DISCUSSION PAPER SERIES

DP13250

SOCIAL NORMS IN NETWORKS

Philip Ushchev and Yves Zenou

PUBLIC ECONOMICS

ISSN 0265-8003

SOCIAL NORMS IN NETWORKS

Philip Ushchev and Yves Zenou

Discussion Paper DP13250
Published 16 October 2018
Submitted 16 October 2018

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

SOCIAL NORMS IN NETWORKS

Abstract

Although the linear-in-means model is the workhorse model in empirical work on peer effects, its theoretical properties are understudied. In this paper, we investigate how social norms affect individual effort, aggregate effort, and welfare. While individual productivity always positively affects own effort and utility, we show that taste for conformity has an ambiguous effect on individual outcomes and depends on whether an individual is above or below her own social norm. Equilibria are usually inefficient and, to restore the first best, the planner subsidizes (taxes) agents whose neighbors make efforts above (below) the social norms in equilibrium. Thus, provision of more subsidies to more central agents is not necessarily efficient.

JEL Classification: D85, J15, Z13

Keywords: networks, Social norms, welfare

Philip Ushchev - ph.ushchev@gmail.com
National Research University Higher School of Economics

Yves Zenou - yves.zenou@monash.edu
Monash University and CEPR

Acknowledgements
We thank the participants at the Brown Bag seminar at Monash University, especially Arthur Campbell and Michelle Rendall. Philip Ushchev gratefully acknowledges that the study has been funded by the Russian Academic Excellence Project '5-100'.

Social Norms in Networks*

Philip Ushchev† Yves Zenou‡

October 16, 2018

Abstract

Although the linear-in-means model is the workhorse model in empirical work on peer effects, its theoretical properties are understudied. In this paper, we investigate how social norms affect individual effort, aggregate effort, and welfare. While individual productivity always positively affects own effort and utility, we show that taste for conformity has an ambiguous effect on individual outcomes and depends on whether an individual is above or below her own social norm. Equilibria are usually inefficient and, to restore the first best, the planner subsidizes (taxes) agents whose neighbors make efforts above (below) the social norms in equilibrium. Thus, provision of more subsidies to more central agents is not necessarily efficient.

Keywords: Networks, social norms, local average model, welfare, Markov chain.
JEL Classification: D85, J15, Z13.

---

*We thank the participants at the Brown Bag seminar at Monash University, especially Arthur Campbell and Michelle Rendall. Philip Ushchev gratefully acknowledges that the study has been 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 (Epple and Romano, 2011; Sacerdote, 2011), crime (Ludwig et al., 2001; 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; Cawley et al., 2017), environmentally friendly behavior (Brekke et al., 2010; Czajkowski et al., 2017), 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 in 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. This is why this game is often referred to as the *local-average model*.

[1]For example, in relation to crime, x_{ig} is the criminal behavior 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]Equation (1) is difficult to estimate because of the endogeneity of $\frac{\theta}{(N_g-1)} \sum_{j=1,j\neq i}^{N_g} x_{jg}$. In particular, 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. Thus, 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).

Surprisingly, the theoretical properties of this model in terms of comparative statics, welfare, and policies have not been investigated. On the contrary, the literature on games on networks[4] (Ballester et al., 2006; Jackson and Zenou, 2015; Bramoullé et al., 2014; 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. We are the first to study the theoretical properties of the local-average model and show that the properties are very different to those of the local-aggregate model.[7]

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, and aggregate effort are the weighted sums of productivity, whereby the weights are non-linear functions of the taste for conformity. In order to understand these results, we first 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 complete conformism, all agents make 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.

[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 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). In their seminal study, Bramoullé et al. (2009) where the first to provide conditions for identification in the local-average model.

[7]Observe that, in this study, we are interested only in *positive* peer effects, which is why we compare the local-aggregate model to the local-average one, as both are games with *strategic complementarities*; that is, an increase in the effort of a neighbor increases the marginal utility of own effort. The main difference between these two models is that, in the local-aggregate model, the social interaction effect is given by $\theta \sum_{j=1,j\neq i}^{N_g} x_{jg}$, while in the local-average model, it is equal to $\frac{\theta}{(N_g-1)} \sum_{j=1,j\neq i}^{N_g} x_{jg}$. This turns out to have very different implications since, in the former model, an increase in θ always has a positive impact on the marginal utility of own effort while, in the latter model, it depends on whether the effort is above or below the social norm. See (10) below.

Second, to understand the abovementioned results further, we reinterpret the model in a probabilistic way whereby the adjacency matrix of the network is now a transition probability matrix of a Markov chain with n finite states, in which each state is the location of an agent in the network, and vice versa. In this model, each agent follows a random walk whereby, with some probability, she exerts an effort level equal to her productivity while, with complementary probability, she mimics the behavior of one of her neighbors. We show that the network game and the Markov chain model are observationally equivalent, even though, in the former, agents are rational and maximize their utility function under perfect information while, in the latter, agents behave like stochastic automata. The Markov chain model provides some intuition of our results and turns out to be very useful when deriving and proving complicated comparative statics and welfare results. In particular, we show that the case of total conformism in our game corresponds to the stationary distribution of the states (location of agents in the network) in the Markov chain model.

Third, we provide clear-cut comparative statics of equilibrium efforts with respect to changes in productivity. We show that an increase in own productivity always increases own equilibrium effort and utility. However, an increase in an individual's neighbor's productivity does not always make this person better off. This is only the case when she is sufficiently productive.

We also study the impact of the taste for conformity λ on equilibrium effort. There are, first, two direct effects at work, a *direct negative productivity effect* and a *direct positive social-norm effect*, since an increase in λ decreases own productivity and increases own social norm, as individuals pay more attention to their neighbors than to themselves. However, there is also an *indirect social-norm effect*, which reflects the fact that social norms are endogenous, implying general equilibrium effects. Indeed, when λ increases, individual i changes her effort level (direct effect), which affects her own social norm. Thus, for individual j, a neighbor of i, her own social norm changes, and thus, she changes her effort, which, in turn, affects the effort of individual i (indirect effect). As a result, a complex interplay between these three effects may result in a non-monotonic relationship between the taste for conformity and the individual equilibrium effort. Whether an individual is above or below her own social norm is key for understanding the shape of this relationship. We also show that, in regular networks, aggregate effort remains neutral to changes in the taste for conformity and is equal to aggregate productivity.

Finally, 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 (above) her own social norm, an increase in i's effort increases the social norm of j, which has a negative (positive) impact on j's conformist utility, because j's effort is now further away from (closer to) her own social norm. In this case, agent i exerts a negative (positive) externality on her neighbor j. To restore the first best, the planner taxes (subsidizes) agents who exert negative (positive) externalities on their neighbors. This is very different from the policy implications of the local-aggregate model whereby 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. Fo r 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.

Other researchers have studied the local-average (conformist) model in network games.[8] Patacchini and Zenou (2012) and Liu et al. (2014) characterized Nash equilibrium and showed that it exists and is unique; Blume et al. (2015) introduced imperfect information; Boucher (2016) embedded the local-average model into a network formation model; and Olcina et al. (2017) embedded it into a learning model.[9] To the best of our knowledge, we are the first to study the properties of the local-average network model, that is, to establish the equivalence between the local-average and the Markov-chain model. We are also the first to derive the comparative-statics properties of the model, to determine the first best, and to characterize the subsidy and tax policies that restore the first best.

The rest of the paper unfolds as follows. In Section 2, we study the local-average model, determine the condition for its existence and uniqueness, and provide an equivalent model in terms of Markov chain. 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. In Section 5, we examine the policy implications of our

[8]Some literature has 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); Fershtman and Weiss (1998).

[9]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).

results. Finally, Section 6 concludes. All proofs are found in Appendix A, while, in Appendix B, we provide additional results and examples.

