Technology Adoption and Social Norms

Philip Ushchev* Yves Zenou†

November 10, 2017

Abstract

Agents, embedded in a social network, first decide whether or not to adopt a new costly technology, and, then, choose their level of productivity effort. The latter choice is affected by the social norm of each individual so that she loses utility from failing to conform to the average effort of her peers (local-average model). Contrary to the local-aggregate model, we show that, in the second stage, if agents are ex ante identical but have different positions in the network, they all exert the same effort level, which corresponds to the first best. We also demonstrate that multiple equilibria may arise in the two-stage game. We show under which conditions symmetric and asymmetric subgame-perfect Nash equilibria emerge and why they are inefficient. Finally, we propose different subsidy policies that can restore the first-best solutions.

Keywords: Technology adoption, networks, social norms, local average model
JEL Classification: D85, J15, Z13.

*National Research University Higher School of Economics, Russia. Email: fuschev@hse.ru, ph.ushchev@gmail.com.

†Monash University, IFN, and CEPR. E-mail: yves.zenou@monash.edu.

1 Introduction

Conformism is the idea that the easiest and hence best life is attained by doing one's very best to blend in with one's surroundings and to do nothing eccentric or out of the ordinary in any way. It may well be best expressed in the old saying, "When in Rome, do as the Romans do." People do indeed conform to social norms and this is particularly true in education (Liu et al., 2014), crime (Liu et al., 2012), participation in extracurricular activities (Boucher, 2016), teenage pregnancy, obesity (Cawley et al., 2017), smoking (Hsieh and Lin, 2017), tax evasion (Fortin et al., 2007), etc.[1]

We model this idea by considering the so-called *local average model* in network games (Patachini and Zenou, 2012; Liu et al., 2014; Fortin and Boucher, 2016), which assumes that each individual wants to conform to the social norm of his/her peers. Interestingly, most theoretical contributions in network games (Ballester et al., 2006; Jackson and Zenou, 2015; Bramoulle and Kranton, 2016; Bramoulle et al., 2016) have been using the *local aggregate model*, where it is the sum of friends that matters and not the social norm, while most empirical studies have been testing the *local average model* (Bramoulle et al., 2009; Lin, 2010; Blume et al., 2014; Boucher et al., 2014; Fortin and Boucher, 2016; Hsieh and Lin, 2017, among others). This discrepancy between the theoretical analysis and the empirical applications in the network literature has thus motivated us to study the theoretical properties of the local average model, which, unfortunately, remain understudied. This is the *first* contribution of this paper.

Our main findings can be summarized as follows. First, we characterize the Nash equilibrium in the local-average model and highlight its fundamental feature: when agents are ex ante identical, i.e. when they have the same productivities, they all exert the same effort level regardless of the agent's position in the network relative to the others. We show that this effort level corresponds to the first best. These results stand in sharp contrast with the local-aggregate model, where the equilibrium effort level is usually not efficient, even when agents have the same productivity levels. We also study the case when agents can be classified into two categories: high-skilled

[1] Conformity can be so strong that, even in cases where people can obviously determine that others are incorrect, they still conform to the social norm. This is what Asch (1955, 1956) have found in his famous experiments. Asch presented subjects with two cards, one contained a single reference line and the other contained three lines of various lengths (one was the same length as the reference line). Asch manipulated the social situation by occasionally having two confederates publicly answer incorrectly prior to the subject providing an answer. The subject heard the incorrect responses of the others and was asked to publicly declare his answer as well. Asch found that the degree of conformity was relatively high.

and low-skilled individuals.[2] We show that high-(low-)skilled agents always choose a level of effort above (below) their social norm.

Second, we provide clear-cut comparative statics of equilibrium efforts with respect to changes in productivities. In the case when agents are either high- or low-skilled, we show that an increase in the productivity of high-skilled agents is always beneficial for them but, always detrimental for the low-skilled agents in terms of individual utility. Third, we provide a complete welfare analysis of the local-average model. We derive a necessary and sufficient condition for the equilibrium to be socially optimal. When this condition does not hold, we propose different tax/subsidy policies that can restore the first best.

The *second* contribution of this paper is to extend the previous model to a two-stage game where, in the first stage, each agent makes a binary $\{0, 1\}$ decision while, in the second stage, each agent chooses her effort level in the local-average model. Indeed, settings with binary actions and positive network effects are ubiquitous: the choice to adopt a technology or platform, such as in social media, where the value of adopting the technology/platform is increasing in the adoption by friends (Van den Bulte and Stremersch, 2004);[3] the choice to partake in crime, when the proficiency of crime, and thus the likelihood of not getting caught, is increasing in the criminality of accomplices (Warr, 2002; Haynie, 2001; Bayer et al., 2009; Patacchini and Zenou, 2012); or, the implementation of defense or anti-immigration policies, when the likelihood of attack or the influx of migrants depends on policies employed in neighboring countries (Dequiedt and Zenou, 2013; Mangin and Zenou, 2016). This paper studies coordination in these environments by focusing on whether or not to adopt a new technology. We assume that there is a cost of adopting the new technology but agents who adopt are more productive than the ones who do not adopt the new technology.

First, we show under which condition, there exists a symmetric subgame-perfect Nash equilibrium (SPNE) where either everybody adopts the new technology (Adopting Equilibrium), or nobody adopts (Non-Adopting Equilibrium) or both equilibria co-exist. In particular, we show that the Adopting (Non-Adopting) Equilibrium is a unique symmetric SPNE if and only if the cost of adopting is small (large) compared to the productivity cost of not adopting. Not surprisingly, under perfect information about the quality of the technology, there are multiple equilibria when the cost of

[2]This case is important in our context since it corresponds to the second stage of the technology adoption game we consider below.

[3]For products such as software, mobile phones, video game consoles, communication apps, etc., these peer-effects are technological in nature: consumers need to adopt technologies compatible with those of their peers in order to have effective interactions.

adopting has intermediate values because of coordination problems (as in the global game literature; see Morris and Shin, 2003). We then study asymmetric SPNE where some agents in the network adopt and some do not adopt. We provide a necessary condition for an asymmetric SPNE to exist and provide a general characterization of asymmetric SPNE. In particular, we show that, in a star network, there never exists an asymmetric SPNE while, in a chain network, the reverse is always true.

The rest of the paper unfolds as follows. In the next section, we relate our model to the relevant literature. In Section 3, we study the local-average model, determine the condition for its existence and uniqueness and investigate its welfare properties. In Section 4, we study the two-stage game where, in the first stage, each agent decides whether or not to adopt a new technology while, in the second stage, each agent chooses her level of effort in the local-average model framework, and determine the conditions under which symmetric and asymmetric subgame-perfect Nash equilibria exist and are unique. In Section 5, we examine the policy implications of our results. Finally, Section 6 concludes. All proofs can be found in Appendix A while, in Appendix B, we provide additional results.

2 Related literature

2.1 Local-average versus local aggregate model

The *local-average* model in network games has been developed to capture the role of *social norms*, e.g. 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 *local-aggregate* model aims at understanding the role of knowledge spillovers on outcomes (Ballester et al., 2006, 2010; Bramoulle et al., 2014; De Marti and Zenou, 2015). In the local-average model, deviating from the average of efforts of one's peers negatively affects the utility of an individual. The closer each individual's effort is to the average of her friends' efforts, the higher is her utility. In the local aggregate model, in contrast, it is the sum of the efforts of one's peers that positively affects the utility of each individual. When peers exert more effort, the utility derived from own effort increases.

