, McGill University.

[38] Epple, D., and Romano, R. (2011), "Peer effects in education: A survey of the theory and evidence," In: Benhabib, J., Bisin, A., and Jackson, M.O. (Eds.), *Handbook of Social Economics, Vol. 1B*, Amsterdam: Elsevier, pp. 1053-1163.

[39] Fershtman, C., and Weiss, Y. (1998), "Social rewards, externalities and stable preferences," *Journal of Public Economics* 70, 53–73.

[40] Fortin, B., Lacroix, G., and Villeval, M.-C. (2007), "Tax evasion and social interactions," *Journal of Public Economics* 91, 2089–2112.

[41] Ghiglino, C., and Goyal, S. (2010), "Keeping up with the neighbors: Social interaction in a market economy," *Journal of the European Economic Association* 8, 90–119.

[42] Golub, B., and Jackson, M.O. (2010), "Naïve learning in social networks and the wisdom of crowds" *American Economic Journal: Microeconomics* 2, 112–149.

[43] Golub, B., and Morris, S. (2017), "Expectations, networks, and conventions," Unpublished manuscript, Harvard University.

[44] Golub, B. and Sadler, E. (2016), “Learning in social networks,” In: Bramoullé, Y., Galeotti, A., and Rogers, B.W. (Eds.), *The Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, pp. 504–542.

[45] Grabisch, M., Poindron, A., and Rusinowska, A. (2017), “A model of anonymous influence with anti-conformist agents,” CES Working Paper No. 2017.47.

[46] Grabisch, M. and A. Rusinowska (2010a), “A model of influence in a social network,” *Theory and Decision* 69, 69–96.

[47] Grabisch, M. and A. Rusinowska (2010b), “A model of influence with an ordered set of possible actions,” *Theory and Decision* 69, 635–656.

[48] Hamilton, J.D. (1994), *Time Series Analysis*, Princeton: Princeton University Press.

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

[50] Herbst, D., and Mas, A. (2015), “Peer effects in worker output in the laboratory generalize to the field,” *Science* 350, 6260.

[51] Horn, R.A., and Johnson, C.R. (1985), *Matrix Analysis*, Cambridge: Cambridge University Press.

[52] Hsieh, C.-S., and Lin, X. (2017), “Gender and racial peer effects with endogenous network formation,” *Regional Science and Urban Economics* 67, 135–147.

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

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

[55] Jackson, M.O., Rogers, B., and Zenou, Y. (2017), “The economic consequences of social network structure”, *Journal of Economic Literature* 55, 49–95.

[56] Jackson, M.O., and Zenou, Y. (2015), “Games on networks”, In: Young, P., and Zamir, S. (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier Publisher, pp. 91–157.

[57] Kandel, E., and Lazear, E.P. (1992), “Peer pressure and partnerships,” *Journal of Political Economy* 100, 801–817.

[58] Kline, B., and Tamer, E. (2018), "Econometric analysis of models with social interactions," In: Graham, B. and De Paula, A. (Eds.), *The Econometric Analysis of Network Data*, Amsterdam: Elsevier, forthcoming.

[59] Kline, P., and Moretti, E. (2014), "People, places, and public policy: Some simple welfare economics of local economic development programs," *Annual Review of Economics* 6, 629–662.

[60] Kling J.R., Liebman, J.B., and Katz, L.F. (2007), "Experimental analysis of neighborhood effects," *Econometrica* 75, 83–119.

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

[62] Lee, L.-F., Liu, X., Patacchini, E., and Zenou, Y. (2018), "Who is the key player? A network analysis of juvenile delinquency," Unpublished manuscript, Monash University.

[63] LeSage, J.P., and Pace, R.K. (2009), *Introduction to Spatial Econometrics*, Boca Raton: CRC Press.

[64] Lindquist, M.J. and Zenou, Y. (2019), "Crime and networks: 10 policy lessons," *Oxford Review of Economic Policy*, forthcoming.

[65] Liu, X., Patacchini, E., and Zenou, Y. (2014), "Endogenous peer effects: Local aggregate or local average?" *Journal of Economic Behavior and Organization* 103, 39–59.

[66] Ludwig, J., Duncan, G.J., and Hirschfield, P. (2001), "Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment," *Quarterly Journal of Economics* 116, 655–679.

[67] Manski, C. (1993), "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies* 60, 531–542.

[68] Mobius M., and Rosenblat, T. (2014), "Social learning in economics," *Annual Review of Economics* 6, 827–847.

[69] Neumark, D., and Simpson, H. (2015), "Place-based policies," In: Duranton, G., Henderson, V., and Strange, W. (Eds.), *Handbook of Regional and Urban Economics*, Vol. 5B, Amsterdam: Elsevier Publisher, pp. 1197–1288.

[70] Olcina, G., Panebianco, F., and Zenou, Y. (2017), "Conformism, social norms and the dynamics of assimilation," CEPR Discussion Paper No. 12166.

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

[72] Sacerdote, B.I. (2011), "Peer effects in education: How might they work, how big are they and how much do we know thus far?" In: Hanushek, E. Machin, S., and Woessmann, L. (Eds.), *Handbook of the Economics of Education*, Dordrecht: Elsevier, pp. 249–277.

[73] Topa, G., and Zenou, Y. (2015), "Neighborhood and network effects," In: Duranton, G., Henderson, V., and Strange, W. (Eds.), *Handbook of Regional and Urban Economics*, Vol. 5A, Amsterdam: Elsevier, pp. 561–624.

[74] Warr, M. (2002), *Companions in Crime: The Social Aspects of Criminal Conduct*, Cambridge: Cambridge University Press.

[75] Zenou, Y. (2016), "Key players," In: Bramoullé, Y., Rogers, B.W., and Galeotti, A. (Eds.), *Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press, pp. 244–274.

Online Appendix

A Proofs of the results in the main text

Proof of Proposition 1.
Proof of part (i). The FOC is given by

$$x_i = (1 - \lambda)\,\alpha_i + \lambda \overline{x}_i. \tag{A.1}$$

Plugging the definition (2) of the social norm $\overline{x}_i$ into (A.1) implies (11). Restating (11) in matrix form, we obtain (12). By solving (12) for $\mathbf{x}$, we verify that the Nash equilibrium $\mathbf{x}^*$ is indeed defined by (13). The existence and uniqueness of the Nash equilibrium $\mathbf{x}^*$ is guaranteed by the fact that, for any $\lambda > 0$, we have:

$$\lambda < \rho(\widehat{\mathbf{G}}) = 1,$$

where $\rho(\widehat{\mathbf{G}})$ stands for the spectral radius of $\widehat{\mathbf{G}}$, and $\rho(\widehat{\mathbf{G}}) = 1$ holds true because $\widehat{\mathbf{G}}$ is a row-normalized matrix with non-negative entries. This proves part (i). ■

Proof of part (ii). Using (13) and observing that $\overline{\mathbf{x}}^* = \widehat{\mathbf{G}}\mathbf{x}^*$, we obtain equation (15) for the equilibrium social norms. This proves part (ii). ■

Proof of part (iii). We now prove that the equilibrium utility levels are given by (16). To do this, we use (A.1) to express agent i's equilibrium social norm $\overline{x}_i^*$ as follows:

$$\overline{x}_i^* = \frac{1}{\lambda} x_i^* - \frac{1 - \lambda}{\lambda} \alpha_i.$$

Plugging this expression into the utility function (6), and simplifying it, we obtain

$$U_i(x_i^*, \mathbf{x}_{-i}^*, \mathbf{g}) = \frac{1}{2}\left[\alpha_i^2 - \frac{1}{\lambda}\left(\alpha_i - x_i^*\right)^2\right]. \tag{A.2}$$

Plugging agent i's equilibrium effort

$$x_i^* = \sum_{j=1}^{n} \widehat{m}_{ij} \alpha_j \tag{A.3}$$

into (A.2) yields (16) and proves part (iii).

The proof of Proposition 1 is now completed. ∎

Proof of Lemma 1. First, restate the FOC as

$$x_i^* - \overline{x}_i^* = (1-\lambda)(\alpha_i - \overline{x}_i^*). \tag{A.4}$$

We have:

$$\begin{aligned}(1-\lambda)(\alpha_i - \overline{x}_i^*) &= (1-\lambda)(\alpha_i - x_i^* + x_i^* - \overline{x}_i^*) \\ &= (1-\lambda)(\alpha_i - x_i^*) + (1-\lambda)(x_i^* - \overline{x}_i^*).\end{aligned}$$

Combining this with (A.4) yields:

$$\lambda(x_i^* - \overline{x}_i^*) = (1-\lambda)(\alpha_i - x_i^*).$$

Hence,

$$x_i^* \gtreqless \overline{x}_i^* \iff \alpha_i \gtreqless x_i^*.$$

Using (A.3), this is equivalent to

$$x_i^* \gtreqless \overline{x}_i^* \iff \alpha_i \gtreqless \sum_{j=1}^{n} \widehat{m}_{ij}\alpha_j = \widehat{m}_{ii}\alpha_i + \sum_{j=1, j\neq i}^{n} \widehat{m}_{ij}\alpha_j,$$

which leads to (19). ∎

Proof of Proposition 2: We start with the following lemma.

Lemma A2 *The matrixes* $\widehat{\mathbf{M}}$ *and* $\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ *are row-normalized.*

Proof of Lemma A2 Let us first prove that $\widehat{\mathbf{M}}$ is row-normalized, that is,

$$\widehat{\mathbf{M}}\mathbf{1} = \mathbf{1},$$

where $\mathbf{1} := (1, ..., 1)^T$. Because $\widehat{\mathbf{G}}$ is row-normalized, $\widehat{\mathbf{G}}^k$ for any integer k is also row-normalized. Combining this with the power-series representation of $\widehat{\mathbf{M}}$, we obtain

$$\widehat{\mathbf{M}}\mathbf{1} = (1-\lambda)\sum_{k=0}^{\infty} \lambda^k \widehat{\mathbf{G}}^k \mathbf{1} = (1-\lambda)\sum_{k=0}^{\infty} \lambda^k \mathbf{1} = \mathbf{1}.$$

This proves that $\widehat{\mathbf{M}}$ is row-normalized. Since the product of two row-normalized matrixes is a row-normalized matrix, $\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ is also row-normalized. This completes the proof of Lemma A2. ∎

Proof of part (i): Let us now show that $0 < \frac{\partial x_i^*}{\partial \alpha_i} < 1$. We have

$$\mathbf{x}^* = \widehat{\mathbf{M}}\boldsymbol{\alpha}. \tag{A.5}$$

Hence,

$$\frac{\partial \mathbf{x}^*}{\partial \boldsymbol{\alpha}} = \widehat{\mathbf{M}},$$

which is strictly positive. Since, by Lemma A2, $\widehat{\mathbf{M}}$ is a row-normalized matrix with positive entries, it must be that $\frac{\partial x_i^*}{\partial \alpha_j} < 1$ for any $i, j = 1, \ldots, n$.

Let us now prove that $0 < \frac{\partial \overline{x}_i^*}{\partial \alpha_i} < 1$. By definition of the social norm, we have $\overline{\mathbf{x}} = \widehat{\mathbf{G}}\mathbf{x}$, and hence,

$$\overline{\mathbf{x}}^* = \widehat{\mathbf{G}}\widehat{\mathbf{M}}\boldsymbol{\alpha}. \tag{A.6}$$

As seen from (A.6), $\frac{\partial \overline{x}_i^*}{\partial \alpha_i}$ is the iith entry of $\widehat{\mathbf{G}}\widehat{\mathbf{M}}$. Since $\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ is a non-negative matrix, we have $\frac{\partial \overline{x}_i^*}{\partial \alpha_i} > 0$. Furthermore, by Lemma A2, $\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ is row-normalized. Hence, none of its entries can exceed 1, which implies that $\frac{\partial \overline{x}_i^*}{\partial \alpha_i} < 1$ and proves part (i). ■

Proof of part (ii): The payoff function of individual i is given by (6). Let us first determine $\frac{\partial U_i^*}{\partial \alpha_i}$. In equilibrium, x_i is determined as the maximizer of (6) under $\mathbf{x}_{-i} = \mathbf{x}_{-i}^*$. By the envelope theorem, we obtain

$$\frac{\partial U_i^*}{\partial \alpha_i} = x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*)\frac{\partial \overline{x}_i^*}{\partial \alpha_i}. \tag{A.7}$$

Let us state the following lemma for the rest of the proof.

Lemma A3 *The following inequalities hold for all* $i = 1, ..., n$*:*

$$\max_j \alpha_j \geq \max\{x_i^*, \alpha_i\} \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq \min\{x_i^*, \alpha_i\} \geq \min_j \{\alpha_j\}. \tag{A.8}$$

Proof of Lemma A3: Let us, first, establish the second and the third inequalities in (A.8):

$$\max\{x_i^*, \alpha_i\} \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq \min\{x_i^*, \alpha_i\}.$$

If $x_i^* \geq \overline{x}_i^*$, then it follows from equation (A.7) and Proposition 2 that

$$x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*) \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq x_i^*.$$

If, on the contrary, $x_i^* < \overline{x}_i^*$, then equation (A.7) and Proposition 2 imply

$$x_i^* \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*).$$

In summary, we have:

$$\max\left\{x_i^*,\ x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*)\right\} \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq \min\left\{x_i^*,\ x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*)\right\}. \quad \text{(A.9)}$$

The individual i's FOC can be recast as follows:

$$x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*) = \alpha_i. \quad \text{(A.10)}$$

Combining (A.9) with (A.10) proves that: $\max\{x_i^*, \alpha_i\} \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq \min\{x_i^*, \alpha_i\}$.

Let us now prove that $\max_j \alpha_j \geq \max\{x_i^*, \alpha_i\}$ and $\min\{x_i^*, \alpha_i\} \geq \min_j\{\alpha_j\}$. By Lemma A2, $\widehat{\mathbf{M}}$ is a row-normalized matrix. Since

$$x_i^* = \sum_{j=1}^{n} \widehat{m}_{ij}\alpha_j,$$

while $\widehat{m}_{ij}$ are positive and sum up to one, we have:

$$\min_j \alpha_j < x_i^* < \max_j \alpha_j.$$

This completes the proof of Lemma A3. ■

It remains to observe that Lemma A3 immediately implies that $\frac{\partial U_i^*}{\partial \alpha_i} > 0$.

Proof of part (*iii*): As implied by (16), U_i^* is a strictly concave quadratic function of α_j for any $j = 1, 2, \ldots, n \setminus \{i\}$. Hence, two cases may arise: either U_i^* decreases with α_j for all positive values of α_j, or U_i^* is bell shaped in α_j. Which of the two cases arises depends on the sign of the partial derivative $\partial U_i^*/\partial \alpha_j$ evaluated at $\alpha_j = 0$. Computing this derivative yields

$$\left.\frac{\partial U_i^*}{\partial \alpha_j}\right|_{\alpha_j=0} = \frac{1}{\lambda}\left(\alpha_i - \sum_{k=1}^{n} \widehat{m}_{ik}\alpha_k\right).$$

Hence, U_i^* is bell shaped in α_j if and only if α_i is sufficiently larger than the productivity of agents other than i and j, meaning that the following inequality holds:

$$\alpha_i > \sum_{k\neq i} \frac{\widehat{m}_{ik}}{1-\widehat{m}_{ii}}\alpha_k.$$

Otherwise, U_i^* strictly decreases in α_j.

Let us now obtain a more general result: agent i's equilibrium utility increases (decreases) in response to a small change in α_j, where $j \neq i$, if and only if agent i's equilibrium efforts are above (below) the social norm. For that, we obtain

$$\frac{\partial U_i^*}{\partial \alpha_j} = x_i^* \delta_{ij} + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*)\frac{\partial \overline{x}_i^*}{\partial \alpha_j},$$

where δ_{ij} is the Kronecker delta:

$$\delta_{ij} = \begin{cases} 1, & \text{if} \quad i=j, \\ 0, & \text{if} \quad i\neq j. \end{cases} \tag{A.11}$$

Since, by Proposition 2, $\frac{\partial \overline{x}_i^*}{\partial \alpha_j} > 0$, we obtain the desired result:

$$\operatorname{sign}\left[\frac{\partial U_i^*}{\partial \alpha_j}\right] = \operatorname{sign}\left(x_i^* - \overline{x}_i^*\right). \tag{A.12}$$

This proves part (iii) by using Lemma 1. ▮

Proof of Proposition 3. Define a_{ik}, where $i = 1, 2, \ldots, n$ and $k = 0, 1, \ldots$, as follows:

$$a_{ik} := \sum_{j=1}^{n} \widehat{g}_{ij}^{[k]} \alpha_j,$$

where $\widehat{g}_{ij}^{[k]}$ is the ijth entry of $\widehat{\mathbf{G}}^k$. Then, restating (13) in coordinate form, we can express $x_i^*(\lambda)$ as follows:

$$x_i^*(\lambda) = (1-\lambda)\sum_{k=0}^{\infty} a_{ik}\lambda^k. \tag{A.13}$$

By Lemma C6 in the Online Appendix C, two cases may arise.