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 if $g_{ij} = 1$, then $g_{ji} = 1$, i.e., the network is *undirected*. Furthermore, there are 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, $d_i := \sum_{j=1}^{n} g_{ij}$. For each $\lambda \in [0,1)$, define the following $(n \times n)$-matrix $\widehat{\mathbf{M}} = [\widehat{m}_{ij}]$ as

$$\widehat{\mathbf{M}} := (1-\lambda)\left(\mathbf{I} - \lambda\widehat{\mathbf{G}}\right)^{-1}. \tag{2}$$

Because $\widehat{\mathbf{G}}$ is row-normalized and $0 \leq \lambda < 1$, matrix $\widehat{\mathbf{M}}$ is well defined and can be represented by the Neumann series:[10]

$$\widehat{\mathbf{M}} = (1-\lambda)\sum_{k=0}^{\infty} \lambda^k \widehat{\mathbf{G}}^k, \tag{3}$$

which implies that $\widehat{\mathbf{M}}$ is itself row-normalized. Note that (3) is reminiscent of the Bonacich centrality (Bonacich, 1987; Ballester et al., 2006). We have:[11]

$$\widehat{m}_{ij} = (1-\lambda)\sum_{k=0}^{\infty} \lambda^k \widehat{g}_{ij}^{[k]} = \frac{\widehat{g}_{ij}^{[0]} + \widehat{g}_{ij}^{[1]}\lambda + \widehat{g}_{ij}^{[2]}\lambda^2 + ...}{1 + \lambda + \lambda^2 + ...}, \tag{4}$$

which counts all the walks from i to j with a decreasing weight of λ (so that a walk of length k is weighted by λ^k) and discounted by the infinite sum of λ^k.

[10] 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.

[11] Observe that $\widehat{g}_{ij}^{[0]} = 1$, if $i = j$, and 0, otherwise.

2.2 Preferences

Denote by x_i the effort level that each agent i exerts, and by $\mathbf{x}_{-i}$ the vector of effort levels exerted by the other $n-1$ agents in the network. Individual i's *social norm*, $\overline{x}_i$, is defined as the average effort of her neighbors, that is,

$$\overline{x}_i := \sum_{j=1}^{n} \widehat{g}_{ij} x_j \tag{5}$$

Agent i's utility function is given by

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

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

The utility function (6) 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.[12] This is very different from the local-aggregate model (Ballester et al., 2006) where what matters is the sum of efforts of one's peers. Indeed, for each agent i, the utility in the local-aggregate model is equal to:

$$U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}) = \alpha_i x_i - \frac{1}{2} x_i^2 + \theta \sum_{j=1}^{n} g_{ij} x_i x_j. \tag{7}$$

In (7), the aggregate effort of the peers, $\sum_{j=1}^{n} g_{ij} x_j$ positively affects the utility of each agent i.

Let us return to utility function (6). For the sake of convenience, we parameterize the taste for conformity in a slightly different way by having

$$\lambda := \frac{\theta}{1+\theta}, \qquad 0 \leq \lambda < 1. \tag{8}$$

Then, (6) can be written as

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

[12]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).

The two parameterizations, (6) and (9), are clearly equivalent, since, as seen from (8), λ is a monotone transformation of θ. Note that, while individual productivity α_i, $i = 1, \ldots, n$, is agent specific, the degree λ of conformity is the same across agents, and hence, can be viewed as a characteristic of society as a whole.

Let us understand the properties of this utility function. First, if i and j are neighbors, we have

$$\frac{\partial U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})}{\partial x_j} = \frac{\lambda}{(1-\lambda)} \widehat{g}_{ij} \left(x_i - \overline{x}_i\right) \gtreqless 0 \iff x_i \gtreqless \overline{x}_i. \tag{10}$$

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

$$\frac{\partial^2 U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})}{\partial x_i x_j} = \left(\frac{\lambda}{1-\lambda}\right) \sum_j \widehat{g}_{ji} > 0, \tag{11}$$

which means that the higher is the effort of an individual's friend, 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:

$$\frac{\partial^2 U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})}{\partial x_i \partial \lambda} = -\frac{(x_i - \overline{x}_i)}{(1-\lambda)^2} \lesseqgtr 0 \iff x_i \gtreqless \overline{x}_i. \tag{12}$$

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 distance 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 individual i's effort x_i and own productivity α_i and her peer's productivity α_j, $j \neq i$, are given by

$$\frac{\partial^2 U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})}{\partial x_i \partial \alpha_i} = 1 > 0, \qquad \frac{\partial^2 U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})}{\partial x_i \partial \alpha_j} = 0, \tag{13}$$

respectively. This means that an increase in own productivity increases the impact of own effort on utility but a change in the productivity of an individual's neighbors does not affect the individual's marginal utility of effort. These cross-derivatives are useful for interpreting our comparative statics results in terms of α.

There are two extreme cases of interest to our study: *pure individualism* ($\lambda = 0$) so that i's utility depends only on own productivity α_i, and *total conformism* ($\lambda \to 1$), so that i's utility depends only on others' behavior.

To summarize, the utility function (6)—equivalently, (9)—is the standard way that 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., 2018). 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 Section 4.

2.3 Nash equilibrium

Each individual i chooses x_i to maximize (9) 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:

$$x_i = (1-\lambda)\alpha_i + \lambda \overline{x}_i. \tag{14}$$

After some normalizations, it should be clear that (14) 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 (14) with the definition (5) 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

$$\mathbf{x} = (1-\lambda)\boldsymbol{\alpha} + \lambda \widehat{\mathbf{G}}\mathbf{x}, \tag{15}$$

where $\boldsymbol{\alpha} := (\alpha_1, \ldots, \alpha_n)^T$ is the productivity vector.[13]

[13]Observe that the linear-in-means model (1) is also called 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

Proposition 1 (Equilibrium efforts, norms, and utilities)

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

$$\mathbf{x}^* = \widehat{\mathbf{M}}\boldsymbol{\alpha} = (1-\lambda)\sum_{k=0}^{\infty}\lambda^k \widehat{\mathbf{G}}^k \boldsymbol{\alpha}. \tag{16}$$

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

$$\overline{\mathbf{x}}^* = \widehat{\mathbf{G}}\widehat{\mathbf{M}}\boldsymbol{\alpha} = (1-\lambda)\sum_{k=0}^{\infty}\lambda^k\, \widehat{\mathbf{G}}^{k+1} \boldsymbol{\alpha}. \tag{17}$$

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

$$U_i^*(\boldsymbol{\alpha}, \lambda, \mathbf{g}) = \frac{1}{2}\left[\alpha_i^2 - \frac{1}{\lambda}\left(\alpha_i - \sum_{j=1}^{n}\widehat{m}_{ij}\alpha_j\right)^2\right]. \tag{18}$$

The equilibrium effort in equation (16) has the same flavor as the Bonacich centrality representation of the Nash equilibrium in the *local aggregate model* (Ballester et al., 2006). There are, however, two main differences. First, unlike the local aggregate model, there is *no need to impose any conditions on* $\frac{\lambda}{1-\lambda}$ (except that $\frac{\lambda}{1-\lambda} > 0$) to guarantee that the Nash equilibrium exists, is unique, and is interior. The second main difference is that in the local-average model, the network is represented by its row-normalized adjacency matrix $\widehat{\mathbf{G}}$, not by its $\{0,1\}$ adjacency matrix $\mathbf{G}$. This assumption is far from being innocuous, since it implies that the two models differ enormously from each other in how the network structure $\mathbf{g}$ maps into the structure of equilibria. For example, as pointed out by Patacchini and Zenou (2012), if agents are *ex ante homogeneous* (i.e., $\alpha_i = \alpha_j$ for any $i, j = 1, 2, \ldots, n$), then, regardless of the network structure, *the equilibrium effort levels are the same across agents* (i.e.,

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, in which the autoregression parameter λ is expected to be strictly less than 1 in modulus; see Hamilton, 1994). Equation (15) is clearly equivalent to the spatial-autoregressive model and it gives a microfoundation of the SAR model via the utility function (6) or (9).