Which model is relevant is 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) propose two methods. The first method is to estimate an augmented model, which includes both average peer effects and aggregate peer effects and see which are significant. The second method is to perform the J test that they propose. Using Add Health data, Liu et al. (2014) show that, for study effort in education, the endogenous peer effect is mostly captured by a social-conformity effect rather than a social-multiplier effect. In other words, the local-average model is the most appropriate model for education as measured by the "study effort" of each student. On the other hand, for sport activities, they find that both social-conformity and social-multiplier effects contribute to the endogenous peer effect. Moreover, Liu et al. (2012) study juvenile delinquency and show the local-aggregate model was 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. Using data from an in-house call center of a multi-national mobile network operator, Lindquist et al. (2017) study how co-worker productivity affects worker productivity via network effects. They show that there are indeed strong network effects in worker productivity. A 10% increase in the current productivity of a worker's co-worker network leads to a 1.7% increase in own current productivity. The estimation results show that this effect can be attributed to conformist behavior (local average network effects), which means that workers' productivity tend to be similar to the average productivity of their peers, measured here as co-workers belonging to the same team and working similar hours.

Our contribution is to be the first to fully develop the analysis of the local-average model. Patacchini and Zenou (2012) characterize the Nash equilibrium, Blume et al. (2015) introduce imperfect information in the local-average model and give conditions under which there exists a unique (Bayesian) Nash equilibrium, and Boucher (2016) embeds the local-average model into a network formation model. We are the first to study the comparative-statics properties of the model (impact of α and θ on efforts and utility), determine the first best and study subsidy policies that may restore the first best.

2.2 Two-stage model: (Binary) decision and effort choices

There is a literature on (complete information) games on networks with strategic complements where players can choose one of two actions. This includes coordination games, and generally all sort of games where players choose whether to do something (adopt a new technology, participate in something, provide a public good effort) or not (for overviews, see Jackson, 2008; Jackson and Zenou, 2015). In these games, players are connected in a network and only care about their neighbors' actions so that the payoff function depends on how many neighbors that each individual has. The main objective of this literature is to determine under which condition there will be "contagion" so that everybody adopt. A prominent paper in this literature is Morris (2000), who characterizes equilibrium adoption via the property of "cohesion" within subsets of players.[4] There is also a literature that introduces some randomness or imperfect information in these types of games. For example, Kandori et al. (1993), Young (1993) and Jackson and Watts (2002a,b) use stochastic stability to solve these coordination games. Leister et al. (2017) use the global game framework to study coordination on networks where the state of the world is unknown (for example, the quality of the technology is unknown in the technology adoption example).

Compared to this literature, our model uses a different framework, where, under complete information, individuals first make a binary choice and then choose an effort level in some activity. As a result, we do not study contagion effects and the threshold value above which agents choose an action (such as adopting a new technology or starting a revolution). Because our model is relatively simple, we are able to study the welfare properties of our model and policies that can increase adoption.

More generally, we believe that, in many real-world situations, it is important to study a two stage-game where, in the first stage, there is a *binary* decision, and, in the second stage, agents exert a *continuous* (effort) choice. In fact, this is a complicated problem where the solution is usually very difficult to characterize (see e.g. Ballester et al., 2010, where they analyze a game where in the first stage individuals can become either criminal or not and then, for criminals, how much crime effort). The trick here is that both choices in the first stage (adopting or not adopting a new technology) have an impact on the effort choices in the second stage. In Ballester et al. (2010), this is not the case since, in the first stage, individuals decide whether or not to become a criminal, but, if not criminal, the effort in the second stage is equal to zero and there is therefore a discontinuity in the payoffs of non-criminals. By avoiding this issue, we are able to totally characterize the subgame-perfect Nash equilibria of the two-stage game and determine under which condition it is unique.[5]

[4] See also Elliot et al. (2014) and Acemoglu et al. (2015) who model the clearing of liabilities between institutions and show that contagion offers an alternative prediction to the spread of perturbations over the network.

[5] To the best of our knowledge, there are only two other network models with binary choice decision and effort choices, as in our framework, but where, in the second-stage, players choose efforts in the local-aggregate model (see Helsley and Zenou, 2014 and Verdier and Zenou, 2017). The results are completely different both in terms of equilibrium characterization and welfare implications. For example, in the latter, there is always a unique subgame-perfect Nash equilibrium (SPNE) whereas, in our model, we have multiple SPNE. This is because the local-average model "creates" naturally coordination problems, which doe not appear in the local-aggregate model.

3 The local-average model

3.1 Definitions and notations

There are n individuals (agents) in the economy. The *network* $\mathbf{g}$ is composed of a set $\mathcal{N} = \{1, \ldots, n\}$ of agents, where $n \geq 2$, and a set of *links* or *direct connections* between them. The *adjacency matrix* $\mathbf{G} = [g_{ij}]$ is a $(n \times n)$-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 $g_{ji} = 1$, so the network is *undirected.* By convention, $g_{ii} = 0$.

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

$$\widehat{\mathbf{M}} := \left(\mathbf{I} - \frac{\theta}{(1+\theta)} \widehat{\mathbf{G}} \right)^{-1} = \sum_{k=0}^{\infty} \left(\frac{\theta}{1+\theta} \right)^{k} \widehat{\mathbf{G}}^{k}, \tag{1}$$

and denote by $\widehat{m}_{ij}$ the ijth entry of this matrix. The matrix $\widehat{\mathbf{M}}$ has the same flavor as the matrix in the Bonacich centrality matrix associated with the network $\mathbf{g}$ (Bonacich, 1987; Ballester et al., 2006). The difference, however, is that $\widehat{\mathbf{M}}$ is row-normalized, which implies $\lambda_1(\widehat{\mathbf{M}}) = 1$, where $\lambda_1(\mathbf{A})$ denotes the largest eigenvalue of a squared matrix $\mathbf{A}$ with non-negative entries. This guarantees, together with $0 < \theta/(1+\theta) < 1$, that the series in (1) converges without imposing any additional conditions on θ.

3.2 Preferences

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

$$\overline{x}_i := \sum_{j \in \mathcal{N}} \widehat{g}_{ij} x_j = \frac{1}{d_i} \sum_{j: g_{ij}=1} x_j. \tag{2}$$

Agent i's utility function is given by

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

where $\alpha_i > 0$ is individual i's *productivity*, while the parameter $\theta > 0$ measures the *taste for conformity.* The utility function (3) has two parts. The first one, $\alpha_i x_i - \frac{1}{2} x_i^2$, is the utility of exerting x_i when there is *no* interaction with other individuals. The second part, $-\frac{\theta}{2}(x_i - \overline{x}_i)^2$, captures the effect of peers on agent i's own action. It is such that each individual wants to minimize the social distance between herself and her reference group. Indeed, the individual loses utility $\frac{\theta}{2}(x_i - \overline{x}_i)^2$ from failing to conform to others. This is the standard way economists have been modelling conformity (see, among others, Akerlof, 1980, 1997; Bernheim, 1994; Kandel and Lazear, 1992; Fershtman and Weiss, 1998; Patacchini and Zenou, 2012).

Most of the results of this section hold true for an arbitrary productivity pattern $\boldsymbol{\alpha} := (\alpha_1, ..., \alpha_n)^{\mathrm{T}}$. However, we will proceed with special focus on the case when agents can be separated in two categories: "high productivity" and "low productivity" agents, i.e. for each $i \in \mathcal{N}$ we have $\alpha_i \in \{\alpha^L, \alpha^H\}$, where $\alpha^H > \alpha^L > 0$. The reason for paying additional attention to this special case is its relevance for the analysis of technology adoption behavior, which we postpone until Section 4.

3.3 Nash equilibrium

Each individual i chooses x_i to maximize (3) 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 with respect to x_i for each $i \in \mathcal{N}$, we obtain the following best-reply functions:

$$x_i = \frac{\alpha_i}{1+\theta} + \frac{\theta}{1+\theta}\overline{x}_i. \tag{4}$$

Equivalently, the $(n \times 1)$ vector $\mathbf{x}^* := (x_1^*, ..., x_n^*)^T$ of equilibrium efforts must be a solution to

$$\mathbf{x}^* = \frac{1}{1+\theta}\boldsymbol{\alpha} + \frac{\theta}{1+\theta}\widehat{\mathbf{G}}\mathbf{x}^*, \tag{5}$$

where $\boldsymbol{\alpha}$ is the productivity pattern. As implied by (4), agent i's level of effort x_i^* is a linear function of the efforts of all agents to whom i is directly connected in the network. Solving (5) yields the Nash equilibrium $\mathbf{x}^*$, which is characterized by the following Proposition.