Case 1: $\widehat{\mathbf{G}}$ is ergodic. Subtracting the scalar $\boldsymbol{\pi\alpha} = \sum_{j=1}^{n} \pi_j \alpha_j$ from both parts of (A.13) yields

$$x_i^*(\lambda) - \boldsymbol{\pi\alpha} = (1-\lambda)\sum_{k=0}^{\infty}\left(a_{ik} - \boldsymbol{\pi\alpha}\right)\lambda^k. \tag{A.14}$$

The ergodicity condition (C.4) in the Online Appendix C implies

$$\lim_{k\to\infty} a_{ik} = \boldsymbol{\pi\alpha} \qquad \text{for all} \qquad i = 1, \ldots, n.$$

Furthermore, it is well known (see Jackson, 2008, Chap. 8) that, for any ergodic Markov chain, convergence to the stationary distribution $\boldsymbol{\pi}$ is exponential at the rate of $|\lambda_2|$, where λ_2 is the second largest eigenvalue of $\widehat{\mathbf{G}}$ in absolute value. In other words, there exists a constant $C > 0$, such that, for any $i = 1, 2, \ldots, n$, and for any $k = 0, 1, 2, \ldots$, we have

$$|a_{ik} - \boldsymbol{\pi\alpha}| < C|\lambda_2|^k. \tag{A.15}$$

Inequality (A.15) implies that the series $\sum_{k=0}^{\infty} (a_{ik} - \boldsymbol{\pi\alpha})$ converges. Hence, by Abel's theorem (Courant and John, 2012, Ch. 7, p. 569), the expression $\sum_{k=0}^{\infty} (a_{ik} - \boldsymbol{\pi\alpha}) \lambda^k$ considered as a function of λ has a finite limit when $\lambda \to 1$. Using this, and taking the limit on both sides of (A.14) under $\lambda \to 1$, we obtain (20).

Case 2. $\widehat{\mathbf{G}}$ is periodic. Then, by Lemma C6 in the Online Appendix C, $\mathbf{g} = K_{m,n}$. Combining this with (13), the equilibrium efforts $x_i^*(\lambda)$ belonging to component V_r of the bipartite network $\mathbf{g}$ can be represented as follows:

$$x_i^*(\lambda) = (1-\lambda)\alpha_i + \frac{\lambda}{1+\lambda}\overline{\alpha}_s + \frac{\lambda^2}{1+\lambda}\overline{\alpha}_r,$$

where $r, s = 1, 2$, $r \neq s$, while $\overline{\alpha}_r$ is defined by (F.2). When $\lambda \to 1$, we obtain:

$$\lim_{\lambda\to 1} x_i^*(\lambda) = \frac{\overline{\alpha}_1 + \overline{\alpha}_2}{2} = \frac{1}{2m}\sum_{j\in V_1}\alpha_j + \frac{1}{2n}\sum_{j\in V_2}\alpha_j. \tag{A.16}$$

Combining (C.3) with (C.6) in the Online Appendix C, it is readily verified that, when $\mathbf{g} = K_{m,n}$, we obtain:

$$\boldsymbol{\pi} = \left(\underbrace{\frac{1}{2m}, \ldots, \frac{1}{2m}}_{m \text{ times}}, \underbrace{\frac{1}{2n}, \ldots, \frac{1}{2n}}_{n \text{ times}} \right).$$

This, together with (A.16), implies (20) and completes the proof. ■

Proof of Proposition 5:

(i) Multiplying both parts of (13) by $\boldsymbol{\pi}$ from the left, we obtain:[1]

$$\boldsymbol{\pi}\mathbf{x}^* = \boldsymbol{\pi}\boldsymbol{\alpha}. \tag{A.17}$$

Differentiating both parts of (A.17) with respect to λ leads to

$$\boldsymbol{\pi}\frac{\partial \mathbf{x}^*}{\partial \lambda} = 0 \iff \sum_{j=1}^{n} \pi_j \frac{\partial x_j^*}{\partial \lambda} = 0.$$

By the Perron–Frobenius theorem, $\pi_j > 0$, for all j. Thus, for any $\lambda \in (0, 1)$, if $\frac{\partial x_i^*}{\partial \lambda} > 0$ for some i, then it has to be that $\frac{\partial x_j^*}{\partial \lambda} < 0$ for some $j \neq i$. This proves part (i).

(ii) By totally differentiating both parts of (11) with respect to λ and setting $\lambda = 0$, we obtain

$$\left.\frac{\partial x_i^*}{\partial \lambda}\right|_{\lambda=0} = \sum_{j=1}^{n} \widehat{g}_{ij}\alpha_j - \alpha_i.$$

Hence, (23) holds at $\lambda = 0$. By continuity, (23) also holds in the vicinity of $\lambda = 0$, that is,when λ is positive but not too large. This proves part (ii).

(iii) First, as implied by Proposition 2, the inequality $\sum_{j=1}^{n} \pi_j\alpha_j \leq \alpha_i$ is equivalent to $x_i^*(0) \geq x_i^*(1)$. Second, owing to (23), the inequality $\alpha_i < \sum_{j=1}^{n} \widehat{g}_{ij}\alpha_j$ is equivalent to

$$\left.\frac{\partial x_i^*}{\partial \lambda}\right|_{\lambda=0} > 0,$$

meaning that $x_i^*(\lambda)$ strictly increases in λ in the vicinity of $\lambda = 0$. Combining this with $x_i^*(0) \geq x_i^*(1)$, we conclude that $x_i^*(\lambda)$ has an interior global maximizer over $[0, 1]$, and hence, it is non-monotone in λ. This proves part (iii).

(iv) The proof of part (iv) repeats verbatim that of part (iii), up to reverting all the inequalities. ∎

Proof of Proposition 6.

Let Δ_{ij} denote the operator that maps a vector or a matrix into the difference between the values of this vector or matrix before and after adding the link ij. For example, $\Delta_{ij}\mathbf{x}^*$ is the n-dimensional column vector that captures how the equilibrium efforts change, while $\Delta_{ij}\widehat{\mathbf{G}}$ is the $(n\times n)$-matrix that captures how the row-normalized adjacency matrix $\widehat{\mathbf{G}}$ social network changes after agents i and j form a link. Also,

[1] Observe that $\boldsymbol{\pi}\widehat{\mathbf{M}} = \boldsymbol{\pi}$.

denote by $\widehat{\mathbf{G}}$ the original row-normalized adjacency matrix without the link i-j and by $\widehat{\mathbf{G}}^{[+ij]}$ the row-normalized adjacency matrix when the link i-j has been added. Similarly, we denote by $\mathbf{x}^*$ the equilibrium effort without the link i-j and by $\mathbf{x}^{[+ij]*}$ the equilibrium effort when the link i-j has been added.

Applying Δ_{ij} to both sides of the equilibrium condition (12), we obtain:

$$\Delta_{ij}\mathbf{x}^* = \lambda\widehat{\mathbf{G}}\left(\Delta_{ij}\mathbf{x}^*\right) + \lambda\left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^*. \tag{A.18}$$

Observe that

$$\begin{aligned}
\Delta_{ij}\mathbf{x}^* &= \mathbf{x}^{[+ij]*} - \mathbf{x}^* \\
&= \lambda\left(\widehat{\mathbf{G}}^{[+ij]}\mathbf{x}^{[+ij]*} - \widehat{\mathbf{G}}\mathbf{x}^*\right) \\
&= \lambda\left(\widehat{\mathbf{G}}^{[+ij]}\mathbf{x}^{[+ij]*} - \widehat{\mathbf{G}}^{[+ij]}\mathbf{x}^* + \widehat{\mathbf{G}}^{[+ij]}\mathbf{x}^* - \widehat{\mathbf{G}}\mathbf{x}^*\right) \\
&= \lambda\left(\widehat{\mathbf{G}}^{[+ij]}\left(\Delta_{ij}\mathbf{x}^*\right) + \left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^*\right).
\end{aligned}$$

Since $\widehat{\mathbf{G}}^{[+ij]} := \widehat{\mathbf{G}} + \left(\Delta_{ij}\widehat{\mathbf{G}}\right)$, we have:

$$\begin{aligned}
\Delta_{ij}\mathbf{x}^* &= \lambda\left(\widehat{\mathbf{G}}^{[+ij]}\left(\Delta_{ij}\mathbf{x}^*\right) + \left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^*\right) \\
&= \lambda\left[\left(\widehat{\mathbf{G}} + \left(\Delta_{ij}\widehat{\mathbf{G}}\right)\right)\left(\Delta_{ij}\mathbf{x}^*\right) + \left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^*\right] \\
&= \lambda\widehat{\mathbf{G}}\left(\Delta_{ij}\mathbf{x}^*\right) + \lambda\left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^* + \lambda\left(\Delta_{ij}\widehat{\mathbf{G}}\right)\left(\Delta_{ij}\mathbf{x}^*\right).
\end{aligned}$$

Denote $\widehat{\mathbf{M}}^{[+ij]} := (1-\lambda)\left(\mathbf{I} - \lambda\widehat{\mathbf{G}}^{[+ij]}\right)^{-1}$. Then,

$$\Delta_{ij}\mathbf{x}^* = \frac{\lambda}{1-\lambda}\widehat{\mathbf{M}}^{[+ij]}\left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^*. \tag{A.19}$$

Let us calculate $\left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^*$. For the sake of the presentation and without loss of generality, assume that $i = 1$ and $j = 2$. Before the link $1-2$ is created, the row-normalized adjacency matrix is equal to:

$$\widehat{\mathbf{G}} = \begin{pmatrix}
0 & 0 & \frac{g_{13}}{d_1} & \cdots & \frac{g_{1n}}{d_1} \\
0 & 0 & \frac{g_{23}}{d_2} & \cdots & \frac{g_{2n}}{d_2} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{g_{n1}}{d_n} & \frac{g_{n2}}{d_n} & \frac{g_{n3}}{d_n} & \cdots & 0
\end{pmatrix}.$$

After the link 1−2 is created, the row-normalized adjacency matrix is equal to:

$$\widehat{\mathbf{G}}^{[+12]} = \begin{pmatrix} 0 & \frac{1}{d_1+1} & \frac{g_{13}}{d_1+1} & \cdots & \frac{g_{1n}}{d_1+1} \\ \frac{1}{d_2+1} & 0 & \frac{g_{23}}{d_2+1} & \cdots & \frac{g_{2n}}{d_2+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{g_{n1}}{d_n} & \frac{g_{n2}}{d_n} & \frac{g_{n3}}{d_n} & \cdots & 0 \end{pmatrix},$$

where d_i is the degree (i.e., number of links) of agent i. Therefore,

$$\widehat{\mathbf{G}}^{[+12]} - \widehat{\mathbf{G}} := \Delta_{12}\widehat{\mathbf{G}} = \begin{pmatrix} 0 & \frac{1}{d_1+1} & \frac{-g_{13}}{d_1(d_1+1)} & \cdots & \cdots & \frac{-g_{1n}}{d_1(d_1+1)} \\ \frac{1}{d_2+1} & 0 & \frac{-g_{23}}{d_2(d_2+1)} & \cdots & \cdots & \frac{-g_{2n}}{d_2(d_2+1)} \\ 0 & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Thus,

$$\left(\Delta_{12}\widehat{\mathbf{G}}\right)\mathbf{x}^* = \begin{pmatrix} Z_1 \\ Z_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where $Z_1 := \left(\left(\Delta_{12}\widehat{\mathbf{G}}\right)\mathbf{x}^*\right)_1$ and $Z_2 := \left(\left(\Delta_{12}\widehat{\mathbf{G}}\right)\mathbf{x}^*\right)_2$, where the subscripts 1 and 2 indicate the first and second element of the vector $\left(\Delta_{12}\widehat{\mathbf{G}}\right)\mathbf{x}^*$. We have:

$$\begin{aligned} Z_1 &= \frac{1}{d_1+1}x_2^* - \frac{g_{13}}{d_1(d_1+1)}x_3^* - \cdots - \frac{g_{1n}}{d_1(d_1+1)}x_n^* \\ &= \frac{1}{d_1+1}x_2^* - \frac{1}{d_1+1}\sum_{j=3}^{n}\frac{g_{1j}}{d_1}x_j^* \\ &= \frac{1}{d_1+1}x_2^* - \frac{1}{d_1+1}\sum_{j=3}^{n}\widehat{g}_{1j}x_j^* \\ &= \frac{(x_2^* - \overline{x}_1^*)}{(d_1+1)}. \end{aligned}$$

In a similar way, it is straighforward to show that:

$$Z_2 = \frac{(x_1^* - \overline{x}_2^*)}{(d_2+1)}.$$

The same reasoning can be applied for the creation of any $i{-}j$ link, so that

$$Z_i := \left(\left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^*\right)_i = \frac{(x_j^* - \overline{x}_i^*)}{(d_i+1)} \qquad \text{and} \qquad Z_j := \left(\left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^*\right)_j = \frac{(x_i^* - \overline{x}_j^*)}{(d_j+1)}.$$

Therefore,

$$\left(\Delta_{ij}\widehat{\mathbf{G}}\right)\mathbf{x}^* = \frac{(x_j^* - \overline{x}_i^*)}{(d_i+1)}\mathbf{e}_i + \frac{(x_i^* - \overline{x}_j^*)}{(d_j+1)}\mathbf{e}_j,$$

where $\mathbf{e}_i$ is the n-dimensional vector whose ith component is equal to 1 while all the others are equal to zero. By plugging this expression into (A.19), we obtain:

$$\Delta_{ij}\mathbf{x}^* = \frac{\lambda}{(1-\lambda)}\left(\frac{(x_j^* - \overline{x}_i^*)}{(d_i+1)}\widehat{\mathbf{M}}^{[+ij]}\,\mathbf{e}_i + \frac{(x_i^* - \overline{x}_j^*)}{(d_j+1)}\widehat{\mathbf{M}}^{[+ij]}\,\mathbf{e}_j\right). \tag{A.20}$$

Because $\widehat{\mathbf{M}}^{[+ij]}$ is a strictly positive matrix, the vectors $\widehat{\mathbf{M}}^{[+ij]}\,\mathbf{e}_i$ and $\widehat{\mathbf{M}}^{[+ij]}\,\mathbf{e}_j$ are both strictly positive. Hence, the vector $\Delta_{ij}\mathbf{x}^*$ of changes in equilibrium effort levels is:

(*i*) strictly positive, if $x_j^* > \overline{x}_i^*$ and $x_i^* > \overline{x}_j^*$;

(*ii*) strictly negative, if $x_j^* < \overline{x}_i^*$ and $x_i^* < \overline{x}_j^*$;

(*iii*) may involve both positive and negative components, otherwise.

It remains to prove that the inequalities $x_j^* > \overline{x}_i^*$ and $x_i^* > \overline{x}_j^*$ are equivalent to, respectively, (26) and (27). Using (11), we find that:

$$x_i^* > \overline{x}_j^* \iff x_j^* < (1-\lambda)\alpha_j + \lambda x_i^*, \qquad x_j^* > \overline{x}_i^* \iff x_i^* < (1-\lambda)\alpha_i + \lambda x_j^*.$$

Combining this with (A.3), we find after simplifications that $x_j^* > \overline{x}_i^*$ is equivalent to condition (26), while $x_i^* > \overline{x}_j^*$ is equivalent to condition (27). This completes the proof. ▮

Proof of Proposition 7.

Using (6), it is readily verified that, for any $\lambda \in [0,1)$ the welfare functional (28) can be written as

$$\mathcal{W} = \boldsymbol{\alpha}^T\mathbf{x} - \frac{1}{2}\mathbf{x}^T\,\mathbf{H}(\lambda)\,\mathbf{x}, \tag{A.21}$$

where the Hessian matrix $\mathbf{H}(\lambda)$ of the welfare functional is given by

$$\mathbf{H}(\lambda) := \mathbf{I} + \frac{\lambda}{1-\lambda}\left(\mathbf{I} - \widehat{\mathbf{G}}\right)^T\left(\mathbf{I} - \widehat{\mathbf{G}}\right). \tag{A.22}$$

The following lemma summarizes the properties of $\mathbf{H}(\lambda)$.