$x_i^* = x_j^*$ for any $i, j = 1, 2, \ldots, n$). This result displays significant differences between the local-average and local-aggregate models. In the former model, when productivity is the same, the outcome does not depend on the network structure, and not even on the number of agents. By contrast, in the latter model, efforts are proportional to the Bonacich centralities of agents, which depend crucially on the agents' positions in the network, no matter whether productivity is the same or not. Thus, in the local aggregate model, the network structure $\mathbf{g}$ has a substantial *own effect*, whereas, in our model, the network affects only the equilibrium via the *cross-effect* with the vector of productivity.

The last result of Proposition 1 gives 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 $i \in N$, we have*

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

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 productivity in the network. The weights $\widehat{m}_{ij}/(1-\widehat{m}_{ii})$ count all the walks from i to j conditional on not having a self-loop. For example, in a star network, if the central agent has the highest productivity in the network, 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 understand some important results in Sections 3 and 4.

2.4 A probabilistic interpretation

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 (4) 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 (4), 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:

$$\mathbb{E}\left[\mathbf{X}\right] = \widehat{\mathbf{M}}\boldsymbol{\alpha} = (1-\lambda)\sum_{k=0}^{\infty}\lambda^k \widehat{\mathbf{G}}^k\boldsymbol{\alpha}, \tag{20}$$

where $\mathbf{X} := (X_1, X_2, \ldots, X_n)^T$ is the vector of efforts in the probabilistic model. By comparing (20) and (16), we see that:

$$\mathbf{x}^* = \mathbb{E}\left[\mathbf{X}\right], \tag{21}$$

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.

We now use the observational equivalence between the two models to compare the outcomes generated by perfect individualism ($\lambda = 0$) and total conformism ($\lambda = 1$). 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

$$\boldsymbol{\pi}\widehat{\mathbf{G}} = \boldsymbol{\pi}, \tag{22}$$

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

$$\pi_i = \frac{d_i}{\sum_{j=1}^{n} d_j}, \quad \text{for all } i = 1, 2, \ldots n, \tag{23}$$

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*:

$$\lim_{k \to \infty} \widehat{\mathbf{G}}^k = \begin{pmatrix} \boldsymbol{\pi} \\ \boldsymbol{\pi} \\ \vdots \\ \boldsymbol{\pi} \end{pmatrix}. \tag{24}$$

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 2 *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 $m = 1$.*

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.

Let us now consider the limit case when $\lambda \to 1$, which means that the society is *totally conformist*, that is, conformity is infinite so that only peers matter.

Proposition 2 (Totally conformist agents) *For any network structure, regardless off whether the corresponding Markov chain is ergodic or not, we have*

$$\lim_{\lambda \to 1} x_i^*(\lambda) = \boldsymbol{\pi}\boldsymbol{\alpha} = \sum_{j=1}^{n} \pi_j \alpha_j, \qquad \textit{for all} \qquad i = 1, \ldots, n. \tag{25}$$

Proposition 2 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 understand the totally conformist society, which is otherwise difficult to characterize.

For example, for a star network with n agents and for which agent 1 is the star, the following is easily verified:

$$\pi_1 = \frac{1}{2}, \qquad \pi_2 = \pi_3 = ... = \pi_n = \frac{1}{2\,(n-1)}$$

This implies that

$$\lim_{\lambda \to 1} \mathbf{x}^*(\lambda) = \frac{1}{2} \begin{pmatrix} \alpha_1 + \frac{1}{(n-1)}\alpha_2 + ... + \frac{1}{(n-1)}\alpha_n \\ \vdots \\ \alpha_1 + \frac{1}{(n-1)}\alpha_2 + ... + \frac{1}{(n-1)}\alpha_n \end{pmatrix}$$

For any regular network with n agents, we have

$$\pi_1 = \pi_2 = ... = \pi_n = \frac{1}{n},$$

which implies that

$$\lim_{\lambda \to 1} \mathbf{x}^*(\lambda) = \frac{1}{n} \begin{pmatrix} \alpha_1 + \alpha_2 + ... + \alpha_n \\ ... \\ \alpha_1 + \alpha_2 + ... + \alpha_n \end{pmatrix} \tag{26}$$

We observe that, depending on the network structure (here the star versus the regular network), each individual effort is different, because $\boldsymbol{\pi}$ changes with the network structure. Let us now compare the pure individualist society ($\lambda \to 0$) and the pure conformist society ($\lambda \to 1$).

Proposition 3 (Individualist versus conformist society)

(*i*) *Individual effort:*

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

(*ii*) *Aggregate effort:*

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

Part (i) of Proposition 3 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 3 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,[14] then perfect conformity increases aggregate effort.

2.5 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)). Basically, one estimates the effect of a group

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

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

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

average's outcome (e.g., the average test score in a classroom or school or the average crime rate in a neighborhood) on the same outcome of an individual belonging to the same group (e.g., own test score in the same classroom or school or own crime rate in the same neighborhood). In other words, the linear-in-means model captures an *average* effect so that peer effects are conceived as an average intra-group externality that identically affects all members of a given group. In reality, the same average effect can have a very different impact on own and aggregate outcome, depending on other moments of the distribution, in particular, the variance.[15] 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.

In order to illustrate this, we provide a simple example that shows how a mean-preserving spread of the productivity impacts own and aggregate outcome. Consider a star-shaped network with $n = 3$ agents in which the individual productivity is given by: $\alpha_1 = 1 + 2t$, $\alpha_2 = 1 - t$ and $\alpha_3 = 1 - t$, where $t \in (-1/2, 1)$.[16] As for locations in the network, let $i = 1$ be the "star" agent, and let $i = 2, 3$ be "periphery" agents, as shown by Figure 1.

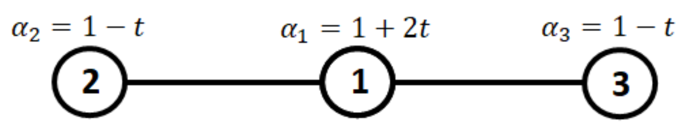

Figure 1: Mean-preserving productivity changes in a chain network with $n = 3$

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

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

[15]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).

[16]This domain is chosen for all individual productivity to remain positive.

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 (16), we can compute the Nash equilibrium in effort for each individual. We obtain

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

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. In addition, 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 2 illustrates these results by contrasting the case in which $\lambda = 0.25 < 0.5$ and when $\lambda = 0.75 > 0.5$.

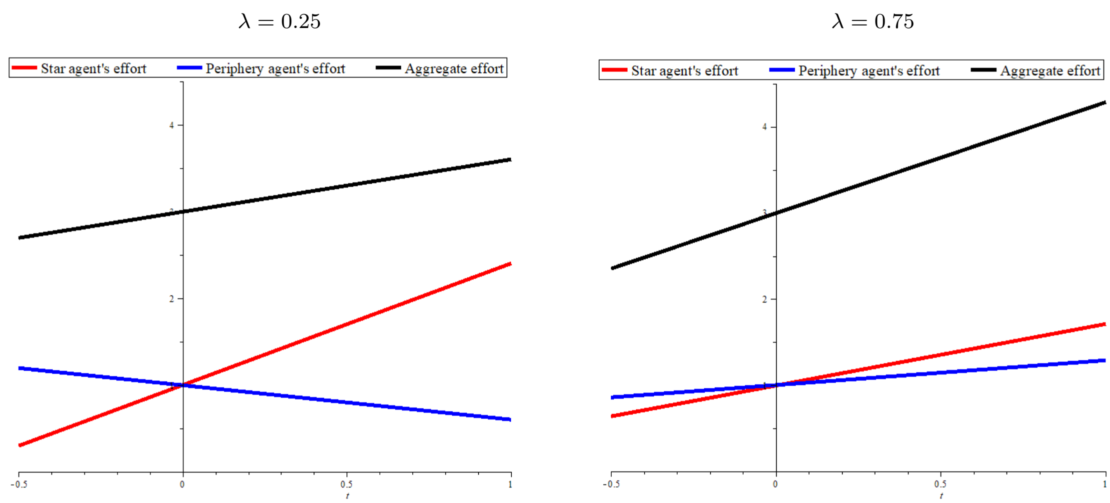

Figure 2: Impact of a mean-preserving spread of productivity on individual and aggregate effort in a star network

As seen from Figure 2, the effort of the central agent, $i = 1$, always increases with t, and so does the aggregate effort. Moreover, the effort of the periphery agents,

$i = 2, 3$, decrease with t when conformity is relatively low (i.e., when $\lambda = 0.25$). This is mainly because their productivity, α_2 and α_3, linearly decreases with t. 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 (given by the effort of the central agent). As a result, the net effect of t on own effort is positive, because the positive indirect effect of an increase in t, which is the peer-effect generated by the star agent (whose productivity increases with t), dominates the negative direct effect of the reduction of own productivity. Therefore, the increase in aggregate effort is steeper when agents are more conformist ($\lambda = 0.75$) than when they are less conformist ($\lambda = 0.25$).