Proposition 1 *[Equilibrium efforts]*

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

$$\mathbf{x}^* = \frac{1}{1+\theta}\widehat{\mathbf{M}}\,\boldsymbol{\alpha} = \frac{1}{1+\theta}\sum_{k=0}^{\infty}\left(\frac{\theta}{1+\theta}\right)^k \widehat{\mathbf{G}}^k\boldsymbol{\alpha}. \tag{6}$$

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

$$\overline{\mathbf{x}}^* = \frac{1}{1+\theta}\widehat{\mathbf{G}}\widehat{\mathbf{M}}\boldsymbol{\alpha} = \frac{1}{1+\theta}\sum_{k=0}^{\infty}\left(\frac{\theta}{1+\theta}\right)^k \widehat{\mathbf{G}}^{k+1}\boldsymbol{\alpha}. \tag{7}$$

(*iii*) *For each* $i \in \mathcal{N}$*, agent* i*'s equilibrium utility level* $U_i^*(\boldsymbol{\alpha}, \mathbf{g}) := U_i(x_i^*, \mathbf{x}_{-i}^*, \mathbf{g})$ *is given by*

$$U_i^*(\boldsymbol{\alpha}, \mathbf{g}) = \frac{1}{2}\left[\alpha_i^2 - \frac{1+\theta}{\theta}\left(\alpha_i - \sum_{j\in\mathcal{N}}\frac{\widehat{m}_{ij}(\theta)}{1+\theta}\alpha_j\right)^2\right]. \tag{8}$$

(*iv*) *Assume there are two levels of productivities, i.e. for each* $i \in \mathcal{N}$*, we have* $\alpha_i \in \{\alpha^L, \alpha^H\}$*, where* $\alpha^H > \alpha^L > 0$*. Then, agent* i*'s equilibrium* x_i^* *effort is above (below) her social norm* $\overline{x}_i^*$ *if and only if agent* i *has a high (low) productivity, i.e.* $\alpha_i = \alpha^H$ *(i.e.* $\alpha_i = \alpha^L$*).*

The equilibrium effort in equation (6) 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* θ (except that $\theta > 0$) to guarantee that the Nash equilibrium exists, is unique, and is interior. Second, 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 $i, j \in \mathcal{N}$), then, regardless of the network structure, *the equilibrium effort levels are the same across agents* (i.e. $x_i^* = x_j^*$ for any $i, j \in \mathcal{N}$). This result displays significant differences between the local average model and the local aggregate model. In the former, when productivities are the same, the outcome does not depend on the network structure, and not even on the number of agents! In contrast, in the latter, efforts are proportional to the Bonacich centralities of agents, which depend crucially on the agents' positions in the network, no matter whether productivities are 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 only affects the equilibrium via the *cross-effect* with the vector of productivities.

To highlight this intuition, let us restate (6) in the coordinate form, which yields for each $i \in \mathcal{N}$:

$$x_i^* = \sum_{j \in \mathcal{N}} \frac{\widehat{m}_{ij}}{1+\theta} \alpha_j. \tag{9}$$

Define the *mean productivity* $\overline{\alpha}$, i.e. $\overline{\alpha} := \frac{1}{n} \sum_{j \in \mathcal{N}} \alpha_j$, and the vector $\mathbf{m}_i \in \Delta_{n-1}$ of *relative positions* of agent i with respect to other agents in the network by: $\mathbf{m}_i := \frac{1}{1+\theta}(\widehat{m}_{i1}, ..., \widehat{m}_{in})$. Then, (9) can be written as follows:

$$x_i^* = \overline{\alpha} + n \times \operatorname{cov}(\mathbf{m}_i, \boldsymbol{\alpha}), \tag{10}$$

where $\operatorname{cov}(\cdot, \cdot)$ stands for the covariance operator. Equation (10) clarifies why the network structure does not matter when $\boldsymbol{\alpha} = \alpha \mathbf{1}$. In this case, the distribution of individual productivities is degenerate, hence, it is uncorrelated with the distribution of agent i's positions in the network. As a result, the idiosyncratic covariance term, $\operatorname{cov}(\mathbf{m}_i, \boldsymbol{\alpha})$, which fully captures the impact of the network structure on x_i^*, vanishes for all $i \in \mathcal{N}$. This suggests that, *in the local average model, the network* $\mathbf{g}$ *does not affect the equilibrium per se, but only through the fact that its structure is correlated with the distribution* $\boldsymbol{\alpha}$ *of productivities.*

Finally, the last result of Proposition 1 states that, when there are only two productivity levels, agents with the highest (lowest) productivity level will provide an equilibrium effort above (below) the average effort of her direct neighbors, i.e. her social norm. This is quite intuitive since the average effort of one's friends is a convex combination of efforts from high- and low-productivity individuals. It would interesting and quite straightforward to empirically test this result.

3.4 Comparative statics

Let us now perform some comparative statics exercises of the Nash equilibrium with respect to productivities $\boldsymbol{\alpha}$.

Proposition 2 *[Comparative statics]*

(*i*) *For all* $i, j \in \mathcal{N}$, *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,

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

(*ii*) *The equilibrium utility* $U_i^*(\alpha_i,\ \boldsymbol{\alpha}_{-i})$ *of each individual* $i \in \mathcal{N}$ *is increasing with her own productivity* α_i,

$$\frac{\partial U_i^*}{\partial \alpha_i} > 0.$$

(*iii*) *For any* $j \neq i$, *agent* i*'s equilibrium utility* U_i^* *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* $\overline{x}_i^*$.

(*iv*) *Assume there are only two levels of productivities, i.e. for each* $i \in \mathcal{N}$ *we have* $\alpha_i \in \{\alpha^L, \alpha^H\}$, *where* $\alpha^H > \alpha^L > 0$. *Then, an increase in* α^H *makes all agents with high productivity (low productivity) better (worse) off, while an increase in* α^L *generates the opposite effect.*

The first two results are relatively intuitive and yet 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 own 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 $\overline{x}_i^*$.

In Appendix B, we also consider the comparative statics with respect to the conformity parameter θ. In particular, we show that $\frac{\partial \overline{x}_i^*}{\partial \theta}$ does not depend on i, i.e. a change in θ leads to a change in all social norms in the same direction.

3.5 Welfare and first best

Let us compare equilibrium outcomes and socially optimal outcomes. Let us first calculate the first-best outcome of this economy and then determine the taxes/subsidies that can restore the first best.

3.5.1 First best

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

$$\mathcal{W} := \sum_{i \in \mathcal{N}} U_i(x_i, \mathbf{x}_{-i}, \mathbf{g}). \tag{11}$$

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 3 *[First best]*

(*i*) *For each* $i \in \mathcal{N}$, *the first best* $\mathbf{x}^O$ *is a solution to*

$$x_i^o = \frac{\alpha_i}{1+\theta} + \frac{\theta}{1+\theta}\overline{x}_i + \frac{\theta}{1+\theta} \sum_{j \in \mathcal{N} \setminus \{i\}} \widehat{g}_{ji} \left(x_j^o - \overline{x}_j^o \right),$$

or, in matrix form,

$$\mathbf{x}^O = \frac{1}{1+\theta}\boldsymbol{\alpha} + \frac{\theta}{1+\theta}\left(\widehat{\mathbf{G}} + \widehat{\mathbf{G}}^T - \widehat{\mathbf{G}}^T\widehat{\mathbf{G}}\right)\mathbf{x}^O. \tag{12}$$

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

$$\widehat{\mathbf{G}}^T(\mathbf{I} - \widehat{\mathbf{G}})\widehat{\mathbf{M}}\boldsymbol{\alpha} = \mathbf{0}. \tag{13}$$

(*iii*) *If agents are ex ante homogeneous in productivity, i.e.* $\alpha_i = \alpha_j = \alpha$ *for all* $i, j \in \mathcal{N}$, *then the Nash equilibrium in efforts are always optimal. Furthermore, if* $\det(\widehat{\mathbf{G}}) \neq 0$, *the converse is also true.*

This proposition gives an exact condition on the productivities αs that ensures that the Nash equilibrium in efforts are always optimal. It also shows that, in general, ex ante homogeneity is *not necessary* for the Nash equilibrium in efforts to be optimal. To illustrate, we provide some examples.