Lemma A4

(*i*) *For any* $\lambda \in [0,1)$*, the matrix* $\mathbf{H}(\lambda)$ *is positive definite.*

(*ii*) *For any* $\lambda \in (0,1)$*,* $\mathbf{H}(\lambda)$ *has a prime eigenvalue equal to one, the corresponding eigenspace being the span of* $\mathbf{1}$*; other* $n-1$ *eigenvalues of* $\mathbf{H}(\lambda)$ *are strictly greater than one, and they unboundedly grow as* $\lambda \to 1$*.*

(*iii*) *We have*

$$\lim_{\lambda \to 1} \mathbf{H}^{-1}(\lambda) = \frac{1}{n} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \tag{A.23}$$

where $\mathbf{H}^{-1}(\lambda)$ *is the inverse to* $\mathbf{H}(\lambda)$*.*

Proof of Lemma A4.

(*i*) For any $\mathbf{x} \in \mathbb{R}^n$, we have

$$\mathbf{x}^T \mathbf{H}(\lambda) \mathbf{x} = \|\mathbf{x}\|^2 + \frac{\lambda}{1-\lambda} \|\mathbf{x} - \widehat{\mathbf{G}}\mathbf{x}\|^2 \geq \|\mathbf{x}\|^2, \tag{A.24}$$

where $\|\cdot\|$ stands for the Euclidean norm. Whenever $\mathbf{x} \neq \mathbf{0}$, we have $\|\mathbf{x}\|^2 > 0$. This proves positive definiteness of $\mathbf{H}(\lambda)$.

(*ii*) Observe that $\mathbf{H}(\lambda)$ is a symmetric matrix:

$$\mathbf{H}^T(\lambda) = \mathbf{H}(\lambda).$$

Hence, all its eigenvalues are real. Furthermore, as shown in (*i*), $\mathbf{H}(\lambda)$ is positive definite, and hence, all its eigenvalues are strictly positive. It is well known (Horn and Johnson, 1985, Ch. 1, p. 34) that the minimum eigenvalue of a symmetric matrix is given by

$$\min_{\mathbf{x} \in \mathbb{R}_n^+} \left\{ \mathbf{x}^T \mathbf{H}(\lambda) \mathbf{x} : \|\mathbf{x}\|^2 = 1 \right\},$$

while the corresponding eigenspace is the span of all the minimizers. When $\|\mathbf{x}\|^2 = 1$, (A.24) implies that $\mathbf{x}^T \mathbf{H}(\lambda) \mathbf{x} \geq 1$. Furthermore, when $\lambda \in (0,1)$ we have

$$\mathbf{x}^T \mathbf{H}(\lambda) \mathbf{x} = 1 \quad \Longleftrightarrow \quad \|\mathbf{x} - \widehat{\mathbf{G}}\mathbf{x}\|^2 = 0 \quad \Longleftrightarrow \quad \mathbf{x} = \widehat{\mathbf{G}}\mathbf{x} \quad \Longleftrightarrow \quad \mathbf{x} = \frac{\mathbf{1}}{\sqrt{n}},$$

where the last equivalence follows from the fact that $\widehat{\mathbf{G}}$ is irreducible, and hence, $\mathbf{1}$ is the only eigenvector (up to a scalar multiple) of $\widehat{\mathbf{G}}$ corresponding to the unitary eigenvalue. All the other eigenvalues of $\widehat{\mathbf{G}}$ are bounded from below by

$$1 + \frac{\lambda}{1-\lambda}\delta,$$

where δ is the smallest strictly positive eigenvalue of $\left(\mathbf{I} - \widehat{\mathbf{G}}\right)^T \left(\mathbf{I} - \widehat{\mathbf{G}}\right)$, given by:

$$\delta := \min_{\mathbf{x}\in\mathbb{R}_n^+} \left\{\mathbf{x}^T\mathbf{H}(\lambda)\mathbf{x} : \|\mathbf{x}\|^2 = 1,\ \mathbf{1}^T\mathbf{x} = 0\right\}.$$

Because $\delta > 0$, we clearly have

$$\lim_{\lambda\to 1}\left(1 + \frac{\lambda}{1-\lambda}\delta\right) = \infty.$$

This proves part (ii).

(iii) Because $\mathbf{H}(\lambda)$ is positive definite for any $\lambda \in [0,1)$, we have

$$\det\left[\mathbf{H}(\lambda)\right] > 0,$$

and hence, $\mathbf{H}^{-1}(\lambda)$ is well defined for any $\lambda \in [0,1)$. Because $\mathbf{H}(\lambda)$ is symmetric, it can be diagonalized. Finally, because the eigenvalues (eigenvectors) of the inverse matrix are the reciprocals of (coincide with) those of the original matrix, part (ii) of the lemma implies that the largest eigenvalue of $\mathbf{H}^{-1}(\lambda)$ equals one, the corresponding eigenspace being the span of $\mathbf{1}$, while other $n-1$ eigenvalues of $\mathbf{H}^{-1}(\lambda)$ converge to zero as $\lambda \to 1$. Putting all these considerations together, and denoting by $\mathbf{s}_i$ the ith column eigenvector of $\mathbf{H}^{-1}(\lambda)$, chosen so that $\mathbf{s}_i^T\mathbf{s}_i = 1$ for all $i = 1, 2, \ldots, n$, we obtain

$$\lim_{\lambda\to 1}\mathbf{H}^{-1}(\lambda) = \left(\frac{1}{\sqrt{n}}\mathbf{1}, \mathbf{s}_2, \ldots, \mathbf{s}_n\right)^T \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \left(\frac{1}{\sqrt{n}}\mathbf{1}, \mathbf{s}_2, \ldots, \mathbf{s}_n\right).$$

Combining this with the fact that the eigenvectors of a symmetric matrix form an orthonormal basis, i.e., $\mathbf{s}_i^T\mathbf{s}_j = \delta_{ij}$ for all $i, j = 1, 2, \ldots, n$, where δ_{ij} is defined by (A.11), we get (A.23). This completes the proof of Lemma A4. ■

We now proceed with the proof of Proposition 7.

Proof of part (i): Using (A.21), the social planner's problem can be written as follows:

$$\max_{\mathbf{x}\in\mathbb{R}^n_+}\left[\boldsymbol{\alpha}^T\mathbf{x}-\frac{1}{2}\mathbf{x}^T\,\mathbf{H}(\lambda)\,\mathbf{x}\right],$$

where $\mathbf{H}(\lambda)$ is defined by (A.22). The FOC to this problem is given by

$$\mathbf{H}(\lambda)\,\mathbf{x}=\boldsymbol{\alpha}. \tag{A.25}$$

Combining (A.25) with (A.22), it is readily verified that (A.25) is equivalent to (30). Clearly, the solution to (A.25)—or, equivalently, to (30)—is unique and is given by

$$\mathbf{x}^O=\mathbf{H}^{-1}(\lambda)\boldsymbol{\alpha}. \tag{A.26}$$

To finish the proof of part (i), we need to verify that, when the FOC holds, the second-order condition also holds. It suffices to prove that the welfare function $\mathcal{W}$ is strictly concave in $\mathbf{x}$. As observed from (A.21), the Hessian matrix of $\mathcal{W}$ equals $-\frac{1}{2}\mathbf{H}(\lambda)$, which is negative definite by part (i) of Lemma A4. Hence, $\mathcal{W}$ is strictly concave.

We also need to show that $\mathbf{x}^O$ is an interior optimum. However, this is not an issue when the mean productivity is sufficiently large. To make this statement precise, denote by μ_α and σ_α the mean and standard deviation of the individual productivity, respectively:

$$\mu_\alpha:=\frac{1}{n}\sum_{j=1}^{n}\alpha_j,\qquad\qquad \sigma_\alpha:=\sqrt{\frac{1}{n}\sum_{j=1}^{n}(\alpha_j-\mu_\alpha)^2}.$$

Lemma A5 *If $\mu_\alpha>\sqrt{n}\sigma_\alpha$, then $\mathbf{x}^O$ is the interior solution to the social planner's problem.*

Proof of Lemma A5. For all $i=1,2,\ldots,n$, define

$$\varepsilon_i:=\frac{\alpha_i-\mu_\alpha}{\sigma_\alpha\sqrt{n}}.$$

ε_i is the deviation of agent i's individual productivity from the mean, rescaled so that $\|\boldsymbol{\varepsilon}\|=1$, where $\boldsymbol{\varepsilon}:=(\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n)^T$. Clearly, α_i can be decomposed as follows:

$$\alpha_i=\mu_\alpha+\sigma_\alpha\sqrt{n}\,\varepsilon_i,\qquad \frac{1}{n}\sum_{j=1}^{n}\varepsilon_j=0,$$

or, in vector-matrix form,

$$\boldsymbol{\alpha} = \mu_\alpha \mathbf{1} + \sigma_\alpha \sqrt{n}\, \boldsymbol{\varepsilon}. \tag{A.27}$$

Plugging (A.27) into (A.26), we obtain

$$\mathbf{x}^O = \mu_\alpha \mathbf{1} + \sigma_\alpha \sqrt{n}\, \mathbf{H}^{-1}(\lambda)\, \boldsymbol{\varepsilon}.$$

For each $i = 1, 2, \ldots, n$, we have

$$x_i^O \geq \mu_\alpha - \sigma_\alpha \sqrt{n} \max_{j=1,2,\ldots,n} \left| \left(\mathbf{H}^{-1}(\lambda) \boldsymbol{\varepsilon} \right)_j \right| \geq \mu_\alpha - \sigma_\alpha \sqrt{n}\, \| \mathbf{H}^{-1}(\lambda) \boldsymbol{\varepsilon} \|, \tag{A.28}$$

where the second inequality follows from the standard triangle inequality. As implied by part (*ii*) of Lemma A4, the spectral radius of $\mathbf{H}^{-1}(\lambda)$ equals 1. Hence, we obtain

$$\| \mathbf{H}^{-1}(\lambda) \boldsymbol{\varepsilon} \| \leq \| \boldsymbol{\varepsilon} \| = 1.$$

Combining this with (A.28) yields

$$x_i^O \geq \mu_\alpha - \sigma_\alpha \sqrt{n} > 0$$

for all $i = 1, 2, \ldots, n$. This completes the proof of Lemma A5. ■

We have shown that $\mathbf{x}^O$ is a unique global maximizer of the welfare functional $\mathcal{W}$, which is interior provided that the mean productivity is sufficiently high. This proves part (*i*). ■

Proof of part (*ii*): Comparing (30) with (13), we find that $\mathbf{x}^* = \mathbf{x}^O$ if and only if the following condition holds:

$$\widehat{\mathbf{G}}^T \left(\mathbf{I} - \widehat{\mathbf{G}} \right) \mathbf{x}^* = 0. \tag{A.29}$$

Using (13), this is equivalent to

$$\widehat{\mathbf{G}}^T \left(\mathbf{I} - \widehat{\mathbf{G}} \right) \widehat{\mathbf{M}} \boldsymbol{\alpha} = \mathbf{0}.$$

This proves part (*ii*). ■

Proof of part (*iii*): Multiplying both parts of (30) by $\mathbf{1}^T$ from the left leads to

$$\mathbf{1}^T \mathbf{x}^O = (1 - \lambda)\, \mathbf{1}^T \boldsymbol{\alpha} + \lambda \mathbf{1}^T \left(\widehat{\mathbf{G}} + \widehat{\mathbf{G}}^T - \widehat{\mathbf{G}}^T \widehat{\mathbf{G}} \right) \mathbf{x}^O.$$

Because $\widehat{\mathbf{G}}$ is row-normalized, we have $\mathbf{1}^T \widehat{\mathbf{G}}^T = \mathbf{1}^T$. Using this, and simplifying it, we obtain

$$\mathbf{1}^T \mathbf{x}^O = (1 - \lambda)\, \mathbf{1}^T \boldsymbol{\alpha} + \lambda \mathbf{1}^T \mathbf{x}^O \quad \Longrightarrow \quad \mathbf{1}^T \mathbf{x}^O = \mathbf{1}^T \boldsymbol{\alpha}.$$

This completes the proof. ▮

Proof of Proposition 8

We focus on the case in which $\sum_{j=1}^{n} \pi_j \alpha_j < \frac{1}{n} \sum_{j=1}^{n} \alpha_j$. For the other case, the proof goes along the same lines.

Combining (20) with (33), we find that

$$\sum_{j=1}^{n} \pi_j \alpha_j < \frac{1}{n} \sum_{j=1}^{n} \alpha_j \qquad \Longleftrightarrow \qquad \lim_{\lambda \to 1} x_i^*(\lambda) < \lim_{\lambda \to 1} x_i^O(\lambda)$$

for all $i = 1, 2, \ldots, n$. Because $x_i^*(\lambda)$ and $x_i^O(\lambda)$ are all continuous in λ (see proof of part (i) above), the inequalities $x_i^*(\lambda) < x_i^O(\lambda)$ must keep holding when λ is slightly below 1. Setting

$$\underline{\lambda} := \max \left\{ \lambda > 0 \;\middle|\; \min_{i=1,2,\ldots,n} \{ x_i^O(\lambda) - x_i^*(\lambda) \} = 0 \right\}$$

completes the proof. ▮

Proof of Remark 3

Taking the limit as $\lambda \to 1$ on both sides of (A.26), and using part (ii) of Lemma A4, we obtain (33). This proves the result. ■

Proof of Proposition 9 Omitted. ▮

B Local-average versus local-aggregate model: A more detailed comparison

In the local-average model, agent i's payoff, $i = 1, 2, \ldots, n$, is given by (3), which becomes after some rearrangement:

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{1+\theta}{2} x_i^2 + \frac{\theta}{d_i} \sum_{j=1}^{n} g_{ij} x_i x_j - \frac{\theta}{2d_i^2} \left(\sum_{j=1}^{n} g_{ij} x_j \right)^2, \tag{B.1}$$

where g_{ij} is the ijth entry of the (non-row-normalized) adjacency matrix. Note that, although the last term, $\frac{\theta}{2d_i^2} \left(\sum_{j=1}^{n} g_{ij} x_j \right)^2$, in (B.1) is immaterial for equilibrium analysis, it affects the welfare results.

Equation (B.1) allows us to have a convenient comparison between the two models. To see this, consider agent i's payoff in the local-aggregate model:

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{1}{2} x_i^2 + \theta \sum_{j=1}^{n} g_{ij} x_i x_j. \tag{B.2}$$

From a formal viewpoint, both models are special cases of a unifying model in which agent i's payoff is given by:

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{\beta_i}{2} x_i^2 + \gamma_i \sum_{j=1}^{n} g_{ij} x_i x_j - \frac{\delta_i}{2} \left(\sum_{j=1}^{n} g_{ij} x_j \right)^2, \tag{B.3}$$

where α_i, β_i, γ_i, and δ_i, $i = 1, 2, \ldots, n$, are all agent-specific. Indeed, we obtain the local aggregate model if we set in (B.3):

$$\begin{aligned} &\beta_1 = \beta_2 = \ldots = \beta_n = 1, \\ &\gamma_1 = \gamma_2 = \ldots = \gamma_n = \theta < \tfrac{1}{\rho(\mathbf{G})}, \\ &\delta_1 = \delta_2 = \ldots = \delta_n = 0, \end{aligned}$$

For (B.3) to become the local-average model, the following constraints must hold for all $i, j = 1, 2, \ldots, n$:

$$\begin{aligned} &\beta_i - d_i \gamma_i = 1, \qquad && d_i \gamma_i = d_j \gamma_j = \ldots = d_j \gamma_n = \theta, \\ &d_i \delta_i - \gamma_i = 0, \qquad && d_i^2 \delta_i = d_j^2 \delta_j = \ldots = d_j^2 \delta_n, \end{aligned}$$

where d_i denotes agent i's degree.

As it is formulated, model (B.3) is not very informative in terms of both theoretical insights (because it hardly yields any clear-cut predictions) and in terms of empirical applications (because it has too many parameters to be estimated). The only way model (B.3) can be put to work is by imposing economically meaningful restrictions on the coefficients in (B.3).

Indeed, imposing different sets of meaningful restrictions on the coefficients in (B.3) can lead to models telling different stories, displaying different properties (such as comparative statics, welfare and network formation), and giving different policy recommendations. In other words, the local-aggregate model essentially fails to approximate a number of essential qualitative effects that arise in the local-average model, despite the fact that both can be obtained from (B.3) by imposing certain constraints on parameters. This should not come as surprise as the two models tell different stories.