More generally, this example illustrates the fact 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, depending on the value of λ and t. When $t = 0$, all efforts are equal to 1, the average productivity in the network. When $t > 0$ ($t < 0$), the effort of the central agent is always above (below) that of the periphery agents. This difference is larger as λ is higher. 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 dispersion (variance).[17]

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 two key parameters: the productivity vector $\boldsymbol{\alpha}$ and the taste for conformity λ.

3.1 Effect of productivity

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

[17] Observe that the productivity dispersion is not a sufficient statistic either. Indeed, when, for example, $t = 0.25$ and $t = -0.25$, the variance is the same across individuals and equal to $\sigma^2_{\boldsymbol{\alpha}} = 0.125$ but the effort of the central agent is higher (lower) than that of the periphery agents when $t = -0.25$ ($t = 0.25$), for any $\lambda > 0$.

Proposition 4 (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,* $\text{sign}\left[\frac{\partial U_i^*}{\partial \alpha_j}\right] = \text{sign}\left(x_i^* - \overline{x}_i^*\right)$, *or equivalently, using Lemma 1,*

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

The first result is straightforward because, as implied by (16), each x_i^* is a convex combination of productivity. 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 Effects of conformism

How do individual and aggregate efforts change when a society becomes more conformist? To answer this question, we study the comparative statics with respect to the conformity parameter λ.

3.2.1 General network structure

In order to understand the impact of λ on individual effort x_i^*, let us differentiate (14). We obtain

$$dx_i^* = \underbrace{-\alpha_i \, d\lambda}_{\text{productivity effect}} + \underbrace{\overline{x}_i^* \, d\lambda}_{\text{direct social-norm effect}} + \underbrace{\lambda \, (\partial \overline{x}_i^* / \partial \lambda) \, d\lambda}_{\text{indirect social-norm effect}} \tag{27}$$

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 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 (14), (27) 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. This is related to the fact that the effect of λ on the marginal utility of effort is ambiguous and depends on $x_i - \overline{x}_i$ (see (12)). In particular, when λ increases, agents become more conformist, and the distance between x_i and $\overline{x}_i$ matters more.

Recall that $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)$ is the left eigenvector of $\widehat{\mathbf{G}}$ associated with the unitary eigenvalue. $\boldsymbol{\pi}$ is also the (relative) degree of each individual in the network. We have 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*

$$\frac{\partial x_i^*}{\partial \lambda} \gtreqless 0 \iff \alpha_i \lesseqgtr \sum_{j=1}^{n} \widehat{g}_{ij}\alpha_j \tag{28}$$

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

$$\sum_{j=1}^{n} \pi_j \alpha_j \leq \alpha_i < \sum_{j=1}^{n} \widehat{g}_{ij}\alpha_j. \tag{29}$$

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

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

$$\sum_{j=1}^{n} \pi_j \alpha_j \geq \alpha_i > \sum_{j=1}^{n} \widehat{g}_{ij}\alpha_j. \tag{30}$$

Then, agent i*'s individual effort* $x_i^*(\lambda)$ *varies non-monotonically with* λ *and 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 (28) 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.

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 (7)) 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 (6) or (9)) 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 5 below.

Parts (*ii*) and (*iii*) of Proposition 5 provide sufficient (but not necessary) conditions for x_i^* to vary *non-monotonically* with λ.[18] 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 (29) and (30) 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 3, 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.[19]

[18] In sub-sections 3.2.2 and 3.2.3, we give sharper conditions for some specific types of network structures.

[19] 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 $\alpha_6 = 0.5 = \alpha_{-6}$.

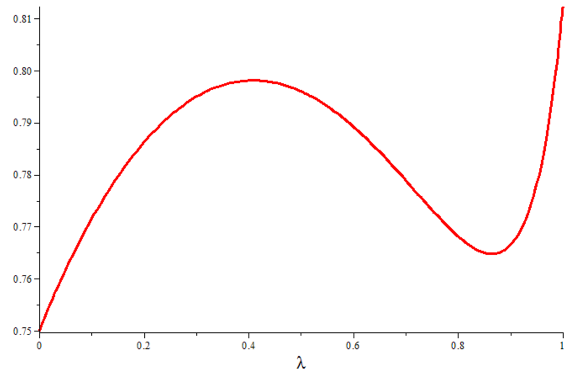

Figure 3: Non-monotonic effect of λ on individual effort for a chain network with $n = 13$

In Proposition 5, we show that the impact of λ on effort is very complex and difficult to sign. In Corollary B.1 in Appendix B.1, we show that the same non-monotonicity results as in Proposition 5 for the aggregate effort, which is an important aspect of this model. 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.

We now consider specific networks in which we can provide more intuition and results on the effect of λ on both individual and aggregate effort.

3.2.2 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 6 (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.

Proposition 6 provides a more precise description of the impact of an increase in λ on x_i^* for the 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.

3.2.3 Regular networks

Let us now 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 7 (Regular networks) *In regular networks, the aggregate effort does not vary with λ, that is,*

$$\sum_{i=1}^{n} x_i^* = \sum_{i=1}^{n} \alpha_i \tag{31}$$

Proposition 7 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. 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 Appendix B.2, we illustrate this result by means of a circular network (which is a regular network of valency $r = 2$). When we rewire the links in this network without changing the network topology, we also 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.

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

$$\mathcal{W} := \sum_{i=1,2,\ldots,n} U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}). \tag{32}$$

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 8 (First best)

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

$$x_i = (1-\lambda)\,\alpha_i + \lambda\,\overline{x}_i + \lambda \sum_{j=1}^{n} \widehat{g}_{ji}\,(x_j - \overline{x}_j)\,, \tag{33}$$

or, in matrix form,

$$\mathbf{x} = (1-\lambda)\,\boldsymbol{\alpha} + \lambda\widehat{\mathbf{G}}\mathbf{x} + \lambda\widehat{\mathbf{G}}^T\left(\mathbf{I} - \widehat{\mathbf{G}}\right)\mathbf{x}. \tag{34}$$

(*ii*) *For the Nash equilibrium in an effort 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:*

$$\widehat{\mathbf{G}}^T(\mathbf{I} - \widehat{\mathbf{G}})\widehat{\mathbf{M}}\boldsymbol{\alpha} = \mathbf{0}. \tag{35}$$

(*iii*) *Moreover, for any network,*

$$\sum_{i=1}^{n} x_i^O = \sum_{i=1}^{n} \alpha_i. \tag{36}$$

Part (*i*) of Proposition 8 clearly shows the difference in effort between the Nash equilibrium (see (14)) and the first best (see (33)). 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 8), 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 $\frac{\partial U_j}{\partial \overline{x}_j} = \left(\frac{\lambda}{1-\lambda}\right)(x_j - \overline{x}_j) \gtreqless 0$. 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 (11)), agents can extert 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 4.4 below.