Example 1. Consider a *star network* with agent 1 being 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:

$$\alpha_1 = \frac{1}{n-1} \sum_{j \in \mathcal{N} \setminus \{1\}} \alpha_j.$$

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

Example 2. Assume now that $\mathbf{g}$ a circular network with $n = 4$, i.e. the row-normalized adjacency matrix is given by:

$$\widehat{\mathbf{G}} = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Then, we have $\mathbf{x}^* = \mathbf{x}^O$ if and only if the average productivities across *maximum independent sets*[6] are the same, i.e.

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

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

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

Corollary 1

(i) If the network is a complete bipartite graph (i.e. $g = K_{m,n}$) with a partition $N = V_1 \cup V_2$ where $|V_1| = m$, $|V_2| = n$, then we have $x^* = x^O$ 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

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

(ii) If $\mathbf{g}$ is a chain network, (13) is equivalent to ex ante homogeneity unless $n = 4r + 3$ for some non-negative integer r, in which case (13) defines a two-dimensional set of productivity vectors $\boldsymbol{\alpha}$ such that the Nash equilibrium in efforts is the first best.

Example 3. Let $\mathbf{g}$ be a chain network involving $n = 7$ agents. Then, the equilibrium is optimal if and only if the individual productivities satisfy the following constraints:

$$\frac{5+2\theta}{4+\theta}\alpha_1 > \alpha_2 > \frac{1+2\theta}{2+3\theta}\alpha_1, \qquad \alpha_3 = -\frac{2\theta+1}{1+\theta}\alpha_1 + \frac{3\theta+2}{1+\theta}\alpha_2,$$
$$\alpha_4 = \frac{2}{1+\theta}\alpha_1 + \frac{\theta-1}{1+\theta}\alpha_2, \qquad \alpha_5 = \frac{2\theta+5}{1+\theta}\alpha_1 - \frac{4+\theta}{1+\theta}\alpha_2,$$
$$\alpha_6 = \frac{4}{1+\theta}\alpha_1 + \frac{\theta-3}{1+\theta}\alpha_2, \qquad \alpha_7 = \frac{3-\theta}{1+\theta}\alpha_1 + \frac{2\theta-2}{1+\theta}\alpha_2,$$

while $\alpha_1 > 0$ is arbitrary.

[6] In graph theory, an *independent set* is a set of nodes in a graph such that no two nodes are adjacent. That is, it is a set S of nodes such that for every two vertices 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.

3.5.2 Restoring the first best

Assume that condition (13) does not hold. Then, to restore the first best, the planner can either subsidize or tax efforts. Let S_i^O denote the optimal subsidy per effort for each agent i, where

$$S_i^O = \frac{\theta}{(1+\theta)} \sum_{j \in \mathcal{N} \setminus \{i\}} \widehat{g}_{ji} \left(x_j - \overline{x}_j\right),$$

or, in matrix form:

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

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

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

Proposition 4 *[Subsidies] If the social planner gives to each agent i the following tax/subsidy per unit of effort:*

$$S_i^O = \frac{\theta}{(1+\theta)} \sum_{j \in \mathcal{N} \setminus \{i\}} \widehat{g}_{ji} \left(x_j - \overline{x}_j\right) \tag{14}$$

then the first best is restored.

By doing so, the planner will restore the first best and will subsidize (resp. tax) agents *who are connected to other agents who provide efforts above (resp. below) their social norms*. In other words, on needs to subsidize agents who exert effort below that of their neighbors and tax those who exert effort above that of their neighbors.

Let us illustrate this result with an example. Assume a star network with $n = 3$ and where agent 1 is the star. The following row-normalized adjacency matrix describes this network:

$$\mathbf{G} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \text{ and } \widehat{\mathbf{G}} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Take: $\alpha_1 = 2$, $\alpha_2 = \alpha_3 = 1$. It is easily verified that the Nash equilibrium, the social norm and the social optimum in efforts are given by:

$$\mathbf{x}^* = \frac{1}{(2\theta+1)}\begin{pmatrix} 3\theta+2 \\ 3\theta+1 \\ 3\theta+1 \end{pmatrix}, \overline{\mathbf{x}}^* = \frac{1}{(2\theta+1)}\begin{pmatrix} 3\theta+1 \\ 3\theta+2 \\ 3\theta+2 \end{pmatrix}, \mathbf{x}^O = \frac{2}{(9\theta+2)}\begin{pmatrix} 6\theta+2 \\ 6\theta+1 \\ 6\theta+1 \end{pmatrix}$$

This implies that $x_1^* > x_1^O$ and $x_2^* \gtreqless x_2^O \Leftrightarrow \theta \gtreqless 1/3$ so that

$$\mathbf{S}^O = \frac{1}{2\theta^2+3\theta+1}\begin{pmatrix} -2\theta \\ \theta/2 \\ \theta/2 \end{pmatrix}$$

In other words, to obtain the first best the planner needs to tax the central agent and subsidy the peripheral ones, which is exactly the opposite result than the one obtained with the local-aggregate model (Helsley and Zenou, 2016). This result strongly depends on the value of the productivities. Assume, for example, that $\alpha_1 = 0.5$, $\alpha_2 = \alpha_3 = 1$ so that the productivity of the central agent is the lowest one. Then,

$$\mathbf{x}^* = \frac{1}{(4\theta+2)}\begin{pmatrix} 3\theta+1 \\ 3\theta+2 \\ 3\theta+2 \end{pmatrix}, \overline{\mathbf{x}}^* = \frac{1}{(4\theta+2)}\begin{pmatrix} 3\theta+2 \\ 3\theta+1 \\ 3\theta+1 \end{pmatrix} \text{ and } \mathbf{x}^O = \frac{1}{(9\theta+2)}\begin{pmatrix} 7.5\theta+1 \\ 7.5\theta+2 \\ 7.5\theta+2 \end{pmatrix}$$

and

$$\mathbf{S}^O = \frac{3\theta}{4\theta+2}\begin{pmatrix} 1 \\ -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$$

So now it is optimal to subsidy the central agent and tax the peripheral agents.

3.5.3 Equilibrium welfare

The equilibrium welfare $\mathcal{W}^*$ is defined as follows:

$$\mathcal{W}^* := \sum_{i\in\mathcal{N}} U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}, \theta, \mathbf{g}).$$

What happens to the equilibrium welfare when the individual productivity α_i of agent i increases, while productivities $\boldsymbol{\alpha}_{-i}$ of the other agents remain constant? The answer is given by the following proposition.

Proposition 5 *[Equilibrium welfare] .*

(*i*) *There exists a strictly positive threshold level, $\widetilde{\theta} > 0$, of the conformity parameter θ, such that, for any $\theta < \widetilde{\theta}$ and for any $i \in \mathcal{N}$, the equilibrium welfare unboundedly grows with α_i,*

$$\frac{\partial \mathcal{W}^*}{\partial \alpha_i} > 0, \qquad \lim_{\alpha_i \to \infty} \mathcal{W}^* = +\infty. \tag{15}$$

(*ii*) *When α_i increases, apart from agent i, all the other agents in the network start getting worse off at some point. Moreover, for every agent $j \in \mathcal{N} \setminus \{i\}$ her individual equilibrium utility unboundedly falls with α_i:*

$$\lim_{\alpha_j \to \infty} U_i^* = \begin{cases} +\infty, & \text{if } i = j, \\ -\infty, & \text{if } i \neq j. \end{cases} \tag{16}$$

(*iii*) *If we replace $\mathcal{W}^*$ with the Rawlsian welfare measure $\mathcal{W}^{\mathrm{R}} := \min_{j \in \mathcal{N}} U_j^*$, social welfare always falls when an individual productivity α_i increases, i.e.*

$$\lim_{\alpha_i \to \infty} \mathcal{W}^{\mathrm{R}} = -\infty.$$

To sum up, Proposition 5 shows that the behavior of equilibrium welfare in settings with social norms depends significantly on the choice of the welfare criterion.