Assume, for example, that the number of direct neighbors of agent i doubles, while individual effort of each neighbor remains roughly the same. Then, the local-aggregate model predicts an increase in agent i's effort by 2θ (which might be a considerable amount if θ is not small), whereas the local-average model predicts a zero reaction of agent i. Thus, while the former puts forwards spillover effects, the latter highlights the importance of conformity, i.e., blending with one's social environment. Therefore, it is quite natural that the two models differ so much in terms of outcomes.

C A probabilistic interpretation of the model

It should be clear from Proposition 1 and Lemma 1 that the change in $\widehat{m}_{ij}$ with respect to the key parameters of the model is key for understanding equilibrium behavior. However, equation (17) that defines $\widehat{m}_{ij}$ is not easy to interpret. To gain more intuition, we reformulate the local-average model in *probabilistic* terms. In this interpretation, $\widehat{\mathbf{G}} = [\widehat{g}_{ij}]$ is now a *transition probability matrix* of a Markov chain with n finite states, where each state is the location of each agent in the network. Since there are n agents, there are n states.

A discrete-time Markov chain is a sequence of random variables $Z_1, Z_2, Z_3, \ldots$, with the Markov property that the probability of moving to the next state depends only on the present state. We have

$$\widehat{g}_{ij} = \mathbb{P}\left\{Z_{k+1} = j \mid Z_k = i\right\}$$

Consider individual i who chooses effort X_i. With probability $1-\lambda$, she chooses to exert α_i units of effort while, with probability λ, she mimics the behavior of one of her neighbors (or direct links), say individual j, which is given by α_j. Then, with probability λ, agent i adopts this behavior (chooses α_j) while, with probability $1-\lambda$, she chooses to talk to one of j's neighbors, and so forth. In this interpretation, λ is still a measure of conformity but helps each individual to collect information about the productivity of other agents.

To formalize this process, denote by X_i the effort of agent i, which is a *random variable* defined by

$$\mathbb{P}\left\{X_i = \alpha_j\right\} = (1-\lambda)\sum_{k=0}^{\infty} \lambda^k \widehat{g}_{ij}^{[k]}.$$

By combining this equation with (17), we find:

$$\mathbb{P}\left\{X_i = \alpha_j\right\} = \widehat{m}_{ij}(\lambda).$$

Thus, $\widehat{m}_{ij}(\lambda)$ is *the probability that, starting from i, the random walk terminates at j.* In other words, $\widehat{m}_{ij}(\lambda)$ is the probability that agent i ends up mimicking the behavior of agent j. The expected value $\mathbb{E}\left[X_i\right]$ of agent i's effort is given by:

$$\mathbb{E}\left[X_i\right] = \sum_{j=1}^{n} \widehat{m}_{ij}\alpha_j.$$

In matrix form, we have:

$$\mathbb{E}\left[\mathbf{X}\right] = \widehat{\mathbf{M}}\boldsymbol{\alpha} = (1-\lambda)\sum_{k=0}^{\infty} \lambda^k \widehat{\mathbf{G}}^k \boldsymbol{\alpha}, \tag{C.1}$$

where $\mathbf{X} := (X_1, X_2, \ldots, X_n)^T$ is the vector of efforts in the probabilistic model. By comparing (C.1) and (13), we see that:

$$\mathbf{x}^* = \mathbb{E}\left[\mathbf{X}\right], \tag{C.2}$$

where $\mathbf{x}^*$ is the Nash equilibrium in the local-average model. Thus, the two models are *observationally equivalent.* This is quite remarkable since, in the former, agents are perfectly rational and solve a game with peer effects, while, in the latter, agents make decisions stochastically, a little bit like in models of evolutionary game theory, in which agents act like robots or automas and then converge to some behavior. We show that, on average, the two types of behavior (Nash equilibrium and stochastic decision) lead to exactly the same outcomes.

The interpretation of the equilibrium in terms of Markov chain is relatively familiar in the network literature. In particular, in the learning literature (see, e.g., DeGroot, 1974; DeMarzo et al., 2003; Golub and Jackson, 2012; for recent overviews, see Mobius and Rosenblat, 2014; Golub and Sadler, 2016), the social structure of a society is described by a weighted directed network, which captures the pattern of communication and transmission of information between agents. Agents have beliefs about some common state of the world and, at each date, agents communicate with their neighbors in the social network and update their beliefs. An agent's updated belief at time t is the (weighted) average of his or her neighbors' beliefs from the previous period $t-1$. This defines a Markov chain in which the states are the beliefs each agent has on the state of the world. In our model, the interpretation is different since there is no time variation and each state is the location of each agent in the network. Moreover, we believe that we are the first to show the equivalence between the Nash equilibrium in effort of a network static game and the expected effort of a Markov chain.

Remember that, in the probabilistic model, $\widehat{\mathbf{G}} = [\widehat{g}_{ij}]$ is the transition probability matrix of a Markov chain with n finite states. It should then be clear that, given that network $\mathbf{g}$ is connected, this Markov chain is *irreducible.* Hence, by the Perron–Frobenius theorem, the stationary distribution $\boldsymbol{\pi}$ can be uniquely defined as the *left eigenvector* of $\widehat{\mathbf{G}}$ associated with the unitary eigenvalue, that is

$$\boldsymbol{\pi}\widehat{\mathbf{G}} = \boldsymbol{\pi}, \tag{C.3}$$

where $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_n)$ is a $(1 \times n)$ vector, while $\pi_i > 0$ for all $i = 1, 2, \ldots, n$. As pointed out by DeMarzo et al. (2003) and Golub and Jackson (2010), because our network is undirected, we have

$$\pi_i = \frac{d_i}{\sum_{j=1}^{n} d_j}, \quad \text{for all } i = 1, 2, \ldots n,$$

where $d_i := \sum_{j=1}^{n} g_{ij}$ is the degree of i. In other words, π_i is the *relative degree* of individual i.

Definition. *The Markov chain with transition matrix* $\widehat{\mathbf{G}}$ *is ergodic if and only if the following condition is satisfied*:

$$\lim_{k\to\infty} \widehat{\mathbf{G}}^k = \begin{pmatrix} \boldsymbol{\pi} \\ \boldsymbol{\pi} \\ \vdots \\ \boldsymbol{\pi} \end{pmatrix}. \tag{C.4}$$

In general, an irreducible Markov chain need not be ergodic. However, as stated by the following lemma, in our model, most (although not all) network structures give rise to ergodic Markov chains.

Lemma C6 *The Markov chain with transition matrix* $\widehat{\mathbf{G}}$ *is not ergodic if and only if* $\mathbf{g} = K_{m,n}$*, that is,* $\mathbf{g}$ *is a complete bipartite graph, with a partition* $V_1 \cup V_2$ *where* $|V_1| = m$*,* $|V_2| = n$*. A star-shaped network is a special case with* $m = 1$*.*

Proof of Lemma C6. If an irreducible Markov chain is non-ergodic, it must be *periodic*, meaning that, up to a simultaneous permutation of rows and columns, its transition matrix $\widehat{\mathbf{G}}$ can be represented as follows (see Horn and Johnson, 1985, Ch. 8, p. 512.):

$$\widehat{\mathbf{G}} = \begin{pmatrix} \mathbf{0} & \mathbf{A}_{12} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_{23} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_{K-1,K} \\ \mathbf{A}_{K1} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}, \tag{C.5}$$

where $K > 1$ is an integer that shows the period of the Markov chain, $\mathbf{A}_{ij}$ are matrixes with positive entries, while $\mathbf{0}$ is a zero matrix of appropriate dimension. Combining this with the fact that

$$\widehat{g}_{ij} > 0 \iff \widehat{g}_{ji} > 0,$$

we find that the only case when $\widehat{\mathbf{G}}$ has a structure satisfying (C.5) is when $\mathbf{g} = K_{m,n}$. Indeed, in this case, the row-normalized adjacency matrix $\widehat{\mathbf{G}}$ is given by

$$\widehat{\mathbf{G}} = \begin{pmatrix} \mathbf{0}_{m\times m} & \frac{1}{n}\mathbf{1}_{m\times n} \\ \frac{1}{m}\mathbf{1}_{n\times m} & \mathbf{0}_{n\times n} \end{pmatrix}, \tag{C.6}$$

where $\mathbf{0}_{p\times q}$ and $\mathbf{1}_{p\times q}$ stand for $(p \times q)$-matrixes of zeros and ones, respectively. This completes the proof. ∎

This lemma states that, for most networks, the Markov chain is ergodic, apart from when we have a complete bipartite graph, since, in that case, they are cycles that prevent the Markov chain from being aperiodic.

D Linear-in-means model and heterogeneity: An example

Consider a star-shaped network with $n = 3$ agents where agent 1 is the star and in which the individual productivities are given by: $\alpha_1 = 1 + 2t$, $\alpha_2 = 1 - t$ and $\alpha_3 = 1 - t$, where $t \in (-1/2, 1)$.[1] Clearly, the mean productivity, $\mu_{\boldsymbol{\alpha}}$, in the network is independent of t and equal to $\mu_{\boldsymbol{\alpha}} := (\alpha_1 + \alpha_2 + \alpha_3)/3 = 1$. Furthermore, it is readily verified that the variance, $\sigma^2_{\boldsymbol{\alpha}}$, of productivity across agents is given by

$$\sigma^2_{\boldsymbol{\alpha}} := \frac{(\alpha_1 - \mu_{\boldsymbol{\alpha}})^2 + (\alpha_2 - \mu_{\boldsymbol{\alpha}})^2 + (\alpha_3 - \mu_{\boldsymbol{\alpha}})^2}{3} = 2t^2.$$

When $t = 0$, there is no heterogeneity in productivity, as $\alpha_i = 1$ for each agent $i = 1, 2, 3$. As t increases in absolute value (no matter in which direction), a mean-preserving spread in productivity occurs, as it does not affect the average but increases the variance of the productivity across agents.

Using (13), the Nash equilibrium in efforts is given by:

$$x_1^* = 1 + \frac{(2-\lambda)}{1+\lambda}\, t, \qquad x_2^* = 1 - \frac{(1-2\lambda)}{1+\lambda}\, t, \qquad x_3^* = 1 - \frac{(1-2\lambda)}{1+\lambda}\, t. \tag{D.1}$$

While the star agent's effort level, x_1^*, always increases in t, the effort exerted by peripheral agents, x_2^* and x_3^*, increases (decreases) when the taste λ for conformity is above (below) 0.5.

We can relate the variance of equilibrium efforts, $\sigma^2_{\mathbf{x}}$, to the variance of productivities, $\sigma^2_{\boldsymbol{\alpha}}$. Indeed, it is readily verified that:

$$\sigma^2_{\mathbf{x}} = 2\left(\frac{1-\lambda}{1+\lambda}\right)^2 t^2 = \left(\frac{1-\lambda}{1+\lambda}\right)^2 \sigma^2_{\boldsymbol{\alpha}}.$$

There is clearly a positive relationship between $\sigma^2_{\mathbf{x}}$ and $\sigma^2_{\boldsymbol{\alpha}}$. We have:

$$\left(\frac{1-\lambda}{1+\lambda}\right)^2 \gtreqless 1 \quad \Longleftrightarrow \quad \lambda \lesseqgtr 0.$$

This implies that our game *dampens* heterogeneity in productivities if agents are *conformists* ($\lambda > 0$) while it *magnifies* heterogeneity in productivities if agents are *anti-conformists* ($\lambda < 0$).[2]

[1] This domain is chosen for all individual productivities to remain positive.

[2] We study our game with *anti-conformism* agents ($\lambda < 0$) in Section 5.3 and in the Online Appendix G.3.

Using (D.1), the aggregate effort is given by

$$x_1^* + x_2^* + x_3^* = 3 + \frac{3\lambda}{1+\lambda}\,t,$$

which always increases in t for any $\lambda > 0$. Figure D1 illustrates these results by contrasting the case in which $\lambda = 0.25 < 0.5$ and when $\lambda = 0.75 > 0.5$.

As seen from Figure D1, the effort of the central agent, $i = 1$, always increases with t, and so does the aggregate effort. However, the effort of the periphery agents, $i = 2, 3$, decrease with t when conformity is relatively low (i.e., when $\lambda = 0.25$). More surprisingly, when conformity is relatively high (i.e., when $\lambda = 0.75$), the efforts of the periphery agents increase with t, even though their individual productivity decreases with t. This is because, when periphery agents are very conformist, they care a lot about their social norm, which is equal to the effort of the central agent. This gives rise to a positive indirect effect of an increase in t, which is the peer-effect generated by the star agent (whose productivity increases with t), which dominates the negative direct effect of the reduction of own productivity. As a result, the net effect of t on all individual efforts is positive. Therefore, the increase in aggregate effort is steeper when agents are more conformist ($\lambda = 0.75$) than when they are less conformist ($\lambda = 0.25$).

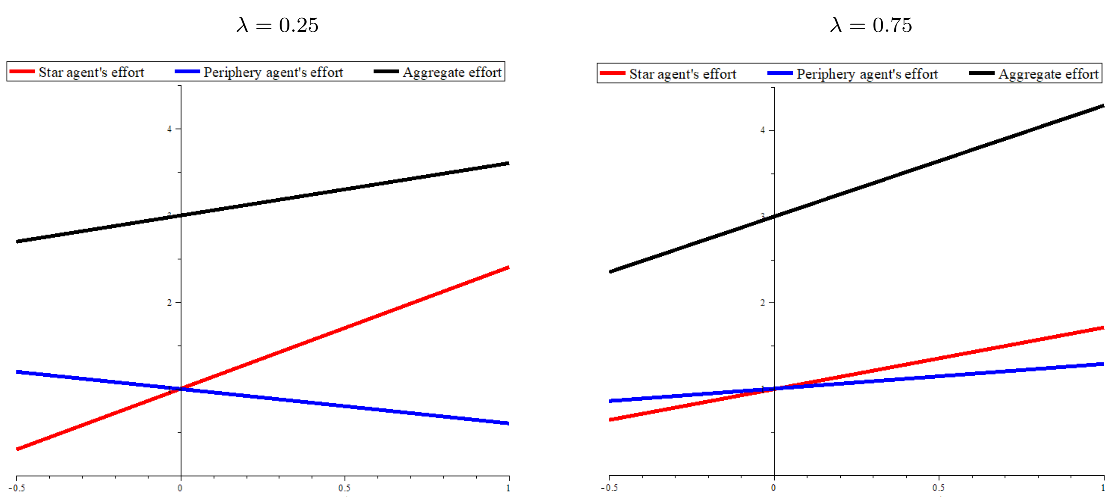

Figure D1: Impact of a mean-preserving spread of productivity on individual and aggregate effort in a star network

E Comparative statics exercises of λ: Additional results and examples

E.1 Impact of conformism on aggregate effort: General results

In the main text, we study the impact of λ on *individual effort*. Here, we study the impact of λ on *aggregate effort*.

Corollary E.1 (Non-monotonicity of aggregate effort in conformism)

(i) *Assume that* $\text{Corr}(\boldsymbol{\pi}, \boldsymbol{\alpha}) > 0$ *and that and agents are, on average, more productive than their direct neighbors. Then, the aggregate effort in the network varies non-monotonically with* λ *and has an interior global minimum in* λ*.*

(ii) *Assume that* $\text{Corr}(\boldsymbol{\pi}, \boldsymbol{\alpha}) < 0$ *and that agents are, on average, less productive than their direct neighbors. Then, the aggregate effort in the network varies non-monotonically with* λ *and has an interior global maximum in* λ*.*

Proof of Corollary E.1

(i) First, by totally differentiating both parts of (11) with respect to λ and setting $\lambda = 0$, we obtain after summation across all $i = 1, 2, \ldots, n$:

$$\frac{\partial}{\partial \lambda}\left(\sum_{i=1}^{n} x_i^*\right)\Bigg|_{\lambda=0} = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} \widehat{g}_{ij}\alpha_j - \alpha_i\right).$$

Hence, when $\lambda = 0$, we have

$$\text{sign}\left\{\frac{\partial}{\partial \lambda}\sum_{i=1}^{n} x_i^*\right\} = \text{sign}\left\{\frac{1}{n}\sum_{i,j=1}^{n} \widehat{g}_{ij}\alpha_j - \frac{1}{n}\sum_{i=1}^{n}\alpha_i\right\}. \tag{E.1}$$

That agents are, on average, more productive than their neighbors means that the right-hand side of (E.1) is negative. Hence, the aggregate effort decreases with λ in the vicinity of $\lambda = 0$.