Part (*ii*) of Proposition 8 gives an exact condition for 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 we show below. Finally, in part (*iii*), we show 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 that, for *regular networks*, using (31), 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.

Remark 3 *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.

4.2 Examples

We now illustrate condition (35) 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:

$$\alpha_1 = \frac{1}{n-1} \sum_{j=1}^{n} \alpha_j. \tag{37}$$

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*[20] is the same, that is,

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

[20]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.

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 4 *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*

$$\overline{\alpha}_r := \frac{1}{|V_r|} \sum_{k \in V_r} \alpha_k. \tag{38}$$

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.

4.3 First best in a sufficiently conformist society

We now focus on the case in which λ, the taste for conformity, is sufficiently large.

Proposition 9 (First best in a sufficiently conformist society)

(*i*) *In a perfectly conformist society,*

$$\lim_{\lambda \to 1} x_i^O = \frac{1}{n} \sum_{j=1}^{n} \alpha_j, \qquad \textit{for all } i = 1, 2, \ldots, n \tag{39}$$

(*ii*) *Moreover, if* $\sum_{j=1}^n \pi_j \alpha_j < \frac{1}{n} \sum_{j=1}^n \alpha_j$ $\left(\sum_{j=1}^n \pi_j \alpha_j > \frac{1}{n} \sum_{j=1}^n \alpha_j\right)$*, 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.*

The first 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 level equal to the average productivity in the network. This implies that, unless the network is regular, the equilibrium in effort is never optimal when $\lambda \to 1$.

The second result combines (39) and (25) to show that, when λ, the taste for conformity is high enough, and then if the average productivity in the network is smaller (greater) than the weighted productivity in the network, whereby the weights are the degrees (number of links) of the agents, then the agents overprovide (underprovide) effort in equilibrium compared to the first best. Recall that (see footnote 14)

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

where $\mathrm{Corr}\left(\boldsymbol{\pi}, \boldsymbol{\alpha}\right)$ is the correlation between $\boldsymbol{\pi}$ and $\boldsymbol{\alpha}$. This result implies that more central agents make higher (lower) effort and exert stronger externalities on their neighbors. As a result, all agents overprovide (underprovide) 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 part (*ii*) of Proposition 9 to hold (see, in particular, footnote 21).

4.4 Restoring the first best

Let us return to the general case in which λ can take any value and assume that condition (35) 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,

$$U_i^{S_i^O} = \left(\alpha_i + S_i^O \right) x_i - \frac{1}{2} x_i^2 - \frac{1}{2} \left(\frac{\lambda}{1-\lambda} \right) \left(x_i - \overline{x}_i \right)^2 \qquad (40)$$

Observe that, when each agent i chooses x_i that maximizes (40), 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 10 (Subsidies) *The first best is restored if the social planner gives to each agent i the following tax/subsidy per unit of effort:*

$$S_i^O = \frac{\lambda}{(1-\lambda)} \sum_{j \neq i} \widehat{g}_{ji} \left(x_j^O - \overline{x}_j^O \right) \tag{41}$$

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 subsidizse (resp. taxes) agents *whose neighbors make efforts above (resp. 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 1 is the star. Let us show that agents who are taxed or subsidized depend mainly on their productivity.

Let us assume that $\alpha_1 = 2$, $\alpha_2 = \alpha_3 = 1$, so that the star is more productive than the peripheral agents are. We observe that (37) is not verified, since $\alpha_1 = 2 > 1 = (\alpha_2 + \alpha_3)/2$, and thus, the Nash equilibrium is not optimal. Indeed, it is easily verified that

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

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 $\lambda \sum_{j=1}^n \widehat{g}_{ji} (x_j - \overline{x}_j)$ (see (33)) 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.[21] 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

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

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

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:

$$\mathbf{S}^O = \frac{\lambda}{(1+4\lambda)} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \tag{42}$$

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.

4.5 Second best with a balanced budget

Up to now, we have assumed no cost for restoring the first best. Let us now consider a second-best policy with a balanced budget. In the second-best policy, the social planner does not directly choose the effort levels of individuals but indirectly affects the equilibrium efforts by subsidizing/taxing efforts. Denote $\mathbf{S} = (s_1, s_2, \ldots, s_n)^T$ as the vector of subsidies/taxes per unit of effort. However, now, the budget must be balanced, that is, the total amount of subsidies and taxes must equal zero. In other words, the subsidies given to certain agents must be financed by the taxes collected from other agents. Hence, the budget constraint is given by

$$\mathbf{S}^T \widehat{\mathbf{M}} (\boldsymbol{\alpha} + \mathbf{S}) = 0, \tag{43}$$

where, as implied by (16), the left-hand side is the total amount $\mathbf{S}^T \mathbf{x}^*$ of subsidies.

Using (18), we find that the equilibrium level of total welfare under subsidies $\mathbf{s}$ can be written as follows:

$$\mathcal{W}(\mathbf{S}) = (\boldsymbol{\alpha} + \mathbf{S})^T \left[\mathbf{I} - \left(\mathbf{I} - \widehat{\mathbf{M}} \right)^T \left(\mathbf{I} - \widehat{\mathbf{M}} \right) \right] (\boldsymbol{\alpha} + \mathbf{S}) \,. \tag{44}$$

which means that $Corr\,(\boldsymbol{\pi}, \boldsymbol{\alpha}) > 0$, since the star (most central agent) has higher productivity α than the peripheral agents do. Thus, our results confirm Proposition 9 since, 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.

The planner's problem is to choose a subsidy/tax vector $\mathbf{s}$ that maximizes (44) subject to (43). In general, the solution to this problem can be obtained only numerically, since both the objective function (44) and the constraint (43) are quadratic. Therefore, we provide only an illustrative example. Consider the same star network with $n = 3$ as in sub-section 4.4 in which $\alpha_1 = 2$, $\alpha_2 = \alpha_3 = 1$. For simplicity, set $\lambda = 0.25$. Then, the subsidy/tax vector with a balanced budget is given by

$$\mathbf{S}^{SB} = \begin{pmatrix} -1.2977 \\ 0.4285 \\ 0.4285 \end{pmatrix}.$$

Owing to budget constraint (43), the central agent is taxed much more heavily than in the first best, while the periphery agents are more generously subsidized. To observe this, plug $\lambda = 0.25$ into (42) to obtain

$$\mathbf{S}^{O} = \begin{pmatrix} -0.25 \\ 0.125 \\ 0.125 \end{pmatrix}.$$

As illustrated by this example, the second-best policy is much less (more) beneficial for the star (periphery) agent than the first-best policy is. More generally, this example shows that, in order to have a balanced-budget policy that maximizes total welfare, it is necessary to tax/subsidize more (less) heavily the agents who exert large (low) externalities on their neighbors.

5 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.[22] In the local-average model, deviating from the average effort of one's peers negatively affects the utility of an individual (see (6)). The closer each individual's effort is to the average of her friends' efforts, the higher is

[22]There is also another well-studied model in network games, which is a game with strategic substituability (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 our topic of analysis in this study, since we focus on the linear-in-means model in which peer effects are supposed to be positive.

her utility. 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 (7)). 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 they have different policy implications. In order 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.

5.1 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 disadvantage kids 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.[23]

[23]See also Curto and Fryer (2014), who study the SEED schools, which are boarding schools

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.

5.2 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. 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. In other words, 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.[24]

5.3 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

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.

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

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) foundind 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.

5.4 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 4 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.