4 The two-stage model

Let us now consider a two-stage model where, in the first stage, each agent $i \in \mathcal{N}$ makes a $\{0, 1\}$ decision (for example, adopt a new technology or not) while, in the second stage, each agent chooses her level of effort in some activity as in the previous section. We assume that the agents who adopt the new technology are more productive than the ones who do not adopt the new technology. To be more precise, *adopters* (i.e. agents who choose to adopt the new technology) have a productivity of $\alpha^H := \alpha > 0$ but have a cost of adopting the technology equal to c while *non-adopters* have a productivity of $\alpha^L : \alpha - t > 0$ but have no cost of adopting the technology. In other words, c is the *cost of adopting the new technology* while t is the *productivity cost* of not adopting the new technology. We will see below that the relationship between these two parameters plays a key role in determining the pattern of equilibria.

4.1 Subgame-perfect equilibria: Existence, uniqueness, and characterization

In this section, we characterize subgame-perfect Nash equilibria (SPNE).

Definition 2 *A* ***symmetric*** *SPNE is when either all agents in a network are adopters (Adopting Equilibrium) or when they are all non-adopters (Non-Adopting Equilibrium). Otherwise, it is an* ***asymmetric*** *SPNE.*

4.1.1 Symmetric equilibria

Define

$$\overline{c}(\theta, \mathbf{g}) := \alpha t - \frac{t^2}{2}\left[1 - \frac{(1+\theta)}{\theta}\left(1 - \frac{1}{1+\theta}\mu(\theta, \mathbf{g})\right)^2\right] \tag{17}$$

and

$$\underline{c}(\theta, \mathbf{g}) := \alpha t - \frac{t^2}{2}\left[1 + \frac{(1+\theta)}{\theta}\left(1 - \frac{1}{1+\theta}\mu(\theta, \mathbf{g})\right)^2\right], \tag{18}$$

where $\mu(\theta, \mathbf{g})$ is the measure of (Bonacich) centrality of the "most central" agent:

$$\mu(\theta, \mathbf{g}) := \max_{i \in \mathcal{N}} \{\widehat{m}_{ii}(\theta, \mathbf{g})\} = \max_{i \in \mathcal{N}} \left\{\sum_{k=0}^{+\infty} \left(\frac{\theta}{1+\theta}\right)^k g_{ii}^{[k]}\right\}. \tag{19}$$

Observe that the impact of the network structure $\mathbf{g}$ on $\overline{c}(\theta, \mathbf{g})$ and $\underline{c}(\theta, \mathbf{g})$ is fully captured by the maximum centrality $\mu(\theta, \mathbf{g})$. We have the following result:

Proposition 6 *[Symmetric SPNE] All symmetric subgame-perfect Nash equilibria (SPNE) are characterized as follows:*

(*i*) *If $c \leq \underline{c}(\theta, \mathbf{g})$, the Adopting Equilibrium is the unique SPNE.*

(*ii*) *If $\underline{c}(\theta, \mathbf{g}) < c \leq \overline{c}(\theta, \mathbf{g})$, there are two SPNE: the Adopting Equilibrium and the Non-Adopting Equilibrium.*

(*iii*) *If $c > \overline{c}(\theta, \mathbf{g})$, the Non-Adopting Equilibrium is the unique SPNE.*

Proposition 6 has three immediate implications. First, at least one symmetric SPNE always exists. Second, the Adopting (Non-Adopting) Equilibrium is a unique symmetric SPNE if and only if c is small (large) enough, as implied by parts (*i*) and (*iii*) of Proposition 6. Third, when c is neither too large nor too small, i.e.

$\underline{c}(\theta, \mathbf{g}) < c < \overline{c}(\theta, \mathbf{g})$, multiple symmetric SPNE emerge. Indeed, in this case, Adopting Equilibrium and Non-Adopting Equilibrium are both SPNE, as implied by part (ii) of Proposition 6. Hence, when $\underline{c}(\theta, \mathbf{g}) < c < \overline{c}(\theta, \mathbf{g})$, technology adoption is described by a coordination game.

Figure 1 illustrates these results.

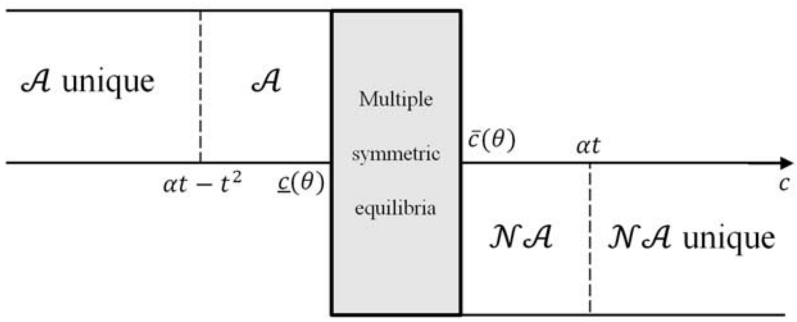

Figure 1: Characterization of symmetric SPNE

Corollary 3 *A necessary and sufficient condition for having multiple symmetric SPNE is:*

$$\mu(\theta, \mathbf{g}) < 1 + \theta - \frac{1}{t}\sqrt{2\theta(1+\theta)\left|c - \alpha t + \frac{t^2}{2}\right|}. \tag{20}$$

where $\mu(\theta, \mathbf{g})$, defined in (19), is value of the centrality of the agent who has the highest Bonacich centrality in the network. This condition says that *multiple symmetric SPNE emerge if and only if* $\mu(\theta, \mathbf{g})$ *is sufficiently small* so that there is no agent that is "too" central. We can still have multiple symmetric SPNE in a star network but this will depend on

4.1.2 Asymmetric equilibria

Let us check if there are some *asymmetric* SPNE, meaning that some agents adopt and some do not adopt in the same network.

Proposition 7 *[Non-existence of asymmetric SPNE] Assume that either* $c \leq \alpha t - t^2$ *or* $c > \alpha t$*. Then, asymmetric SPNE do not exist.*

In other words, a necessary condition for an asymmetric SPNE to exist is that *the adoption cost c is neither too large nor too small.*

We now provide a simple condition of whether a particular outcome is an asymmetric SPNE. To do so, let $\mathcal{A} \subseteq \mathcal{N}$ be the set of adopters. Then, it follows that $\mathcal{N} \setminus \mathcal{A}$ is the set of non-adopters. Consider an arbitrary partition of $\mathcal{N}$ into $\mathcal{A}$ and $\mathcal{N} \setminus \mathcal{A}$ such that both $\mathcal{A}$ and $\mathcal{N} \setminus \mathcal{A}$ are non-empty, and define

$$\overline{c}_{\mathcal{A}}(\theta, \mathbf{g}) := \alpha t - \frac{t^2}{2}\left[1 - \left(\frac{1+\theta}{\theta}\right)\min_{i \in \mathcal{A}}\left\{\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)\sum_{j \neq i}^{n}\frac{\widehat{m}_{ij}}{1+\theta}I_{\mathcal{A}}(j)\right\}\right] \tag{21}$$

and

$$\underline{c}_{\mathcal{A}}(\theta, \mathbf{g}) := \alpha t - \frac{t^2}{2}\left[1 - \left(\frac{1+\theta}{\theta}\right)\max_{i \in \mathcal{N} \setminus \mathcal{A}}\left\{\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)\sum_{j \in \mathcal{N} \setminus \{i\}}^{n}\frac{\widehat{m}_{ij}}{1+\theta}I_{\mathcal{A}}(j)\right\}\right], \tag{22}$$

where

$$I_{\mathcal{S}}(i) := \begin{cases} 1, & i \in \mathcal{S} \\ -1, & i \in \mathcal{N} \setminus \mathcal{S} \end{cases}$$

We have:

Proposition 8 *[Characterization of asymmetric SPNE] Let* $\mathcal{A} \subseteq \mathcal{N}$ *be such that both* $\mathcal{A}$ *and* $\mathcal{N} \setminus \mathcal{A}$ *are non-empty. Then, the asymmetric SPNE in which* $\mathcal{A}$ *is the set of adopters and* $\mathcal{N} \setminus \mathcal{A}$ *is the set of non-adopters exists if and only if*

$$\underline{c}_{\mathcal{A}}(\theta, \mathbf{g}) < c < \overline{c}_{\mathcal{A}}(\theta, \mathbf{g}). \tag{23}$$

As implied by Proposition 8, a sufficient condition for an asymmetric SPNE to exist is as follows: there exists a $\mathcal{A} \neq \varnothing, \mathcal{N}$, such that $\underline{c}_{\mathcal{A}}(\theta, \mathbf{g}) < \overline{c}_{\mathcal{A}}(\theta, \mathbf{g})$, or equivalently

$$\min_{i \in \mathcal{A}}\left\{\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)\sum_{j \neq i}^{n}\frac{\widehat{m}_{ij}}{1+\theta}I_{\mathcal{A}}(j)\right\} > \max_{i \in \mathcal{N} \setminus \mathcal{A}}\left\{\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)\sum_{j \neq i}^{n}\frac{\widehat{m}_{ij}}{1+\theta}I_{\mathcal{A}}(j)\right\} \tag{24}$$

It is easily verified that, for *regular networks*, there are no asymmetric SPNE. Our conjecture is that, for *star networks*, there are no asymmetric SPNE. A similar result is obtained in Leister et al. (2017) but in a coordination game with strategic complementarities (like here) but with the local-aggregate framework, no effort choice and imperfect information on the common state of the world (i.e. the quality of the technology in our technology adoption example). Indeed, using a global game approach where there is a unique cut-off value of the value of technology below which an individual adopts, Leister et al. (2017) show that, in a star network, all agents will either adopt or not adopt (i.e. there is no asymmetric Nash equilibrium) at the limit when the error term of the noisy signal goes to zero. However, the reason of this result is very different to ours since it is due to the fact that the star agent cannot have beliefs compatible with the fact that she will adopt while the peripheral agents will not. Leister et al. (2017) also show that, when the core of a core-periphery network increases (the core is equal to 1 in a star network), then an asymmetric Nash equilibrium will emerge where the agents in the core adopt while the periphery agents will not if the core is large enough. This is due to the fact that, because of strategic complementarities due to the local-aggregate framework, when the core is large enough, each individual in the core obtains enough positive externalities from her neighbors in order to adopt while the agents in the periphery find it optimal not to adopt.

4.2 Examples

Let us illustrate our results on asymmetric SPNE with examples.

4.2.1 Star-shaped network

Consider a star network with $n = 3$ where agent 2 is the star. Assume $\alpha = 10$ and $t = 6$. Then, $\alpha t = 60$, $\alpha t - t^2 = 24$, $\underline{c}(0.5, \mathbf{g}) = 38.625$ and $\overline{c}(0.5, \mathbf{g}) = 45.375$. A complete characterization of SPNE in this network is as follows:

Proposition 9 *Consider a star network with $n = 3$ where agent 2 is the star. For any value of θ, α, t, there never exists an asymmetric SPNE. Furthermore, assume $\alpha = 10$, $t = 6$ and $\theta = 0.5$.*

(*i*) *If $c \leq 38.625$, the Adopting Equilibrium is the unique SPNE.*

(*ii*) *If $38.625 < c \leq 45.375$, there are two SPNE: the Adopting Equilibrium and the Non-Adopting Equilibrium.*

(ii) If $c > 45.375$, the Non-Adopting Equilibrium is the unique SPNE.

4.2.2 Chain network

Consider a chain network with $n = 4$ where we start at agent 1 (extreme left), then agent 2, etc. For this case, the adjacency matrix $\mathbf{G}$ and the row-normalized adjacency matrix $\widehat{\mathbf{G}}$ are given, respectively, by

$$\mathbf{G} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ and } \widehat{\mathbf{G}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

Assume $\mathcal{A} = \{2,3\}$ and $\mathcal{N} \setminus \mathcal{A} = \{1,4\}$. Then, we have:

$$\min_{i \in \mathcal{A}} \left\{ \left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \sum_{j \neq i}^{n} \frac{\widehat{m}_{ij}}{1+\theta} I_{\mathcal{A}}(j) \right\} = -\frac{\theta^3 \left(4\theta^2 + 9\theta + 4\right)}{\left(6\theta^3 + 19\theta^2 + 16\theta + 4\right)^2}$$

and

$$\max_{i \in \mathcal{N} \setminus \mathcal{A}} \left\{ \left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \sum_{j \neq i}^{n} \frac{\widehat{m}_{ij}}{1+\theta} I_{\mathcal{A}}(j) \right\} = -\frac{\theta^2 \left(5\theta^2 + 10\theta + 4\right) \left(3\theta^3 + 13\theta^2 + 14\theta + 4\right)}{\left(6\theta^3 + 19\theta^2 + 16\theta + 4\right)^2 (1+\theta)}.$$

Hence, the condition (24) is equivalent to

$$15\theta^5 + 91\theta^4 + 199\theta^3 + 199\theta^2 + 92\theta + 16 > 0$$

which is true for any $\theta \geq 0$. Thus *there always exists an asymmetric equilibrium where the central agents adopt and the periphery agents do not adopt*, i.e. where $\mathcal{A} = \{2,3\}$ and $\mathcal{N} \setminus \mathcal{A} = \{1,4\}$.

Assume further that $\alpha = 10$ and $t = 6$. Then, we have:

$$\underline{c}_{\mathcal{A}}(\theta, \mathbf{g}) = 42 - 18 \frac{\theta \left(5\theta^2 + 10\theta + 4\right) \left(3\theta^3 + 13\theta^2 + 14\theta + 4\right)}{\left(6\theta^3 + 19\theta^2 + 16\theta + 4\right)^2},$$

$$\overline{c}_{\mathcal{A}}(\theta, \mathbf{g}) = 42 - 18 \frac{\theta^2 (1+\theta) \left(4\theta^2 + 9\theta + 4\right)}{\left(6\theta^3 + 19\theta^2 + 16\theta + 4\right)^2},$$

which means that $\mathcal{A} = \{2,3\}$ and $\mathcal{N} \diagdown \mathcal{A} = \{1,4\}$ is an asymmetric SPNE if and only if the adoption cost c satisfies the following inequality:

$$42 - 18\frac{\theta\left(5\theta^2+10\theta+4\right)\left(3\theta^3+13\theta^2+14\theta+4\right)}{\left(6\theta^3+19\theta^2+16\theta+4\right)^2} < c < 42 - 18\frac{\theta^2(1+\theta)\left(4\theta^2+9\theta+4\right)}{\left(6\theta^3+19\theta^2+16\theta+4\right)^2}.$$

Furthermore, using (17) – (19), we have:

$$\mu(\theta, \mathbf{g}) = \frac{2\left(\theta+1\right)^2\left(\theta^2+4\theta+2\right)}{6\theta^3+19\theta^2+16\theta+4},$$

$$\underline{c}(\theta, \mathbf{g}) = 42 - 18\frac{(1+\theta)}{\theta}\left(1 - \frac{2\left(\theta+1\right)\left(\theta^2+4\theta+2\right)}{6\theta^3+19\theta^2+16\theta+4}\right)^2.$$

$$\overline{c}(\theta, \mathbf{g}) = 42 + 18\frac{(1+\theta)}{\theta}\left(1 - \frac{2\left(\theta+1\right)\left(\theta^2+4\theta+2\right)}{6\theta^3+19\theta^2+16\theta+4}\right)^2.$$

It is readily verified that

$$\underline{c}_{\mathcal{A}}(\theta, \mathbf{g}) < \overline{c}_{\mathcal{A}}(\theta, \mathbf{g}) < \overline{c}(\theta, \mathbf{g})$$

for all positive values of θ. However, as it can be seen from Figure 2, $\underline{c}_{\mathcal{A}}(\theta, \mathbf{g}) > \underline{c}(\theta, \mathbf{g})$ only when $\theta > \theta_0 \approx 10$, otherwise we have $\underline{c}_{\mathcal{A}}(\theta, \mathbf{g}) < \underline{c}(\theta, \mathbf{g})$.