Second, using the standard definition of the correlation coefficient $\text{Corr}(\cdot\,,\cdot)$, we have

$$\text{sign}\left\{\text{Corr}\left(\boldsymbol{\pi}, \boldsymbol{\alpha}\right)\right\} = \text{sign}\left\{\sum_{j=1}^{n} \pi_j\alpha_j - \frac{1}{n}\sum_{j=1}^{n}\alpha_j\right\}.$$

Combining this with part (*ii*) of Proposition 4 yields

$$\mathrm{Corr}(\boldsymbol{\pi}, \boldsymbol{\alpha}) > 0 \quad \Longleftrightarrow \quad \sum_{j=1}^{n} x_j^*(1) > \sum_{j=1}^{n} x_j^*(0),$$

that is, aggregate effort under total conformism ($\lambda = 1$) is higher than that under pure individualism ($\lambda = 0$). Combining this with the fact that the aggregate effort decreases when λ is small, we conclude that $\sum_{i=1}^{n} x_i^*(\lambda)$ is non-monotone λ and has an interior global minimum. This proves part (*i*).

(*ii*) The proof of part (*ii*) repeats verbatim that of part (*i*), up to reverting all the inequalities. ▮

E.2 Impact of conformism on aggregate effort: Regular networks

Let us study regular networks. Recall that a network is regular if each agent has the same number of neighbors. Specifically, a network is regular of valency r, where $r < n$ is a positive integer, if each individual has exactly r neighbors. For example, a circular network is regular of valency $r = 2$, while a complete network of n individuals is regular of valency $r = n - 1$. In our model, regular networks have the following remarkable property.

Proposition E1 (Regular networks) *In regular networks, the aggregate effort does not vary with λ, that is,*

$$\sum_{i=1}^{n} x_i^* = \sum_{i=1}^{n} \alpha_i. \tag{E.2}$$

Proof of Proposition E1.

Given (A.17), for a regular network, we have $\boldsymbol{\pi} = (1/n, \ldots, 1/n)$, which we plug into (A.17) and multiply both parts of the resulting equality by n, yielding

$$\sum_{j=1}^{n} x_i^* = \sum_{j=1}^{n} \alpha_j.$$

Because this equality holds for any λ, the proof is complete. ▮

Proposition E1 shows that, even if each individual effort varies non-trivially with λ in simple regular networks, the total effort is not affected by a change in the taste for conformity.

E.3 Impact of conformism on individual and aggregate effort: Star-shaped networks

Consider a star-shaped network in which $i = 1$ is the star agent. Denote

$$\alpha^{\mathrm{s}} := \alpha_1, \qquad \alpha^{\mathrm{p}} := \frac{\alpha_2 + \ldots + \alpha_n}{n-1}.$$

In other words, α^{s} is the productivity of the star agents, while α^{p} is the average productivity of all periphery agents.

Proposition E2 (Star-shaped networks) *Consider a star network.*

(i) Assume $\alpha^{\mathrm{s}} < \alpha^{\mathrm{p}}$. Then, the effort x_1^* of the star agent increases with λ but the aggregate effort $\sum_i x_i^*$ decreases in λ. For any periphery agent $i \geq 2$, we obtain

- *(ia) if $\alpha_i \leq \alpha^{\mathrm{s}}$, then x_i^* increases with λ;*
- *(ib) if $\alpha^{\mathrm{s}} < \alpha_i < (3\alpha^{\mathrm{p}} + \alpha^{\mathrm{s}})/4$, then x_i^* is U shaped in λ;*
- *(ic) if $\alpha_i \geq (3\alpha^{\mathrm{p}} + \alpha^{\mathrm{s}})/4$, then x_i^* decreases with λ.*

(ii) Assume $\alpha^{\mathrm{s}} > \alpha^{\mathrm{p}}$. Then, the effort of the star agent $x_i^{\mathrm{s}*}$ decreases with λ but the aggregate effort $\sum_i x_i^*$ increases in λ. For a periphery agent $i \geq 2$, we obtain

- *(iia) if $\alpha_i \leq (3\alpha^{\mathrm{p}} + \alpha^{\mathrm{s}})/4$, then x_i^* increases with λ;*
- *(iib) if $(3\alpha^{\mathrm{p}} + \alpha^{\mathrm{s}})/4 < \alpha_i < \alpha^{\mathrm{s}}$, then x_i^* is bell shaped in λ;*
- *(iic) if $\alpha_i \geq \alpha^{\mathrm{s}}$, then x_i^* decreases with λ.*

Proof of Proposition E2.

Let us start with the following lemma:

Lemma E7 *Assume that* $\mathbf{g}$ *is a star-shaped network where $i = 1$ is the star agent. Then, for all $\lambda \in (0, 1)$, we have*

$$\frac{\partial x_1^*}{\partial \lambda} = \frac{\left(\alpha^P - \alpha^S\right)}{\left(1+\lambda\right)^2}$$

and for the periphery agents $i = 2, .., n$,

$$\frac{\partial x_i^*}{\partial \lambda} = \alpha^P - \alpha_i + \frac{\left(\alpha^S - \alpha^P\right)}{\left(1+\lambda\right)^2}$$

Proof of Lemma E7: It is readily verified that, for a star-shaped network with n agents, $i = 1$ being the star node, the row-normalized adjacency matrix $\widehat{\mathbf{G}}$ and its square $\widehat{\mathbf{G}}^2$ are given by

$$\widehat{\mathbf{G}} = \begin{pmatrix} 0 & \frac{1}{n-1} & \dots & \frac{1}{n-1} \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix}, \qquad \widehat{\mathbf{G}}^2 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{1}{n-1} & \dots & \frac{1}{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{1}{n-1} & \dots & \frac{1}{n-1} \end{pmatrix}. \tag{E.3}$$

Furthermore, we have

$$\widehat{\mathbf{G}}^3 = \begin{pmatrix} 0 & \frac{1}{n-1} & \dots & \frac{1}{n-1} \\ 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix} = \widehat{\mathbf{G}}, \tag{E.4}$$

In other words, the matrix $\widehat{\mathbf{G}}$ is cyclical with the cycle length equal to 2. Combining (E.3)–(E.4) with (11) and simplifying it, we obtain

$$\mathbf{x}^*(\lambda) = (1-\lambda)\boldsymbol{\alpha} + \frac{\lambda}{1+\lambda}\widehat{\mathbf{G}}\boldsymbol{\alpha} + \frac{\lambda^2}{1+\lambda}\widehat{\mathbf{G}}^2\boldsymbol{\alpha}.$$

Further simplification yields that the effort of the star agent, $i = 1$, is given by

$$x_1^*(\lambda) = \alpha^{\mathrm{p}} + \frac{\alpha^{\mathrm{s}} - \alpha^{\mathrm{p}}}{1+\lambda}, \tag{E.5}$$

while the efforts of periphery agents, $i = 2, \dots, n$ are as follows:

$$x_i^*(\lambda) = (1-\lambda)\alpha_i + \frac{\lambda}{1+\lambda}\alpha^{\mathrm{s}} + \frac{\lambda^2}{1+\lambda}\alpha^{\mathrm{p}} \tag{E.6}$$

Differentiating (E.5)–(E.6) with respect to λ, we obtain

$$\frac{\partial x_1^*}{\partial \lambda} = -\frac{(\alpha^{\mathrm{s}} - \alpha^{\mathrm{p}})}{(1+\lambda)^2}, \tag{E.7}$$

$$\frac{\partial x_i^*}{\partial \lambda} = \alpha^{\mathrm{p}} - \alpha_i + \frac{\alpha^{\mathrm{s}} - \alpha^{\mathrm{p}}}{(1+\lambda)^2}. \tag{E.8}$$

respectively. This completes the proof of this lemma. ▮

Using Lemma E7, it is straightforward to characterize the effect of λ on x_i^* for the periphery agents.

Let us prove the impact of λ on the aggregate effort. Using (E.7) and (E.8), we obtain

$$\begin{aligned}
\frac{\partial \sum_i x_i^*}{\partial \lambda} &= \frac{\partial x_1^*}{\partial \lambda} + \sum_{i=1}^{n} \frac{\partial x_i^*}{\partial \lambda} \\
&= -\frac{(\alpha^{\mathrm{s}} - \alpha^{\mathrm{p}})}{(1+\lambda)^2} + (n-1)\,\alpha^{\mathrm{p}} + \frac{(n-1)\,(\alpha^{\mathrm{s}} - \alpha^{\mathrm{p}})}{(1+\lambda)^2} - \sum_{i=2}^{n} \alpha_i \\
&= \frac{(n-2)\,(\alpha^{\mathrm{s}} - \alpha^{\mathrm{p}})}{(1+\lambda)^2}
\end{aligned}$$

since, by definition, $(n-1)\,\alpha^{\mathrm{p}} = \sum_{i=2}^{n} \alpha_i$. This completes the proof. ∎

Proposition E2 provides a more precise description of the impact of a higher taste for conformity on equilibrium efforts in a star network. In particular, it shows that, if the productivity of an agent is high (low), then an increase in the taste for conformity reduces (increases) her effort, because she feels pressured by her friends who provide, on average, lower (higher) effort.

E.4 Impact of conformism on individual and aggregate effort: Circular networks

Consider a circular network (which is a regular network of valency $r = 2$) with $n = 5$ agents in which productivity is given by $\alpha_i = i$ for all $i = 1, \ldots, 5$.

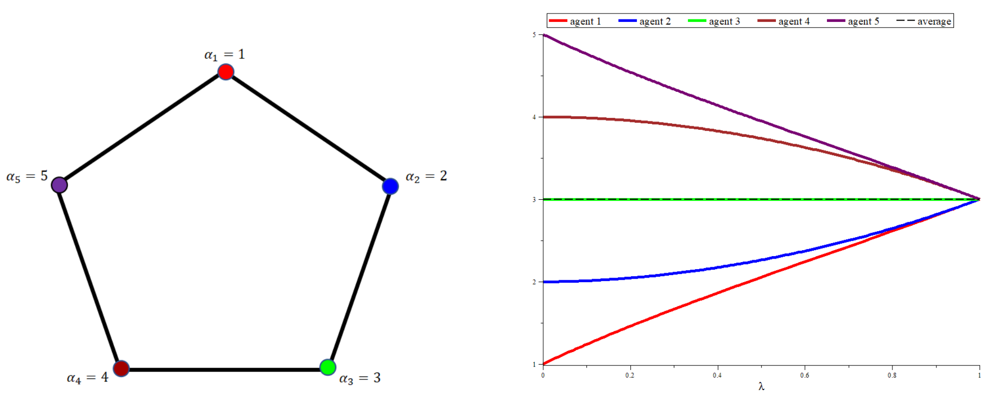

Figure E2: How λ affects effort in a circular network

The left panel of Figure E2 depicts the network structure and productivity pattern, while the right panel shows how individual efforts vary with λ. Clearly, the efforts x_1^* and x_2^* of the low-productive agents 1 and 2 increase with λ, while the efforts x_4^* and x_5^* of highly productive agents 4 and 5 decrease with λ. However, the effort of the "median" agent (individual 3) remains at the same level as the average effort $\sum_{j=1}^{5} x_i^*/5$ in the network and is not affected by a change in λ. This is in accordance with Proposition E1, as the average effort is strictly proportional to the aggregate effort.

Let us now rewire the links in the network of Figure E2 so that, topologically, the new network, displayed in the left panel of Figure E3, is isomorphic to the previous one. In particular, the new network is still a circular (regular) network with $n = 5$ agents but the social norms are very different to those in Figure E2. In particular, the neighbors of agent 2 are now agents 4 and 5, the two most productive individuals in the economy, whereas before, her neighbors were 1 and 3, who are clearly less productive. By contrast, the neighbors of agent 4 are now agents 1 and 2 instead of agents 3 and 5, which means that her neighbors are now less productive. Thus, when we compare the right panels of Figures E2 and E3, we observe that the convergence of agent 2's and agent 4's efforts to the average effort is now faster than in the original network.

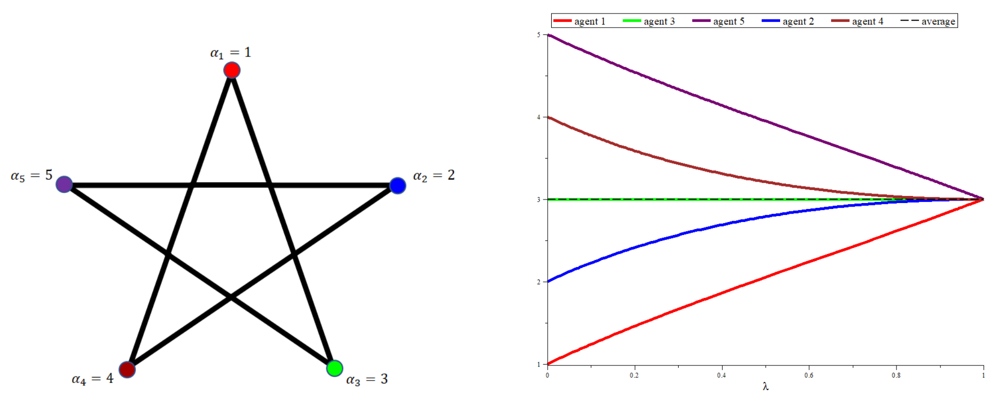

Figure E3: How λ affects effort in a rewired circular network

Finally, let us rewire again the social ties without changing the network topology. We obtain the network depicted in the left panel of Figure E4. Although the network is regular, we observe, in the right panel of Figure E4, that the effect of λ on the effort of the median player (agent 3) is now U shaped and is different from the average effort in the network. Indeed, when λ is small, the immediate impact of direct neighbors dominates the indirect impact of the others on individual i. The neighbors of agent 3 are now agents 1 and 2, who are both less productive than agent 3 is. Therefore, under pure individualism ($\lambda = 0$), we have $x_3^* = \alpha_3 = 3$, but as λ becomes slightly positive, x_3^* decreases. However, when λ becomes larger, the indirect impact of agents 4 and 5, who are more productive than agent 3, becomes sufficiently strong, which results in a higher effort of agent 3.

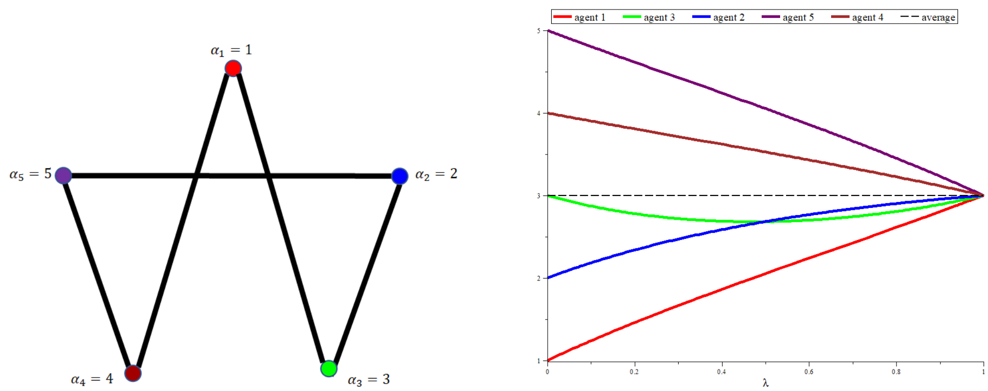

Figure E4: How λ affects effort in another rewired circular network

F Equilibrium versus first best: Specific networks

We illustrate condition (31), which gives a condition for the equilibrium effort to be optimal, for specific networks.

Example 1. Consider a *star network* in which agent 1 is in the center. Then, we have $\mathbf{x}^* = \mathbf{x}^O$ if and only if the star-agent productivity is equal to the average productivity of all periphery agents:

$$\alpha_1 = \frac{1}{n-1}\sum_{j=1}^{n} \alpha_j. \tag{F.1}$$

In particular, when there are two levels of productivity, that is, $\alpha_i \in \{\alpha^L, \alpha^H\}$, $\alpha^H > \alpha^L > 0$, the Nash equilibrium in a star-shaped network is never optimal.