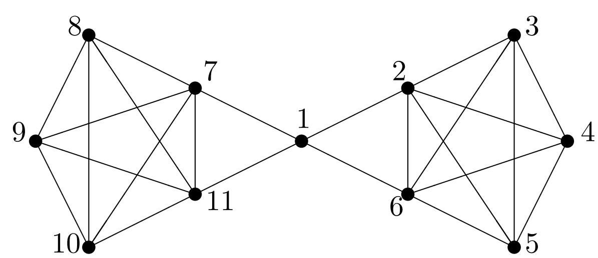

Figure 4: A bridge network

5.4.1 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. In order 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 (7). Then, if $\theta < 1/\mu(\mathbf{g})$ (where $\mu(\mathbf{g})$ is the largest eigenvalue of $\mathbf{g}$),[25] 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 (6). 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 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.

[25] This condition is verified for the network displayed in Figure 4, since $\theta = 0.2 < 0.227 = 1/\mu(\mathbf{g})$.

5.4.2 A group-based policy: Changing the norm

Consider again the network **g** in Figure 4 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.

6 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 the 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). We show that individual productivity positively affects equilibrium effort and utility. We also show that the taste for conformity has an ambiguous effect on individual outcomes and depends on whether an individual is above or below her own social norm. Equilibria are usually inefficient and we provide a condition for the productivity of the agents that ensures that the Nash equilibrium in efforts is always optimal. For example, for a star network, this condition boils down to the fact that the central agent's productivity has to equal the average productivity of all peripheral agents. Because this condition is unlikely to hold, we show that to restore the first best, the planner needs to subsidize agents who are connected to other agents who make efforts above their social norms and to tax agents who are connected to other agents who make efforts below their social norms. This finding implies that the planner does not always want to subsidize central agents, as is the case in the local-aggregate model.

An interesting extension would be to add a first stage in our model in which agents form links with each other. This is clearly a difficult exercise. However, if agents have different productivity, homophily behavior emerges, since high- (low-) productivity agents are likely to form links with other high- (low-) productivity agents, because this reduces their cost of failing to conform to neighbors. In other words, in a network-formation model, if there is a cost of forming links, agents like to minimize the distance between their effort and that of their direct neighbors (social norm) and there is a positive correlation between own productivity and neighbors' productivities. However, the correlation between productivity and the degree of the agents is unclear.[26]

More generally, we believe that our framework is 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.

[26] Jackson (2018) developed a similar two-stage model but within the local-aggregate framework and showed that there is a positive correlation between the productivity α of an agent and her degree. Indeed, because of the complementarities in the local-aggregate model, people with higher αs benefit more from the interactions with others than do those with lower αs and thus, prefer to have a higher degree.

[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] 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.

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

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

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

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

[14] 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.

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

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

[17] 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.

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

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

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

[21] Cawley, J., Han, E., Kim, J., and Norton, E.C. (2017), "Testing for Peer Effects Using Genetic Data," NBER Working Paper No. 23719.

[22] 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.

[23] 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.

[24] Courant, R. and John, F. (2012), *Introduction to Calculus and Analysis I*, Berlin: Springer.

[25] 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.

[26] 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.

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

[28] 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.

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

[30] Dutta, R., Levine, D.K., and Modica, S. (2018), "Interventions when social norms are endogeneous: A critique," Unpublished manuscript, McGill University.

[31] 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.

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

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

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

[35] 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.

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

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

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

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

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

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

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

[43] Jackson, M.O. (2018), "The friendship paradox and systematic biases in perceptions and social norms," *Journal of Political Economy*, forthcoming.

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

[45] 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.

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

[47] 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.

[48] 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.

[49] 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.

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

[51] 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.

[52] 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.

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

[54] 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.

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

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

[57] Sacerdote, B. (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.

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

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

[60] 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.

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 (5) of the social norm $\overline{x}_i$ into the (A.15) implies (14). Restating (14) in matrix form, we obtain (15). By solving (15) for $\mathbf{x}$, we verify that the Nash equilibrium $\mathbf{x}^*$ is indeed defined by (16). 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 (16) and observing that $\overline{\mathbf{x}}^* = \widehat{\mathbf{G}}\mathbf{x}^*$, we obtain equation (17) for the equilibrium social norms. This proves part (ii). ■

Proof of part (iii). We now prove that the equilibrium utility levels are given by (18). To do this, we use the A.15 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 (9), 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 (18) and proves part (iii).

The proof of Proposition 1 is now completed. ∎

Proof of Lemma 1. First, restate the FOC as

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

We have

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

This implies that

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

This implies that

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

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 Lemma 2. 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.):

$$\widehat{\mathbf{G}} = \begin{pmatrix} \mathbf{0} & \mathbf{A}_{12} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_{23} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_{K-1,K} \\ \mathbf{A}_{K1} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}, \tag{A.4}$$

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 (A.4) is when $\mathbf{g} = K_{m,n}$. Indeed, in this case, the row-normalized adjacency matrix $\widehat{\mathbf{G}}$ is given by

$$\widehat{\mathbf{G}} = \begin{pmatrix} \mathbf{0}_{m\times m} & \frac{1}{n}\mathbf{1}_{m\times n} \\ \frac{1}{m}\mathbf{1}_{n\times m} & \mathbf{0}_{n\times n} \end{pmatrix}, \tag{A.5}$$

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. ∎

Proof of Proposition 2. 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 $\widehat{\mathbf{G}}$. Then, restating (16) in coordinate form, we can express $x_i^*(\lambda)$ as follows:

$$x_i^*(\lambda) = (1-\lambda) \sum_{k=0}^{\infty} a_{ik} \lambda^k. \tag{A.6}$$

By Lemma 2, 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.6) yields

$$x_i^*(\lambda) - \boldsymbol{\pi\alpha} = (1-\lambda) \sum_{k=0}^{\infty} \left(a_{ik} - \boldsymbol{\pi\alpha}\right) \lambda^k. \tag{A.7}$$

The ergodicity condition (24) 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

$$|a_{ik} - \boldsymbol{\pi\alpha}| < C|\lambda_2|^k. \tag{A.8}$$

Inequality (A.8) 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.7) under $\lambda \to 1$, we obtain (25).

Case 2. $\widehat{\mathbf{G}}$ is periodic. Then, by Lemma 2, $\mathbf{g} = K_{m,n}$. Combining this with (16), 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 (38). When $\lambda \to 1$, we obtain:

$$\lim_{\lambda\to 1} x_i^*(\lambda) = \frac{\overline{\alpha}_1 + \overline{\alpha}_2}{2} = \frac{1}{2m}\sum_{j\in V_1}\alpha_j + \frac{1}{2n}\sum_{j\in V_2}\alpha_j. \tag{A.9}$$

Combining (22) with (A.5), 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.9), implies (25) and completes the proof. ■

Proof of Proposition 4: We start with the following lemma.

Lemma A.1 *The matrixes* $\widehat{\mathbf{M}}$ *and* $\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ *are row-normalized.*

Proof. 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 A.1. ■

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

$$\mathbf{x}^* = \widehat{\mathbf{M}}\boldsymbol{\alpha}. \tag{A.10}$$

Hence,

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

which is strictly positive. Since, by Lemma A.1, $\widehat{\mathbf{M}}$ 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,

$$\overline{\mathbf{x}}^* = \widehat{\mathbf{G}}\widehat{\mathbf{M}}\boldsymbol{\alpha} \tag{A.11}$$

As seen from (A.11), $\frac{\partial \overline{x}_i^*}{\partial \alpha_i}$ 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 A.1, $\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ 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 (9). Let us first determine $\frac{\partial U_i^*}{\partial \alpha_i}$. In equilibrium, x_i is determined as the maximizer of (9) under $\mathbf{x}_{-i} = \mathbf{x}_{-i}^*$. By the envelope theorem, we obtain

$$\frac{\partial U_i^*}{\partial \alpha_i} = x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*)\frac{\partial \overline{x}_i^*}{\partial \alpha_i}. \tag{A.12}$$

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

Lemma A.2 *The following inequalities hold for all $i = 1, ..., n$:*

$$\max_j \alpha_j \geq \max\{x_i^*, \alpha_i\} \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq \min\{x_i^*, \alpha_i\} \geq \min_j \{\alpha_j\} \tag{A.13}$$

Proof of Lemma A.2: Let us, first, establish the second and the third inequalities in (A.13):

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

If $x_i^* \geq \overline{x}_i^*$, then it follows from equation (A.12) and Proposition 4 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.12) and Proposition 4 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:

$$\max\left\{x_i^*,\ x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*)\right\} \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq \min\left\{x_i^*,\ x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*)\right\}. \tag{A.14}$$

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

$$x_i^* + \frac{\lambda}{1-\lambda}(x_i^* - \overline{x}_i^*) = \alpha_i. \tag{A.15}$$

Combining (A.14) with (A.15) 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 A.1, $\widehat{\mathbf{M}}$ 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 A.2. ■

It remains to observe that Lemma A.2 immediately implies that $\frac{\partial U_i^*}{\partial \alpha_i} > 0$.