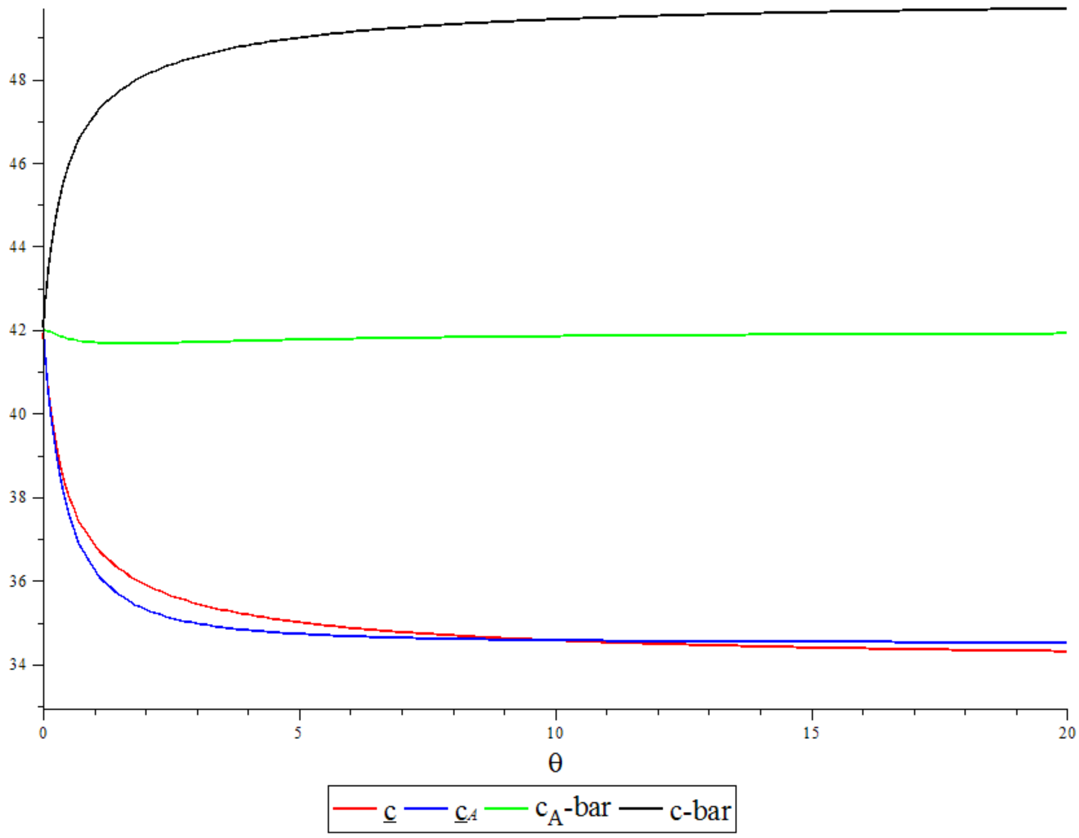

Figure 2: SPNE in a chain network

Therefore, for the chain network with $n = 4$, we have:

Case 1: Assume $\theta < \theta_0 \approx 10$.

(i) when $c < \underline{c}_{\mathcal{A}}(\theta, \mathbf{g})$, Adopting Equilibrium is a unique SPNE;

(ii) when $\underline{c}_{\mathcal{A}}(\theta, \mathbf{g}) < c < \underline{c}(\theta, \mathbf{g})$, there are two SPNE: Adopting Equilibrium and the asymmetric SPNE in which central (peripheral) agents adopt (do not adopt);

(iii) when $\underline{c}(\theta, \mathbf{g}) < c < \overline{c}_{\mathcal{A}}(\theta, \mathbf{g})$, there are three SPNE: Adopting Equilibrium, Not Adopting Equilibrium, and the asymmetric SPNE in which central (peripheral) agents adopt (do not adopt);

(iv) when $\overline{c}_{\mathcal{A}}(\theta, \mathbf{g}) < c < \overline{c}(\theta, \mathbf{g})$, there are two SPNE: Adopting Equilibrium and Not Adopting Equilibrium;

(v) when $c > \overline{c}(\theta, \mathbf{g})$, Not Adopting Equilibrium is the unique SPNE.

Case 2: Assume $\theta > \theta_0 \approx 10$.

(i) when $c < \underline{c}(\theta, \mathbf{g})$, Adopting Equilibrium is a unique SPNE;

(*ii*) when $\underline{c}(\theta, \mathbf{g}) < c < \underline{c}_{\mathcal{A}}(\theta, \mathbf{g})$, there are two SPNE: Adopting Equilibrium and Non-Adopting Equilibrium;

(*iii*) when $\underline{c}_{\mathcal{A}}(\theta, \mathbf{g}) < c < \overline{c}_{\mathcal{A}}(\theta, \mathbf{g})$, there are three SPNE: Adopting Equilibrium, Non-Adopting Equilibrium, and the asymmetric SPNE in which central (peripheral) agents adopt (do not adopt);

(*iv*) when $\overline{c}_{\mathcal{A}}(\theta, \mathbf{g}) < c < \overline{c}(\theta, \mathbf{g})$, there are two SPNE: Adopting Equilibrium and Non-Adopting Equilibrium;

(*v*) when $c > \overline{c}(\theta, \mathbf{g})$, Non-Adopting Equilibrium is the unique SPNE.

5 Policy implications

5.1 Policy implications: local average versus local aggregate model

We believe that it is important to be able to disentangle between different behavioral peer-effect models because they have different policy implications.

5.1.1 Education

There has been some debate in the United States of giving incentives to teachers. It is, however, difficult to determine which incentive to give to teachers in order to improve teacher quality. If the local aggregate model is at work among teachers, then we would need to have a teacher-based incentive policy since teachers will influence each other while, if it is the local average model, then one should implement a school-based incentive policy because this will be the only way to change the social norm of working hard among teachers.

If we now consider the students themselves, then the two models will be useful for policy implications. If the local-aggregate model is important in explaining students' education outcomes (Calvo-Armengol et al., 2009), then any individual-based policy (for example, vouchers) would be efficient. If, on the contrary, we believe that the local-average model is more important, then we should change the social norm in the school or the classroom and try to implement the idea that it is "cool" to work hard at school.[7]

[7] Focussing on charter schools in the United States where the aim is to change the social norm of studying, Angrist et al. (2010, 2012) and Curto and Fryer (2014) show that these schools have strong positive effects on average achievement such as reading and math scores.

To sum-up, 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 should be implemented while, for the local-aggregate model, this would not be necessary and an individual-based policy should instead be implemented.

5.1.2 Crime

It is well-documented that crime is, to some extent, a group phenomenon, and the source of crime is located in the intimate social networks of individuals (see e.g. Sutherland, 1947; Warr, 2002; Bayer et al., 2009).

In the local-aggregate model, a key-player policy (Ballester et al., 2006; Liu et al., 2012), whose aim is to remove the criminal that reduces total crime in a network the most, would be the most effective policy since the effort of each criminal and thus the sum of one's friends crime efforts will be reduced. 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 (see, in particular, Kleiman, 2009; Glaeser et al., 1996; Verdier and Zenou, 2004). That is, as the fraction of individuals participating in a criminal behavior increases, the impact on others is multiplied through social networks. Thus, criminal behaviors 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 will not affect the social norm of each group of friends in the network. 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, one needs to target a group or gang of criminals to drastically reduce crime. This example of crime clear 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.