Example 2. Assume now that $\mathbf{g}$ is a circular network with $n = 4$, so that each agent has two links. Then, we have $\mathbf{x}^* = \mathbf{x}^O$ if and only if the average productivity across *maximum independent sets*[3] is the same, that is,

$$\frac{\alpha_1 + \alpha_3}{2} = \frac{\alpha_2 + \alpha_4}{2}.$$

In particular, when there are two levels of productivity, that is, $\alpha_i \in \{\alpha^L, \alpha^H\}$, $\alpha^H > \alpha^L > 0$, the Nash equilibrium is optimal if and only if there are two highly productive agents and two low-productive agents, and the highly productive agents are all linked to each other.

Examples 1 and 2 are special cases of the following more general results.

Corollary F.1 *If the network is a complete bipartite graph (i.e.,* $\mathbf{g} = K_{m,n}$*) with a partition* $V_1 \cup V_2$ *where* $|V_1| = m$, $|V_2| = n$, *then we have* $\mathbf{x}^* = \mathbf{x}^O$ *if and only if* $\overline{\alpha}_1 = \overline{\alpha}_2$, *where* $\overline{\alpha}_r$ *is the average productivity over* V_r, $r = 1, 2$, *that is*

$$\overline{\alpha}_r := \frac{1}{|V_r|}\sum_{k \in V_r} \alpha_k. \tag{F.2}$$

Proof of Corollary F.1: Let us derive (31) for complete bipartite graphs: $\mathbf{g} = K_{m,n}$. The Nash equilibrium $\mathbf{x}^*$ is the solution to (12). A necessary and

[3] In graph theory, an *independent set* is a set of nodes in a graph such that no two nodes are adjacent. In other words, it is a set S of nodes such that for every two vertexes in S, there is no edge connecting the two. A *maximum independent set* is an independent set of the largest possible size for a given graph.

sufficient condition (31) for the Nash equilibrium $\mathbf{x}^*$ to deliver a first best is given by

$$\widehat{\mathbf{G}}^T(\mathbf{x}^* - \overline{\mathbf{x}}^*) = \mathbf{0}.$$

Using (12), this condition can be equivalently restated as follows:

$$\widehat{\mathbf{G}}^T\boldsymbol{\alpha} = \widehat{\mathbf{G}}^T\overline{\mathbf{x}}^*, \tag{F.3}$$

where $\overline{\mathbf{x}}^* = \widehat{\mathbf{G}}\mathbf{x}^*$ is the vector of equilibrium social norms.

Recall that, when $\mathbf{g} = K_{m,n}$, the row-normalized adjacency matrix $\widehat{\mathbf{G}}$ is given by (C.6). Hence, the best reply functions (11) take the following form:

$$x_i = (1-\lambda)\alpha_i + \begin{cases} \dfrac{\lambda}{n}\displaystyle\sum_{k\in V_2} x_k, & i \in V_1, \\ \dfrac{\lambda}{m}\displaystyle\sum_{k\in V_1} x_k, & i \in V_2. \end{cases} \tag{F.4}$$

Computing the means across all $i \in V_1$ and across all $i \in V_2$, we obtain

$$\frac{1}{m}\sum_{k\in V_1} x_k = \frac{1-\lambda}{m}\sum_{k\in V_1}\alpha_k + \frac{\lambda}{n}\sum_{k\in V_2} x_k, \tag{F.5}$$

$$\frac{1}{n}\sum_{k\in V_2} x_k = \frac{1-\lambda}{n}\sum_{k\in V_2}\alpha_k + \frac{\lambda}{m}\sum_{k\in V_1} x_k. \tag{F.6}$$

respectively. Note that $\frac{1}{n}\sum_{k\in V_2} x_k = \overline{x}_i$ for any individual $i \in V_1$, while $\frac{1}{m}\sum_{k\in V_1} x_k = \overline{x}_j$ for any individual $j \in V_2$. Without loss of generality, let agent $i = 1$ belong to V_1, and let agent $i = 2$ belong to V_2. Then, we have

$$\frac{1}{n}\sum_{k\in V_2} x_k = \overline{x}_1, \qquad \frac{1}{m}\sum_{k\in V_1} x_k = \overline{x}_2.$$

Solving the system (F.5)–(F.6) in terms of $\overline{x}_1$ and $\overline{x}_2$, we obtain

$$\overline{x}_r^* = \frac{\lambda}{1+\lambda}\overline{\alpha}_r + \frac{1}{1+\lambda}\overline{\alpha}_s, \tag{F.7}$$

where $r, s = 1, 2$, $r \neq s$. Finally, observe that the following equalities hold:

$$\widehat{\mathbf{G}}^T\boldsymbol{\alpha} = \begin{pmatrix} \mathbf{0}_{m\times m} & \frac{1}{m}\mathbf{1}_{m\times n} \\ \frac{1}{n}\mathbf{1}_{n\times m} & \mathbf{0}_{n\times n} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \\ \alpha_{m+1} \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \frac{n}{m}\overline{\alpha}_2 \\ \vdots \\ \frac{n}{m}\overline{\alpha}_2 \\ \frac{m}{n}\overline{\alpha}_1 \\ \vdots \\ \frac{m}{n}\overline{\alpha}_1 \end{pmatrix},$$

$$\widehat{\mathbf{G}}^T\mathbf{x}^* = \begin{pmatrix} \mathbf{0}_{m\times m} & \frac{1}{m}\mathbf{1}_{m\times n} \\ \frac{1}{n}\mathbf{1}_{n\times m} & \mathbf{0}_{n\times n} \end{pmatrix} \begin{pmatrix} \overline{x}_1^* \\ \vdots \\ \overline{x}_1^* \\ \overline{x}_2^* \\ \vdots \\ \overline{x}_2^* \end{pmatrix} = \begin{pmatrix} \frac{n}{m}\overline{x}_2^* \\ \vdots \\ \frac{n}{m}\overline{x}_2^* \\ \frac{m}{n}\overline{x}_1^* \\ \vdots \\ \frac{m}{n}\overline{x}_1^* \end{pmatrix}.$$

Hence, the condition (F.3) holds if and only if $\overline{\alpha}_r = \overline{x}_r^*$ for $r = 1, 2$. Using (F.7), we find that this is equivalent to $\overline{\alpha}_1 = \overline{\alpha}_2$. This completes the proof. ∎

This corollary and the examples above show that, in order for the equilibrium efforts to be optimal, there needs to be some compensation for the externalities that agents exert on others. In particular, for bipartite networks, such as the star and circular network, the average productivity of the different agents has to be the same. For example, in the star network, it cannot be that the productivity of the star is much higher than the average productivity of the peripheral agents because, in that case, the externalities that the star exerts on the peripheral agents are not exactly compensated by the externalities created by the peripheral agents on the star.

G Extensions of the benchmark model

G.1 Weighted networks

In the baseline local-average model, the social network was *undirected*, *unweighted* and with *no self-loops*. Consider an extension in which the network may be *directed*, *weighted* and may have *self-loops* as in the standard DeGroot model (Golub and Jackson, 2010). Let $\mathbf{W} = [w_{ij}]$ be an arbitrary $(n \times n)$ row-normalized irreducible matrix with non-negative entries. Each cell w_{ij}, $i, j = 1, 2, \ldots, n$, gives the relative impact (weight) of agent j's effort on agent i's social norm $\overline{x}_i$ defined as follows:

$$\overline{x}_i \equiv \sum_{j=1}^{n} w_{ij} x_j. \tag{G.1}$$

In particular, we do not rule out self loops, i.e., we allow for the possibility that $w_{ii} > 0$ for some i. Agent i's utility function is the same as in the baseline model and given by (3), i.e.

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{W}) = \alpha_i x_i - \frac{1}{2} x_i^2 - \theta \left(x_i - \overline{x}_i \right)^2,$$

but agent i's social norm $\overline{x}_i$ is now equal to (G.1). The best reply mapping is then given by

$$\mathbf{x} = (1 - \lambda)\boldsymbol{\alpha} + \lambda \mathbf{W}\mathbf{x}, \tag{G.2}$$

where $\lambda \equiv \theta/(1+\theta) \in (0, 1)$. Because the weighting matrix $\mathbf{W}$ is row-normalized, non-negative, and irreducible, we have: $\rho(\mathbf{W}) = 1$, where $\rho(\cdot)$ is th spectral radius of matrix $\mathbf{W}$. Combining this with $\lambda \in (0, 1)$, we infer that the best-reply mapping (G.2) is a contraction mapping, hence, a unique (interior) equilibrium $\mathbf{x}^*$ exists. Furthermore, equilibrium efforts $\mathbf{x}^*$ can be represented using the Newman-series formula:

$$\mathbf{x}^* = (1 - \lambda) \sum_{k=0}^{\infty} \lambda^k \mathbf{W}^k \boldsymbol{\alpha}. \tag{G.3}$$

Let us now examine if our main results remain the same with a weighted adjacency matrix $\mathbf{W}$ instead of a non-weighted one $\widehat{\mathbf{G}}$.

Total conformism. Let $\boldsymbol{\pi}$ be the left eigenvector of $\mathbf{W}$ associated with $\rho(\mathbf{W}) = 1$. Then, the analysis of total conformism ($\lambda \to 1$) exactly follows the one of the baseline model, that is

$$\lim_{\lambda \to 1} x_i^* = \boldsymbol{\pi}\boldsymbol{\alpha}, \tag{G.4}$$

Observe, however, that $\boldsymbol{\pi}$ can no longer be expressed in a closed form as a function of $\mathbf{W}$. The proof of (G.4) repeats almost verbatim to that of Proposition 3, up to replacing $\widehat{\mathbf{G}}$ with $\mathbf{W}$ everywhere. Similarly, Proposition 4 remains true, except that $\boldsymbol{\pi}$ is no longer proportional to $(d_1, d_2, \ldots, d_n)$.

Comparative statics. The comparative statics exercises of equilibrium efforts with respect to $\boldsymbol{\alpha}$ remain the same. The comparative statics with respect to the taste for conformity λ is already ambiguous in the baseline model with $\widehat{\mathbf{G}}$ (see Proposition 5), so there is no reason to expect more clear-cut results in the setting with arbitrary weights.

The only aspect of the analysis where this extension with $\mathbf{W}$ buys substantial tractability compared to the baseline model is when we study the effects of changes in the network structure. The reason is that one can use standard tools of calculus to study the consequences of infinitesimal perturbations in $\mathbf{W}$. Indeed, assume that the weighting matrix changes as follows:

$$\mathbf{W} \to \mathbf{W} + \mathrm{d}\mathbf{W}, \tag{G.5}$$

where $\mathrm{d}\mathbf{W} = [\mathrm{d}w_{ij}]$ is the $(n \times n)$ perturbation matrix. Assume that the perturbations $\mathrm{d}w_{ij}$ are arbitrarily small (infinitesimal in the limit). Furthermore, because the new weighting matrix, $\mathbf{W} + \mathrm{d}\mathbf{W}$, must be row-normalized, the perturbation $\mathrm{d}\mathbf{W}$ must satisfy:

$$\mathrm{d}\mathbf{W}\mathbf{1} = \mathbf{0}.$$

Let $\mathrm{d}\mathbf{x}^*$ stand for the response of the equilibrium efforts $\mathbf{x}^*$ to the infinitesimal weight perturbation (G.5). Using (G.3), we get:

$$\mathbf{x}^* = (1-\lambda) \sum_{k=1}^{\infty} k \lambda^k \mathbf{W}^{k-1} \mathrm{d}\mathbf{W} \boldsymbol{\alpha}. \tag{G.6}$$

We have the following result:

Proposition G3 *A small perturbation* (G.5) *of the weighting matrix* $\mathbf{W}$ *leads to:*

(*i*) *an increase in everyone's effort, if*

$$\mathrm{d}\mathbf{W}\boldsymbol{\alpha} > \mathbf{0}; \tag{G.7}$$

(*ii*) *a reduction in everyone's effort, if*

$$\mathrm{d}\mathbf{W}\boldsymbol{\alpha} < \mathbf{0}; \tag{G.8}$$

(iii) an ambiguous outcome, otherwise.

Intuitively, condition (G.7) means that high-productive agents end up having more impact on everyone's norms, while condition (G.7) means the reverse. To see this, let $[\mathrm{d}\mathbf{W}]_i$ denote the ith column of the perturbation matrix $\mathrm{d}\mathbf{W}$, $i = 1, 2, \ldots, n$. Then, we have:

$$\mathrm{d}\mathbf{W}\boldsymbol{\alpha} = \sum_{i=1}^{n} \alpha_i [\mathrm{d}\mathbf{W}]_i.$$

Hence condition (G.7) clearly holds if $[\mathrm{d}\mathbf{W}]_i > \mathbf{0}$ for agents with high productivities.

Welfare results. The exact counterparts of Propositions 7, 8 and 9 hold true. The proofs repeat verbatim those of the baseline model, up to replacing $\widehat{\mathbf{G}}$ with $\mathbf{W}$ throughout.

G.2 Heterogeneous tastes for conformity

In our model, we assume that all individuals have the same taste for conformity λ. We now allow for heterogeneity in taste for conformity. Agent i's payoff (6) becomes

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{1}{2} x_i^2 - \frac{1}{2} \left(\frac{\lambda_i}{1 - \lambda_i} \right) (x_i - \overline{x}_i)^2 , \tag{G.9}$$

where $\lambda_i \in (0, 1)$ is the taste for conformity, which is now agent-specific. For each agent $i = 1, 2, \ldots, n$, her best-reply function is given by

$$x_i = (1 - \lambda_i)\alpha_i + \lambda_i \overline{x}_i. \tag{G.10}$$

or, equivalently, in vector-matrix form

$$\mathbf{x} = (\mathbf{I} - \boldsymbol{\Lambda})\boldsymbol{\alpha} + \boldsymbol{\Lambda}\widehat{\mathbf{G}}\mathbf{x}, \tag{G.11}$$

where $\mathbf{I}$ is the identity matrix, while $\boldsymbol{\Lambda}$ is the $(n \times n)$-matrix defined by

$$\boldsymbol{\Lambda} := \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Observe that the matrix $\mathbf{\Lambda}\widehat{\mathbf{G}}$ has non-negative entries, while its maximum eigenvalue is smaller than one. Combining this with (G.11), it is readily verified that a unique Nash equilibrium exists, it is interior and is given by:

$$\mathbf{x}^* = (\mathbf{I} - \mathbf{\Lambda}\widehat{\mathbf{G}})^{-1}(\mathbf{I} - \mathbf{\Lambda})\boldsymbol{\alpha} = \sum_{k=0}^{\infty}(\mathbf{\Lambda}\widehat{\mathbf{G}})^k(\mathbf{I} - \mathbf{\Lambda})\boldsymbol{\alpha}, \tag{G.12}$$

where the expansion to the Neumann series is justified by the fact that the spectral radius of $\mathbf{\Lambda}\widehat{\mathbf{G}}$ is smaller than one. This implies that:

$$\overline{\mathbf{x}}^* = \widehat{\mathbf{G}}\sum_{k=0}^{\infty}(\mathbf{\Lambda}\widehat{\mathbf{G}})^k(\mathbf{I} - \mathbf{\Lambda})\boldsymbol{\alpha}, \tag{G.13}$$

Furthermore, since x_i^* is a convex combination of α_i and $\overline{x}_i^*$, we have:

$$x_i \gtreqqless \overline{x}_i^* \quad \Longleftrightarrow \quad \alpha_i \gtreqqless \overline{x}_i^*, \tag{G.14}$$

which is the extension of Lemma 1 for the case of heterogeneous taste for conformity. This result proves useful in studying comparative statics of equilibrium with respect to an individual taste for conformity. Indeed, by total differentiation of (G.10) with respect to λ_j, and using the definition of agent i's social norm $\overline{x}_i$, we obtain:

$$\frac{\partial x_i^*}{\partial \lambda_j} = \delta_{ij}(\overline{x}_j^* - \alpha_j) + \lambda_i \frac{\partial \overline{x}_i^*}{\partial \lambda_j},$$

where δ_{ij} is defined by (A.11). In vector-matrix form, this can be written as:

$$\frac{\partial \mathbf{x}^*}{\partial \lambda_j} = (\overline{x}_j^* - \alpha_j)\mathbf{e}_j + \mathbf{\Lambda}\widehat{\mathbf{G}}\frac{\partial \mathbf{x}^*}{\partial \lambda_j}, \tag{G.15}$$

where $\mathbf{e}_j$ is the n-dimensional vector whose jth component is equal to 1 while all the others are equal to zero. Solving the linear system (G.15) for $\partial \mathbf{x}^*/\partial \lambda_j$ yields:

$$\frac{\partial \mathbf{x}^*}{\partial \lambda_j} = (\overline{x}_j^* - \alpha_j)\sum_{k=0}^{\infty}(\mathbf{\Lambda}\widehat{\mathbf{G}})^k\mathbf{e}_j. \tag{G.16}$$

Define the following vector:

$$\mathbf{a}(\boldsymbol{\alpha}, \mathbf{\Lambda}, \mathbf{g}) := \mathbf{x}^* - \overline{\mathbf{x}}^* = (\mathbf{I} - \widehat{\mathbf{G}})\sum_{k=0}^{\infty}(\mathbf{\Lambda}\widehat{\mathbf{G}})^k(\mathbf{I} - \mathbf{\Lambda})\boldsymbol{\alpha}.$$

Proposition G4 *Assume that agent j becomes slightly more conformist, i.e., λ_j marginally increases, then*

(*i*) *the efforts of all agents, including agent j, change in the same direction;*

(*ii*) *the efforts of all agents increase (decrease) if and only if $x_j^* < \overline{x}_j^*$ ($x_j^* > \overline{x}_j^*$) or equivalently $a_j(\boldsymbol{\alpha}, \boldsymbol{\Lambda}, \mathbf{g}) < 0$ ($a_j(\boldsymbol{\alpha}, \boldsymbol{\Lambda}, \mathbf{g}) > 0$).*

Proof of Proposition G4

(*i*) Because the matrix $\boldsymbol{\Lambda}\widehat{\mathbf{G}}$ is irreducible, the vector $\sum_{k=0}^{\infty}(\boldsymbol{\Lambda}\widehat{\mathbf{G}})^k\mathbf{e}_j$ is strictly positive. Hence (G.16) implies that all components of $\partial\mathbf{x}^*/\partial\lambda_j$ have the same sign. This proves (*i*).