Proof of part *(iii)*: As implied by (18), 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

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

Since, by Proposition 4, $\frac{\partial \overline{x}_i^*}{\partial \alpha_j} > 0$, we obtain the desired result:

$$\text{sign}\left[\frac{\partial U_i^*}{\partial \alpha_j}\right] = \text{sign}\left(x_i^* - \overline{x}_i^*\right). \tag{A.16}$$

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

Proof of Proposition 5:
(i) Multiplying both parts of (16) 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 (14) 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, (28) holds at $\lambda = 0$. By continuity, (28) also holds in the vicinity of $\lambda = 0$, that is,when λ is positive but not too large. This proves part (ii).

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

(*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 (28), 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 us start with the following lemma:

Lemma A.3 *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*

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

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

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

Proof of Lemma A.3: 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

$$\widehat{\mathbf{G}} = \begin{pmatrix} 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix}, \qquad \widehat{\mathbf{G}}^2 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \end{pmatrix}. \tag{A.18}$$

Furthermore, we have

$$\widehat{\mathbf{G}}^3 = \begin{pmatrix} 0 & \frac{1}{n-1} & \cdots & \frac{1}{n-1} \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix} = \widehat{\mathbf{G}}, \tag{A.19}$$

In other words, the matrix $\widehat{\mathbf{G}}$ is cyclical with the cycle length equal to 2. Combining (A.18)–(A.19) with (14) 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

$$x_1^*(\lambda) = \alpha^{\mathrm{p}} + \frac{\alpha^{\mathrm{s}} - \alpha^{\mathrm{p}}}{1+\lambda}, \tag{A.20}$$

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

$$x_i^*(\lambda) = (1-\lambda)\alpha_i + \frac{\lambda}{1+\lambda}\alpha^{\mathrm{s}} + \frac{\lambda^2}{1+\lambda}\alpha^{\mathrm{p}} \tag{A.21}$$

Differentiating (A.20)–(A.21) with respect to λ, we obtain

$$\frac{\partial x_1^*}{\partial \lambda} = -\frac{(\alpha^{\mathrm{s}} - \alpha^{\mathrm{p}})}{(1+\lambda)^2}, \tag{A.22}$$

$$\frac{\partial x_i^*}{\partial \lambda} = \alpha^{\mathrm{p}} - \alpha_i + \frac{\alpha^{\mathrm{s}} - \alpha^{\mathrm{p}}}{(1+\lambda)^2}. \tag{A.23}$$

respectively. This completes the proof of this lemma. ▮

Using Lemma A.3, 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 (A.22) and (A.23), 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. ▮

Proof of Proposition 7.

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. ▮

Proof of Proposition 8.

Using (9), it is readily verified that, for any $\lambda \in [0, 1)$ the welfare functional (32) can be written as

$$\mathcal{W} = \boldsymbol{\alpha}^T \mathbf{x} - \frac{1}{2} \mathbf{x}^T \, \mathbf{W}(\lambda) \, \mathbf{x}, \tag{A.24}$$

where $\mathbf{W}(\lambda)$ is an $(n \times n)$-matrix defined by

$$\mathbf{W}(\lambda) := \mathbf{I} + \frac{\lambda}{1-\lambda} \left(\mathbf{I} - \widehat{\mathbf{G}} \right)^T \left(\mathbf{I} - \widehat{\mathbf{G}} \right). \tag{A.25}$$

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

Lemma A.4

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

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

(iii) We have

$$\lim_{\lambda \to 1} \mathbf{W}^{-1}(\lambda) = \frac{1}{n} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}, \tag{A.26}$$

where $\mathbf{W}^{-1}(\lambda)$ is the matrix inverse to $\mathbf{W}(\lambda)$ evaluated at $\lambda \in [0, 1)$.

Proof of Lemma A.4.

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

$$\mathbf{x}^T \, \mathbf{W}(\lambda) \, \mathbf{x} = \|\mathbf{x}\|^2 + \frac{\lambda}{1-\lambda} \|\mathbf{x} - \widehat{\mathbf{G}}\mathbf{x}\|^2 \geq \|\mathbf{x}\|^2, \tag{A.27}$$

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

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

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

Hence, all its eigenvalues are real. Furthermore, as shown in (i), $\mathbf{W}(\lambda)$ 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

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

while the corresponding eigenspace is the span of all the minimizers. When $\|\mathbf{x}\|^2 = 1$, (A.27) implies that $\mathbf{x}^T\mathbf{W}(\lambda)\mathbf{x} \geq 1$. Furthermore, when $\lambda \in (0, 1)$ we have

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

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:

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

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

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

and hence, $\mathbf{W}^{-1}(\lambda)$ is well defined for any $\lambda \in [0, 1)$. Because $\mathbf{W}(\lambda)$ 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 $\mathbf{W}^{-1}(\lambda)$ equals one, the corresponding eigenspace being the span of $\mathbf{1}$, while other $n-1$ eigenvalues of $\mathbf{W}^{-1}(\lambda)$ converge to zero as $\lambda \to 1$. Putting all these considerations together, and denoting by $\mathbf{s}_i$ the ith column eigenvector of $\mathbf{W}^{-1}(\lambda)$, chosen so that $\|\mathbf{s}_i\| = 1$ for all $i = 1, 2, \ldots, n$, we obtain

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

Performing matrix multiplication yields (A.26) and completes the proof of Lemma A.4. ■

We now proceed with the proof of Proposition 8.

Proof of part (*i*): Using (A.24), the social planner's problem can be written as follows:

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

where $\mathbf{W}(\lambda)$ is defined by (A.25). The FOC to this problem is given by

$$\mathbf{W}(\lambda) \mathbf{x} = \boldsymbol{\alpha}. \tag{A.28}$$

Combining (A.28) with (A.25), it is readily verified that (A.28) is equivalent to (34). Clearly, the solution to (A.28)—or, equivalently, to (34)—is unique and is given by

$$\mathbf{x}^O = \mathbf{W}^{-1}(\lambda) \boldsymbol{\alpha}. \tag{A.29}$$

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.24), the Hessian matrix of $\mathcal{W}$ equals $-\frac{1}{2}\mathbf{W}(\lambda)$, which is negative definite by part (*i*) of Lemma A.4. 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:

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

Lemma A.5 *If $\mu_\alpha > \sqrt{n}\sigma_\alpha$, then $\mathbf{x}^O$ is the interior solution to the social planner's problem.*

Proof of Lemma A.5. 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,

$$\boldsymbol{\alpha} = \mu_\alpha \mathbf{1} + \sigma_\alpha \sqrt{n}\, \boldsymbol{\varepsilon}. \tag{A.30}$$

Plugging (A.30) into (A.29), we obtain

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

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

$$x_i^O \geq \mu_\alpha - \sigma_\alpha \sqrt{n} \max_{j=1,2,\ldots,n} \left| \left(\mathbf{W}^{-1}(\lambda)\boldsymbol{\varepsilon}\right)_j \right| \geq \mu_\alpha - \sigma_\alpha \sqrt{n}\, \|\, \mathbf{W}^{-1}(\lambda)\boldsymbol{\varepsilon} \,\|, \tag{A.31}$$

where the second inequality follows from the standard triangle inequality. As implied by part (*ii*) of Lemma A.4, the spectral radius of $\mathbf{W}^{-1}(\lambda)$ equals 1. Hence, we obtain

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

Combining this with (A.31) 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 A.5. ■

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 (34) with (16), we find that $\mathbf{x}^* = \mathbf{x}^O$ if and only if the following condition holds:

$$\widehat{\mathbf{G}}^T \left(\mathbf{I} - \widehat{\mathbf{G}}\right) \mathbf{x}^* = 0. \tag{A.32}$$

Using (16), 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 (34) 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 Corollary 4

Let us derive (35) for complete bipartite graphs: $\mathbf{g} = K_{m,n}$. The Nash equilibrium $\mathbf{x}^*$ is the solution to (15). A necessary and sufficient condition (35) for the Nash equilibrium $\mathbf{x}^*$ to deliver a first best is given by

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

Using (15), this condition can be equivalently restated as follows:

$$\widehat{\mathbf{G}}^T\boldsymbol{\alpha} = \widehat{\mathbf{G}}^T\overline{\mathbf{x}}^*, \tag{A.33}$$

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 (A.5). Hence, the best reply functions (14) take the following form:

$$x_i = (1-\lambda)\alpha_i + \begin{cases} \dfrac{\lambda}{n}\displaystyle\sum_{k\in V_2} x_k, & i \in V_1, \\ \dfrac{\lambda}{m}\displaystyle\sum_{k\in V_1} x_k, & i \in V_2. \end{cases} \tag{A.34}$$

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

$$\frac{1}{m}\sum_{k\in V_1} x_k = \frac{1-\lambda}{m}\sum_{k\in V_1}\alpha_k + \frac{\lambda}{n}\sum_{k\in V_2} x_k, \tag{A.35}$$

$$\frac{1}{n}\sum_{k\in V_2} x_k = \frac{1-\lambda}{n}\sum_{k\in V_2}\alpha_k + \frac{\lambda}{m}\sum_{k\in V_1} x_k. \tag{A.36}$$

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 (A.35)–(A.36) in terms of $\overline{x}_1$ and $\overline{x}_2$, we obtain

$$\overline{x}_r^* = \frac{\lambda}{1+\lambda}\overline{\alpha}_r + \frac{1}{1+\lambda}\overline{\alpha}_s, \tag{A.37}$$

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 (A.33) holds if and only if $\overline{\alpha}_r = \overline{x}_r^*$ for $r = 1, 2$. Using (A.37), we find that this is equivalent to $\overline{\alpha}_1 = \overline{\alpha}_2$. This completes the proof. ∎

Proof of Proposition 9

Proof of part (*i*). Taking the limit as $\lambda \to 1$ on both sides of (A.29), and using part (*ii*) of Lemma A.4, we obtain (39). This proves part (*i*). ■

Proof of part (*ii*). 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 (25) with (39), 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 Proposition 10 Omitted. ▮

B Additional results and examples

B.1 Additional result

We here provide an additional result that is mentioned in the main text.

Corollary B.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 B.1

(*i*) First, by totally differentiating both parts of (14) 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

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

That agents are, on average, more productive than their neighbors means that the right-hand side of (B.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 3 yields

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

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. ∎

B.2 Example: How λ affects effort in a circular regular network

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 $i = 1, \ldots, 5$.

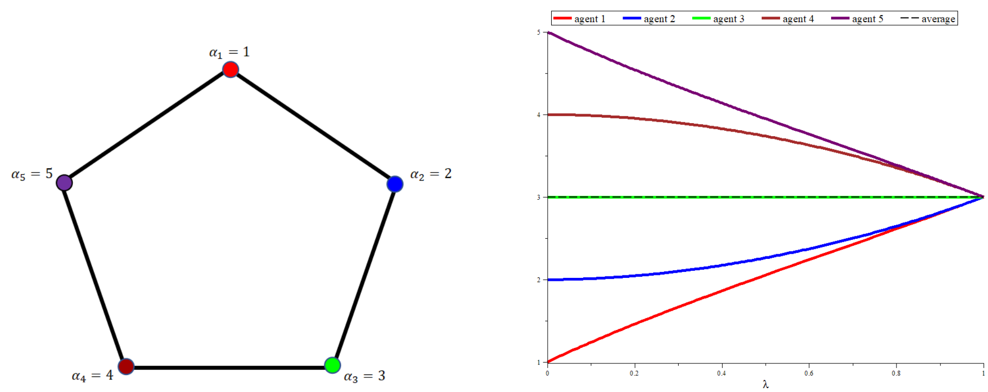

Figure B1: How λ affects effort in a circular network

The left panel of Figure B1 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 7, as the average effort is strictly proportional to the aggregate effort. Observe that, when $\lambda \to 1$, all agents make the same effort (Proposition 2), which is equal to the average effort for regular networks, as shown in (26).

Let us now rewire the links in the network of Figure B1 so that, topologically, the new network, displayed in the left panel of Figure B2, 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 B1. 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 B1 and B2, 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.

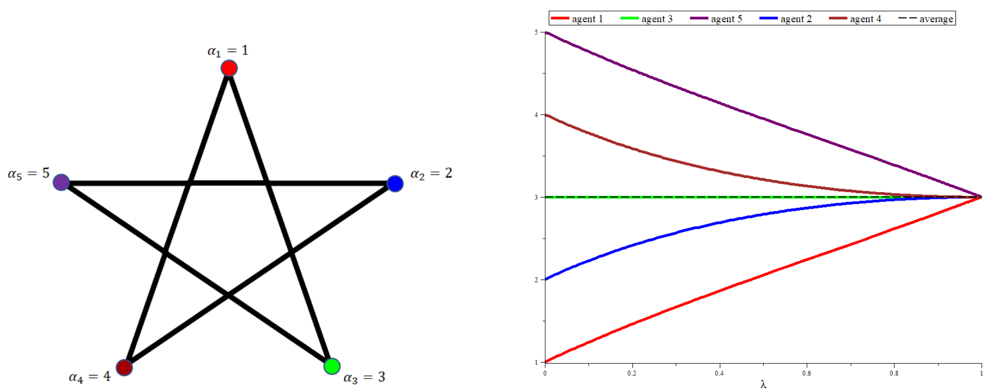

Figure B2: How λ affects effort in a rewired circular network

Finally, let us rewire again the social ties without changing the network topology. We obtain the network depicted in the left panel of Figure B3. Although the network is regular, we observe, in the right panel of Figure B3, 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.

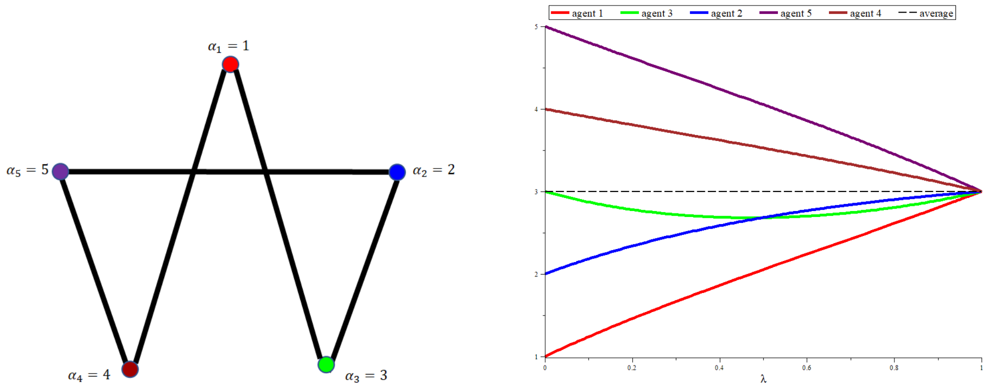

Figure B3: How λ affects effort in another rewired circular network