(*ii*) Equation (G.16) implies that:

$$\frac{\partial \mathbf{x}^*}{\partial \lambda_j} \gtreqless 0 \Leftrightarrow x_j^* \gtreqless \overline{x}_j^*. \tag{G.17}$$

We can restate this in terms of the primitives of the model. Consider $\mathbf{a}(\boldsymbol{\alpha}, \boldsymbol{\Lambda}, \mathbf{g})$. Clearly, each component $a_j(\boldsymbol{\alpha}, \boldsymbol{\Lambda}, \mathbf{g})$ of this vector is a well-defined function of the primitives. Combining (G.17) with (G.14), we obtain (*ii*). ∎

G.3 Anti-conformism

In the baseline local-average model, individual i's utility function is given by (3), that is

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{x_i^2}{2} - \frac{\theta}{2}\left(x_i - \overline{x}_i\right)^2, \tag{G.18}$$

where $\theta > 0$ is taste for conformity. Let us now assume that $\theta < 0$ so that agents are *anti-conformists*. Denote, as above, $\lambda \equiv \theta/(1+\theta)$. We have the following result:

Proposition G5 (Anti-conformism) *Assume $\theta < 0$.*

(*i*) *If $\theta > -1/2$, then, there exists a unique equilibrium, in which player i's best reply is given by:*

$$x_i = \max\{0, (1-\lambda)\alpha_i + \lambda\overline{x}_i\}. \tag{G.19}$$

(*ii*) *If $-1/2 \leq \theta < -1$, player i's best reply is still given by* (G.19) *but multiple equilibria may arise.*

(*iii*) *If $\theta \geq -1$, no pure-strategy equilibrium exists.*

Proof of Proposition G5:

Case (i) Assume $\theta > -1/2$. In this case, most of the equilibrium analysis still works through, unless the productivity dispersion across individuals is very high. As implied by the first-order conditions, at least in the vicinity of the equilibrium, the best-reply mapping has to be given by:

$$\mathbf{x} = (1-\lambda)\alpha + \lambda \widehat{\mathbf{G}}\mathbf{x},$$

where $\lambda \equiv \theta/(1+\theta)$. The condition $\theta > -1/2$ guarantees that $|\lambda| < 1$, hence, the best-reply mapping is still a contraction mapping. Therefore, a unique interior equilibrium exists, and, more importantly, the widely used Neumann-series decomposition

$$\mathbf{x}^* = (1-\lambda)\sum_{k=0}^{\infty} \lambda^k \widehat{\mathbf{G}}^k \boldsymbol{\alpha} \tag{G.20}$$

of equilibrium efforts still holds. Two issues arise, however.

First, in the anti-conformist case, efforts are strategic *substitutes* rather than strategic complements. This implies a more tricky comparative statics of equilibrium efforts with respect to individual productivities $\boldsymbol{\alpha}$. To see this, observe that, because $\lambda < 0$ in the anti-conformist model, the terms of the series in the right-hand-side of (G.20) have alternating signs. As a result, the average productivity across the neighbors of agent i leads to a reduction of x_i^* while the average productivity across her neighbors of distance 2 leads to an increase in x_i^*, etc.[4] Thus, in contrast to Proposition 2, the impact of α_j on x_i^* is a priori ambiguous.

Second, unlike the case with conformity ($\theta > 0$), in general, the interiority of equilibrium is not guaranteed. To see this, let us restate (G.18) as follows:

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = (\alpha_i + \theta \overline{x}_i)x_i - \frac{1+\theta}{2}x_i^2 - \frac{\theta}{2}\overline{x}_i^2. \tag{G.21}$$

Since $\theta < 0$, we have take into account the possibility that $\alpha_i + \theta \overline{x}_i < 0$, in which case we have:

$$\arg\max_{x_i \geq 0} U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = 0.$$

Hence, player i's best reply is thus given by (G.19). Since $|\lambda| < 1$, the best-reply mapping is still a contraction mapping. Hence, the equilibrium is always unique.

To explain why the equilibruim needs not be interior, consider the simplest possible case when two agents ($n = 2$) are linked to each other (the dyad). The candidate

[4] A similar effect, called "my enemy's enemy is my friend", has also been found by Ushchev and Zenou (2018) in a model of price competition in product variety networks.

interior equilibrium is given by

$$x_i^* = \frac{1}{1+\lambda}\alpha_i + \frac{\lambda}{1+\lambda}\alpha_j, \tag{G.22}$$

where $i, j = 1, 2$, and $i \neq j$. Since $\lambda < 0$, for (G.22) to be an interior equilibrium, it is necessary and sufficient that the following inequalities hold:

$$-\lambda < \frac{\alpha_1}{\alpha_2} < -\frac{1}{\lambda}. \tag{G.23}$$

Condition (G.23) can be equivalently restated as:

$$|\lambda| < \frac{\min\{\alpha_1, \alpha_2\}}{\max\{\alpha_1, \alpha_2\}}. \tag{G.24}$$

Condition (G.23) or, equivalently, (G.24), means that agents must be either not too heterogeneous in productivities, or not too non-conformist, or both.

To illustrate this, we contrast the cases when $\lambda = 1/2$ (moderate conformity) and when $\lambda = -1/2$ (moderate anti-conformity). These cases emerge when, respectively, $\theta = 1$ and $\theta = -1/3$, so that the condition $\theta > -1/2$ is satisfied in the anti-conformity case. Without loss of generallity, assume that $\alpha_1 \geq \alpha_2$.

In the conformity case ($\lambda = 1/2$), equation (G.22) takes the form of:

$$x_i^* = \frac{2}{3}\alpha_i + \frac{1}{3}\alpha_j,$$

and defines an interior equilibrium for any values $(\alpha_1, \alpha_2) \in \mathbb{R}^2_{++}$ of productivities. Furthermore, in accordance with part (*i*) of Proposition 2, each agent's effort increases both with her own productivity and her peer's productivity.

In the anti-conformity case ($\lambda = -1/2$), equation (G.22) becomes:

$$x_i^* = 2\alpha_i - \alpha_j.$$

Two observations are in order. First, in contrast to Proposition 2, when the equilibrium is interior, each agent's effort increases with her own productivity but *decreases* with the productivity of the other agent. Second, for the equilibrium to be interior, condition (G.24) must hold. Because $\alpha_1 \geq \alpha_2$ and $\lambda = -1/2$, condition (G.24) takes the form:

$$\alpha_1 < 2\alpha_2.$$

When $\alpha_1 \geq 2\alpha_2$, it is readily verified that the unique equilibrium is given by $(x_1^*, x_2^*) = (3\alpha_1/2, 0)$. Combining this with (G.22), we have:

$$x_1^* = \begin{cases} 2\alpha_1 - \alpha_2, & \text{if } \alpha_1 < 2\alpha_2, \\ \frac{3}{2}\alpha_1, & \text{if } \alpha_1 \geq 2\alpha_2; \end{cases} \qquad x_2^* = \begin{cases} 0, & \text{if } \alpha_2 \leq \frac{\alpha_1}{2}, \\ 2\alpha_2 - \alpha_1, & \text{if } \alpha_2 > \frac{\alpha_1}{2}. \end{cases} \tag{G.25}$$

Case (ii): Assume $-1/2 \leq \theta < -1$. In this case, the best replies are still given by (G.19). However, the best-reply mapping is no longer a contraction maping. As a result, multiple equilibria may arise.

To illustrate this, consider again the dyad example with $n = 2$, but set $\alpha_1 = \alpha_2 = 1$, and $\theta = -2/3$. Then, it is readily verified that the best reply of player i is given by

$$x_i = \max\{0, 3 - 2x_j\}, \tag{G.26}$$

where $i, j = 1, 2$, and $i \neq j$. A simple graphical analysis shows there are three different equilibria (see Figure G5 below), including a symmetric interior equilibrium and two corner equilibria:

$$(x_1^*, x_2^*) = (1, 1), \qquad (x_1^{**}, x_2^{**}) = (3, 0), \qquad (x_1^{***}, x_2^{***}) = (0, 3).$$

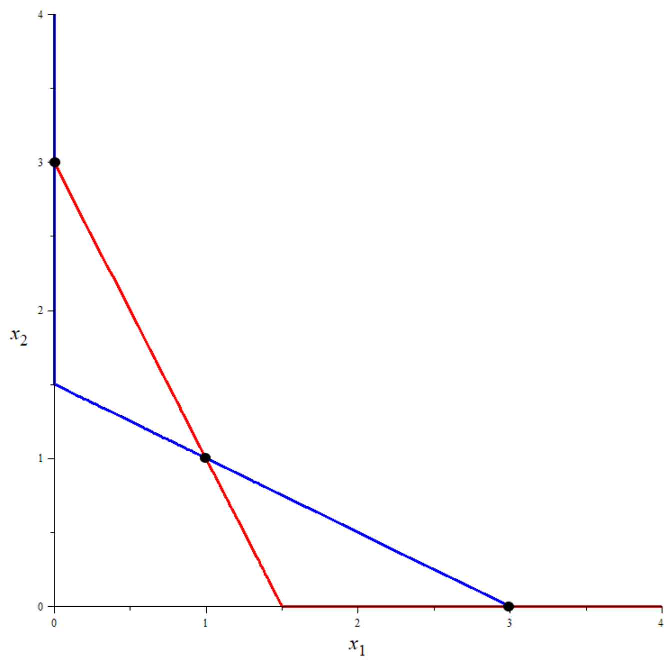

Figure G5: Multiple equilibria under anti-conformism

Case (iii): Assume $\theta \geq -1$. In this case, as implied by (G.18), individual i's payoff function becomes unbounded with respect to x_i. Hence, no pure-strategy equilibrium exists. ∎

G.4 Ambition and social norms

Assume that the utility of individual i is now given by:

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{x_i^2}{2} - \frac{\theta}{2}\left(x_i - \beta_i \overline{x}_i\right)^2, \tag{G.27}$$

where $\beta_i \geq 1$ is agent i's *ambition factor*. Since $\beta_i \geq 1$, the "reference effort" of each individual i is now higher than the social norm $\overline{x}_i$ of her neighbors. Clearly, by setting $\beta_1 = \ldots = \beta_n = 1$, we are back to our benchmark model.

We also impose the following condition on taste λ for conformity:

$$\lambda \beta_{\max} < 1, \tag{G.28}$$

where $\beta_{\max} \equiv \max\{\beta_1, \ldots, \beta_n\}$. It is readily verified that the equilibrium efforts and the socially optimal efforts are given by the same expressions as above, up to replacing $\widehat{\mathbf{G}}$ with $\mathbf{B}\widehat{\mathbf{G}}$, where $\mathbf{B}$ is a diagonal $(n \times n)$-matrix given by

$$\mathbf{B} \equiv \begin{pmatrix} \beta_1 & 0 & \ldots & 0 \\ 0 & \beta_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \beta_n \end{pmatrix}.$$

The condition (G.28) guarantees that the linear mapping $\mathbf{B}\widehat{\mathbf{G}}$, hence the best-reply mapping, is still a contraction mapping, so that we still have existence and uniqueness of (interior) equilibrium.

To gain some intuition about how ambitious behavior affects our welfare results, we focus here on the simplest possible case when:

$$\alpha_1 = \ldots = \alpha_n = \alpha > 0, \qquad \beta_1 = \ldots = \beta_n = \beta > 1, \tag{G.29}$$

i.e., all agents are equally productive and equally ambitious. In this case, if $\lambda\beta < 1$, there is a unique interior equilibrium in efforts, which is given by:

$$x^* \equiv x_1^* = \ldots = x_n^* = \frac{1-\lambda}{1-\lambda\beta}\alpha. \tag{G.30}$$

Observe that, in our benchmark model where $\beta = 1$, if $\lambda < 1$, there exists a unique interior equilibrium given by: $x^* = \overline{x}^* = \alpha$. Since $\beta > 1$, we see that by introducing an ambition factor in the utility function (see (G.27)), each individual exerts an effort above the social norm (average effort) $\overline{x}^* = \alpha$ of their direct friends.

Let us now consider welfare issues. The social planner's first-order condition is given by:

$$\mathbf{x} = (1-\lambda)\alpha\mathbf{1} + \lambda\left(\beta\widehat{\mathbf{G}} + \beta\widehat{\mathbf{G}}^T - \beta^2\widehat{\mathbf{G}}^T\widehat{\mathbf{G}}\right)\mathbf{x}. \tag{G.31}$$

Consider first the case of a *regular network*. We have: $\widehat{\mathbf{G}} = \widehat{\mathbf{G}}^T$, and, after some simple algebra, (G.31) implies that the socially optimal effort levels are identical across agents and equal to:

$$x_1^O = \ldots = x_n^O = x^O := \frac{1-\lambda}{1-\lambda\beta+\lambda\beta(\beta-1)}\alpha. \tag{G.32}$$

Comparing (G.32) with (G.30), we see that all agents *over-invest* in equilibrium compared to the social optimum. Indeed, because $\beta > 1$, we have:

$$x^O = \frac{1-\lambda}{1-\lambda\beta+\lambda\beta(\beta-1)}\alpha < \frac{1-\lambda}{1-\lambda\beta}\alpha = x^*.$$

Assume now that the social network $\mathbf{g}$ is *not regular*. In this case, it is still true under certain conditions that agents overinvest in efforts even with identical productivities.

Proposition G6 *Assume that conditions* (G.28)-(G.29) *hold, and that the social network* $\mathbf{g}$ *is not regular. Then:*

(i) *the decentralized equilibrium is never optimal;*

(ii) *the set of agents overinvesting in equilibrium compared to the first best is non-empty;*

(iii) *there exist two threshold values,* $\underline{\lambda}(\beta,\mathbf{g})$ *and* $\overline{\lambda}(\beta,\mathbf{g})$*, satisfying* $0 < \underline{\lambda}(\beta,\mathbf{g}) \leq \overline{\lambda}(\beta,\mathbf{g}) < 1/\beta$*, and such that, if either* $\lambda \leq \underline{\lambda}(\beta,\mathbf{g})$*, or* $\lambda \geq \overline{\lambda}(\beta,\mathbf{g})$*, then all agents overinvest in equilibrium compared to the first best.*

Proof. (i) To prove that $\mathbf{x}^* \neq \mathbf{x}^O$, we will show that, unlike the equilibrium efforts given by (G.30), the first-best efforts are no longer identical across agents. To see this, assume that, on the contrary, there exists $x^O > 0$, such that $\mathbf{x}^O = x^O\mathbf{1}$. Plugging $\mathbf{x}^O = x^O\mathbf{1}$ into (G.31), we get:

$$x^O\mathbf{1} = \left[(1-\lambda)\alpha + \lambda\beta x^O\right]\mathbf{1} + \lambda\beta(1-\beta)x^O\widehat{\mathbf{G}}^T\mathbf{1}. \tag{G.33}$$

This, in turn, implies that $\widehat{\mathbf{G}}^T\mathbf{1}$ is collinear to $\mathbf{1}$, or, equivalently, that $\mathbf{1}^T$ is the left eigenvector of $\widehat{\mathbf{G}}$. Then, by Frobenius-Perron theorem, because $\mathbf{1}^T$ is a positive vector, it must correspond to the principal eigenvalue of $\widehat{\mathbf{G}}$, which is equal to one. But we know that the left eigenvector of $\widehat{\mathbf{G}}$ corresponding to the unitary eigenvalue is unique up to a scalar multiple (by Frobenius-Perron theorem), and is collinear to $(d_1, d_2, \ldots, d_n)$, where d_i is agent i's degree (see (21)). Thus, $\mathbf{g}$ must be a regular network, which contradicts our assumption. This proves part (i).

(ii) The vector $\Delta\mathbf{x} \equiv \mathbf{x}^O - \mathbf{x}^*$ captures the differences between first best efforts and equilibrium efforts. We need to show that $\Delta\mathbf{x}$ has a strictly negative entry.

Using (G.30) – (G.31), we get after some linear algebra:

$$\Delta\mathbf{x} = -\alpha\lambda\beta(\beta-1)\frac{1-\lambda}{1-\lambda\beta}\left[(1-\lambda)\mathbf{I} + \lambda\left(\mathbf{I}-\beta\widehat{\mathbf{G}}\right)^T\left(\mathbf{I}-\beta\widehat{\mathbf{G}}\right)\right]^{-1}\mathbf{G}^T\mathbf{1}. \quad \text{(G.34)}$$

Because the matrix $(1-\lambda)\mathbf{I} + \lambda\left(\mathbf{I}-\beta\widehat{\mathbf{G}}\right)^T\left(\mathbf{I}-\beta\widehat{\mathbf{G}}\right)$ is positive definite,[5] its inverse is well defined and is also positive definite. Therefore, multiplying both parts of (G.34) by the positive row-vector $\mathbf{1}^T\widehat{\mathbf{G}}$, we get:

$$\mathbf{1}^T\widehat{\mathbf{G}}\Delta\mathbf{x} < 0,$$

which implies that $\Delta\mathbf{x}$ has a strictly negative entry. This proves part (ii).

(iii) Consider the limit case when $\lambda \to 0$. In this case, we have:

$$\lim_{\lambda\to 0}\left[(1-\lambda)\mathbf{I} + \lambda\left(\mathbf{I}-\beta\widehat{\mathbf{G}}\right)^T\left(\mathbf{I}-\beta\widehat{\mathbf{G}}\right)\right]^{-1}\mathbf{G}^T\mathbf{1} = \mathbf{G}^T\mathbf{1} > \mathbf{0}.$$

By continuity, the inequality

$$\left[(1-\lambda)\mathbf{I} + \lambda\left(\mathbf{I}-\beta\widehat{\mathbf{G}}\right)^T\left(\mathbf{I}-\beta\widehat{\mathbf{G}}\right)\right]^{-1}\mathbf{G}^T\mathbf{1} > \mathbf{0} \quad \text{(G.35)}$$

must hold for sufficiently small positive values of λ. Define $\underline{\lambda}(\beta, \mathbf{g})$ as follows:

$$\underline{\lambda}(\beta, \mathbf{g}) := \sup\left\{\mu \in \left(0, \frac{1}{\beta}\right) \,\middle|\, \text{(G.35) holds for all } \lambda < \mu\right\}. \quad \text{(G.36)}$$

Combining (G.36) with (G.34), we find that $\Delta\mathbf{x} < \mathbf{0}$ when $\lambda < \underline{\lambda}(\beta, \mathbf{g})$.

[5] The argument is the same as in the proof of part (i) of Lemma A4 in Appendix A.

Next, consider the other extreme case: $\lambda \to 1/\beta$. Using (G.30), we find that, for all $i = 1, 2, \ldots, n$, we have:

$$\lim_{\lambda \to 1/\beta} x_i^* = \infty. \tag{G.37}$$

Using (G.31), we get:

$$\lim_{\lambda \to 1/\beta} \mathbf{x}^O = \left(1 - \frac{1}{\beta}\right) \alpha \left[\left(1 - \frac{1}{\beta}\right)\mathbf{I} + \frac{1}{\beta}\left(\mathbf{I} - \beta\widehat{\mathbf{G}}\right)^T \left(\mathbf{I} - \beta\widehat{\mathbf{G}}\right)\right]^{-1} \mathbf{1}. \tag{G.38}$$

Because the matrix $\left(1 - \frac{1}{\beta}\right)\mathbf{I} + \frac{1}{\beta}\left(\mathbf{I} - \beta\widehat{\mathbf{G}}\right)^T \left(\mathbf{I} - \beta\widehat{\mathbf{G}}\right)$ is positive definite, its inverse is well defined. Thus, it follows from (G.38) that, for all $i = 1, 2, \ldots, n$, agent i's first best effort level x_i^O converges to a finite limit as $\lambda \to 1/\beta$. Comparing (G.38) to (G.37), we get:

$$\lim_{\lambda \to 1/\beta} \Delta x_i = -\infty.$$

Combining this with (G.34), we infer that (G.35) must hold when λ is sufficiently close to $1/\beta$. Define $\overline{\lambda}(\beta, \mathbf{g})$ as follows:

$$\underline{\lambda}(\beta, \mathbf{g}) := \inf \left\{ \mu \in \left(0, \frac{1}{\beta}\right) \;\middle|\; \text{(G.35) holds for all } \lambda > \mu \right\}. \tag{G.39}$$

Putting (G.39) and (G.34) together, we find that $\Delta\mathbf{x} < \mathbf{0}$ when $\lambda > \overline{\lambda}(\beta, \mathbf{g})$. This proves part (*iii*) and completes the proof of the whole proposition. ■

G.5 Network formation

Consider a two-stage game where, in the first stage, agents create links while, in the second stage, they exert effort as in our model.

Assume, for simplicity, that there are two types of agents: high-productive agents with $\alpha = \alpha^H$ and low-productive agents for which $\alpha = \alpha^L$, with $\alpha^H > \alpha^L > 0$.

The equilibrium concept of network formation is *Pairwise Nash Equilibrium*, as it is standard in this literature (see Bloch and Jackson, 2006, for a precise definition).[6] We follow the standard agreement so that each agents only breaks a link if that makes her strictly better off while a new link is created if and only if it makes both agents weakly better off. Define a *completely homophilous* network $\mathbf{g}^h$ as a network

[6] As Bloch and Jackson (2006) put it: "Pairwise Nash Equilibrium is a refinement of both pairwise stability and Nash equilibrium, where one requires that a network be immune to the formation of a new link by any two agents, and the deletion of *any number of links* by any individual agent".

for which agents of different types are not linked to each other and all agents of the same type form a complete network. It is equivalent to a complete bi-partite network where links only exist between agents of the same type. The following figure displays a completely homophilous network with 10 agents, 5 of each type:

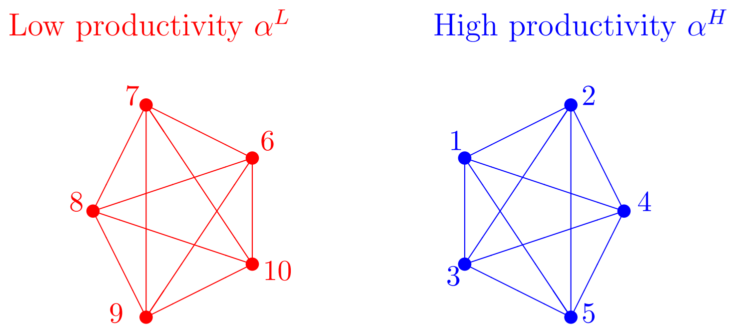

Figure G6: Completely homophilous network

We have the following result:

Proposition G7 *Assume that there are two types of agents with productivities α^L and α^H, where $\alpha^H > \alpha^L > 0$. Assume also that creating a link is costless and so is severing a link. Then:*

(i) *Consider the* ***local-aggregate model****, where each agent's utility is given by* (4)*. Then, the unique Pairwise Nash Equilibrium is the complete network where all agents of any type are linked to each other.*

(ii) *Consider the* ***local-average model****, where each agent's utility is given by* (3)*. Then, the unique Pairwise Nash Equilibrium is the completely homophilous network, which is a network of two disconnected components, where in each component all agents of the same type form a complete network.*

Proof.

(i) Consider the **local-aggregate model**. If there is no cost of forming a link, then because of strategic complementarities and the local-aggregate nature of the model, it is always beneficial for an individual (of any type) to form a link with any other individual (of any type). As a result, the only equilibrium network is the one for which everybody is linked to everybody else (see König et al., 2014).

(ii) Consider now the **local-average model**. In order to prove that the only Pairwise Nash Equilibrium is the completely homophilous network, we proceed in two steps:

[(iia)] We first show that the completely homophilous network $\mathbf{g}^h$ is a pairwise Nash equilibrium;

[(iib)] Then, we show that no other network $\mathbf{g} \neq \mathbf{g}^h$ is a pairwise Nash equilibrium.

(iia) Let us first show that $\mathbf{g}^h$ is pairwise Nash equilibrium. It is readily verified that, when $\mathbf{g} = \mathbf{g}^h$, each agent i, whatever her type, exerts effort $x_i^*(\mathbf{g}^h) = \alpha_i$, follows the social norm $\overline{x}_i^*(\mathbf{g}^h) = \alpha_i$, and gains utility level $U_i^*(\mathbf{g}^h) = \alpha_i^2/2$.

Assume that agent i chooses to break a link with her neighbor j (who must be of the same type as i for such a link to exist). It is readily verified that:

$$U_i^*\left(\mathbf{g}^h - ij\right) = \alpha_i^2/2 = U_i^*(\mathbf{g}^h).$$

Thus, breaking the link with j does not make i strictly better off, hence, there is no reason to do so.

It remains to show that no one is willing to form a link. Assume that, on the contrary, there is an agent i who would benefit from forming a new link with another agent, j. Note that, because i is already linked to everyone who has same type, i and j must be of a different type. Then, i's equilibrium utility becomes:

$$U_i^*(\mathbf{g}^h + ij) = \frac{\alpha_i^2 + 2\theta\alpha_i\overline{x}_i^*(\mathbf{g}^h + ij) - \theta\overline{x}_i^*(\mathbf{g}^h + ij)^2}{2(1+\theta)}. \tag{G.40}$$

Using (G.40), it is readily verified that:

$$U_i^*(\mathbf{g}^h + ij) < \frac{\alpha_i^2}{2} \quad \Longleftrightarrow \quad \overline{x}_i^*(\mathbf{g}^h + ij) \neq \alpha_i.$$

Hence, to show that $U_i^*(\mathbf{g}^h + ij) < U_i^*(\mathbf{g}^h) = \alpha_i^2/2$, it remains to prove that $\overline{x}_i^*(\mathbf{g}^h + ij) \neq \alpha_i$. To see this, observe that, because the network $\mathbf{g}^h + ij$ is connected, its row-normalized adjacency matrix $\widehat{\mathbf{G}}(\mathbf{g}^h + ij)$ is irreducible. Hence, there is a path linking i with any other agent, i.e., for any agent $r \neq i$, there exists an integer $k > 0$, such that:

$$\widehat{g}_{ir}^{[k]}(\mathbf{g}^h + ij) > 0.$$

Combining this with Proposition 1, we find that agent i's new equilibrium social norm $\overline{x}_i^*(\mathbf{g}^h + ij)$ is a strict convex combination (i.e., with strictly positive weights) of $\alpha_1, \alpha_2, \ldots, \alpha_n$. Hence,

$$\alpha^L < \overline{x}_i^*(\mathbf{g}^h + ij) < \alpha^H,$$

which implies $\overline{x}_i^*(\mathbf{g}^h + ij) \neq \alpha_i \in \left\{\alpha^L, \alpha^H\right\}$. This proves (i).

(iib) We now prove that any $\mathbf{g} \neq \mathbf{g}^h$ is not a pairwise Nash equilibrium. Two cases may arise.

Case 1: cross-type links. Assume first that $\mathbf{g}$ has a cross-type link, i.e., there are i and j, such that $\alpha_i \neq \alpha_j$, but $g_{ij} = 1$. Let $\mathbf{g}^{ij}$ be the component of $\mathbf{g}$ agents i and j both belong to. We will show that dropping all links and remaining alone makes agent i strictly better off. Indeed, because agents who are not in $\mathbf{g}^{ij}$ affect neither i's effort nor her well-being, i's equilibrium untility level is given by

$$U_i^*(\mathbf{g}) = U_i^*\left(\mathbf{g}^{ij}\right) = \frac{\alpha_i^2 + 2\theta\alpha_i\overline{x}_i^*\left(\mathbf{g}^{ij}\right) - \theta\left[\overline{x}_i^*\left(\mathbf{g}^{ij}\right)\right]^2}{2(1+\theta)}. \tag{G.41}$$

We now show that $U_i^*(\mathbf{g}) < \alpha_i^2/2$, where $\alpha_i^2/2$ is the utility level gained by agent i in the absence of neighbors. As implied by (G.41), we have: $U_i^*(\mathbf{g}) \leq \alpha_i^2/2$, and

$$U_i^*(\mathbf{g}) = \frac{\alpha_i^2}{2} \quad \Longleftrightarrow \quad \overline{x}_i^*\left(\mathbf{g}^{ij}\right) = \alpha_i.$$

Because $\mathbf{g}^{ij}$ is connected, we find, using the same argument as in the proof of (iia), that $\overline{x}_i^*\left(\mathbf{g}^{ij}\right) \neq \alpha_i$. Hence, $U_i^*(\mathbf{g}) < \alpha_i^2/2$, i.e., breaking unilaterally all the links (including one with j) makes i strictly better off. Hence, no network with links between agents of different types is a pairwise Nash equilibrium.

Case 2: no cross-type links. Assume now that, for any (i, j), we have: $\alpha_i \neq \alpha_j \Rightarrow g_{ij} = 0$. We need to show that, if there are i and $j \neq i$, such that $\alpha_i = \alpha_j$, but $g_{ij} = 0$, then $\mathbf{g}$ is not pairwise Nash equilibrium. Let $\mathbf{g}^i$ and $\mathbf{g}^j$ be, respectively, components of $\mathbf{g}$ agents i and j belong to. Because $\mathbf{g}$ does not have cross-type links, it must be true that all agents in $\mathbf{g}^i$ have the same type, and so do all agents in $\mathbf{g}^j$. Moreover, because $\alpha_i = \alpha_j$, all agents $\mathbf{g}^i + \mathbf{g}^j$ have the same type. Then, this implies:

$$\begin{aligned} x_i^*(\mathbf{g}) &= x_i^*\left(\mathbf{g}^i\right) = \alpha_i, & x_i^*(\mathbf{g} + ij) &= x_i^*\left(\mathbf{g}^i + \mathbf{g}^j\right) = \alpha_i; \\ \overline{x}_i^*(\mathbf{g}) &= \overline{x}_i^*\left(\mathbf{g}^i\right) = \alpha_i, & \overline{x}_i^*(\mathbf{g} + ij) &= \overline{x}_i^*\left(\mathbf{g}^i + \mathbf{g}^j\right) = \alpha_i; \\ U_i^*(\mathbf{g}) &= U_i^*\left(\mathbf{g}^i\right) = \frac{\alpha_i^2}{2}, & U_i^*(\mathbf{g} + ij) &= U_i^*\left(\mathbf{g}^i + \mathbf{g}^j\right) = \frac{\alpha_i^2}{2}. \end{aligned}$$

Thus, adding a link between i and j does not change equilibrium utilities. By the agreement above, if agents are both indifferent between creating or not creating a link, they choose to create it. Hence, $\mathbf{g} \neq \mathbf{g}^h$ is not a pairwise Nash equilibrium in the network formation game. This completes the proof. ■

References for the Online Appendix

Bloch, F. and M.O. Jackson (2006), "Definitions of equilibrium in network formation games," *International Journal of Game Theory* 34, 305–318.

Courant, R. and John, F. (2012), *Introduction to Calculus and Analysis I*, Berlin: Springer.

Golub, B. and Jackson, M.O. (2010), "Naïve learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics* 2, 112–149.

König, M., Tessone, C. and Y. Zenou (2014), "Nestedness in networks: A theoretical model and some applications," *Theoretical Economics* 9, 695–752.

Ushchev, P. and Y. Zenou (2018), "Price competition in product variety networks," *Games and Economic Behavior* 110, 226–247.