5.2 Policy implications: Two-stage decisions

In this paper, we focused on the adoption of a new technology in the first stage and the productivity effort in the second stage. We believe that our framework is general enough so that it can encompass other applications. Here are some examples:

(i) *Assimilation choices*: In the first stage, each ethnic minority has to decide whether they have a low or high level of assimilation. It is, however, more costly ($c > 0$) to have a higher than a lower level of assimilation (learning a language, for example). In the second stage, given the choice of assimilation, more assimilated individuals have higher productivity effort (our α_i) than the low-assimilated ones ($\alpha > \alpha - t$) because they know better the culture of the host country. The social norm $\overline{x}_i$ for each individual i is the share of assimilated people among i's friends. In that case, all our results hold, and our model has strong policy implications in terms of assimilation policies.

(ii) *Tax evasion*: In the first stage, each agent has to decide the level of tax evasion (low or high). For example, they have to decide whether or not they want to put some of their money in a Swiss bank account or in an offshore bank account, given that it is more costly to evade more because the probability of being detected is higher. In the second stage, each agent has to decide how much income to conceal given that it is easier to conceal more income, the higher is the level of tax evasion in the first stage (if part of your money is in a Swiss bank account or an offshore bank account then it is easier to conceal income). In that case, the social norm $\overline{x}_i$ for each individual i is share of tax evaders among i 's friends. Again, all our results go through and our model has strong policy implications in terms of policies fighting against tax evasion.

(iii) *Crime*: In the first stage, each individual decides whether or not to join a gang (or anything that can improve the technology of committing crime). It is costly to join a gang because of the entry cost, but you adopt a better technology of committing crime. In the second stage, each criminal decides individually how much crime to commit given that you are more efficient in committing crime, if you have joined a gang in the first period (for example, because the probability of being caught is lower, etc.). In that case, the social norm $\overline{x}_i$ for each criminal i is share of criminals among i's friends. In that case, our model has strong policy implications in terms of policies fighting against crime.

6 Concluding remarks

We develop a two-stage model where agents, embedded in a social network, first decide whether or not to adopt a new costly technology, and, then, choose their level of productivity effort. The latter choice is affected by the social norm of each individual so that she loses utility from failing to conform to the average effort of her peers (her social norm). This is the so-called local-average model. Contrary to the local-aggregate model, we show that, in the second stage, if agents are ex ante identical but have different positions in the network, they all exert the same effort level, which corresponds to the first best. If agents are ex ante heterogenous in terms of productivities, then agents will exert effort according to her productivity so that individuals with higher productivities will exert higher efforts. We also give a condition on these productivities that ensures that the Nash equilibrium in efforts are always optimal. For example, for a star network, this condition boils down to the fact that the star-agent productivity has to be equal to the average productivity of all periphery agents. We show that to restore the first best the planner needs to subsidize agents who exert effort below that of their neighbors and tax those who exert effort above that of their neighbors. We then characterize the subgame-perfect Nash equilibria (SPNE) of the two-stage game by differentiating between symmetric (all agents within a network either adopt or do not adopt the new technology) and asymmetric (some agents in the network adopt while others do not) SPNE. We give the exact conditions on the cost of adopting the new technology and on the productivity cost of not adopting the new technology for which these equilibria exist. It turns out that, in a star network, only symmetric SPNE exists while, in a chain network, both symmetric and asymmetric SPNE exist.

We believe that our framework is rich enough to encompass many real-world situations where people make a binary choice decision first and then decide some effort level related to this choice in a framework where individuals are conformist and dislike to deviate from the social norm of their friends. We also believe that our results lead to important policy implications that can be tested empirically. We leave this for future research.

References

[1] Acemoglu, D., Ozdaglar, A. and A. Tahbaz-Salehi (2015), "Systemic risk and stability in financial networks," *American Economic Review* 105, 564-608.

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

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

[4] Asch, S. E. (1955), "Opinions and social pressure," *Scientific American* 193, 31-35.

[5] Asch, S. (1956), "Studies of independence and submission to group pressure: I. A minority of one against a unanimous majority," *Psychological Monographs* 70(9).

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

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

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

[9] Bernheim, B.D. (1994), "A Theory of Conformity," *Journal of Political Economy* 102, 841-877.

[10] Bisin, A. and T. Verdier (2000), "Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits," *Quarterly Journal of Economics* 115, 955-988.

[11] Bisin, A. and T. Verdier (2001), "The economics of cultural transmission and the dynamics of preferences," *Journal of Economic Theory* 97, 298-319.

[12] Bisin, A. and T. Verdier (2011), "The economics of cultural transmission and socialization," In J. Benhabib, A. Bisin, and M.O. Jackson (Eds.), *Handbook of Social Economics*, Amsterdam: Elsevier Science, pp. 339-416.

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

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

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

[16] Boucher, V., Bramoulle, Y., Djebbari, H., and B. Fortin (2014), "Do peers affect student achievement? Evidence from Canada using group size variation," *Journal of Applied Econometrics* 29, 91-109.

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

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

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

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

[21] Bramoulle, Y., Rogers, B.W. and A. Galeotti (2016), *The Oxford Handbook of the Economics of Networks*, Oxford: Oxford University Press.

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

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

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

[25] Dequiedt, V. and Y. Zenou (2013), "International migration, imperfect information, and brain drain," *Journal of Development Economics* 102, 62-78.

[26] Elliott, M., Golub, B. and M.O. Jackson (2014), "Financial networks and contagion," *American Economic Review* 104, 3115-3153.

[27] Fershtman, C., and Y. Weiss (1998), "Social Rewards, Externalities and Stable Preferences," *Journal of Public Economics* 70, 53-73.

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

[29] Haynie, D.L. (2001), "Delinquent peers revisited: Does network structure matter?" *American Journal of Sociology* 106, 1013-1057.

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

[31] Hsieh, C.-S. and X. Lin (2017), "Gender and Racial Peer Effects with Endogenous Network Formation," *Regional Science and Urban Economics*, forthcoming.

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

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

[34] Jackson, M.O., Rogers, B. and Y. Zenou (2017), "The economic consequences of social network structure", *Journal of Economic Literature* 55(1), forthcoming.

[35] Jackson M.O. and A. Watts (2002a), "The evolution of social and economic networks," *Journal of Economic Theory* 106, 265-295.

[36] Jackson M.O. and A. Watts (2002b), "On the formation of interaction networks in social coordination games," *Games and Economic Behavior* 41, 265-291.

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

[38] Kandel, E. and, E.P. Lazear (1992), "Peer Pressure and Partnerships," *Journal of Political Economy* 100, 801-817.

[39] Kandori, M., Mailath, G. and R. Rob (1993), "Learning, mutation, and long-run equilibria in games," *Econometrica* 61, 29-56.

[40] Leister, M., Zenou, Y. and J. Zhou (2017), "Coordination on networks", CEPR Discussion Paper No. 12392.

[41] Lin, X. (2010), "Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables," *Journal of Labor Economics* 28, 825-860.

[42] Lindquist, M.J., Sauermann, J. and Y. Zenou (2017), "Network effects on worker productivity," mimeo.

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

[44] Liu, X., Patacchini, E., Zenou, Y. and L-F. Lee (2012), "Criminal networks: Who is the key player?" CEPR Discussion Paper.

[45] Mangin, S. and Y. Zenou (2016), "Illegal migration and policy enforcement," *Economics Letters* 148, 83-86.

[46] Morris, S. (2000), "Contagion," *Review of Economic Studies* 67, 57-78.

[47] Morris, S. and H.S. Shin (2003), "Global games: Theory and applications," In: M. Dewatripont, L. Hansen and S. Turnovsky (Eds.), *Advances in Economics and Econometrics*, Cambridge: Cambridge University Press, pp. 57-114.

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

[49] Van den Bulte, C. and S. Stremersch (2004), "Social contagion and income heterogeneity in new product diffusion: A meta-analytic test," *Marketing Science* 23, 530-544.

[50] Verdier, T. and Y. Zenou (2017), "The role of social networks in cultural assimilation," *Journal of Urban Economics* 97, 15-39.

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

[52] Young, H.P. (1993), "The evolution of conventions," *Econometrica* 61, 57-84.

Appendix A: Proofs

Proof of Proposition 1.
Proof of part (i). By developing the utility function (3), we obtain:

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

The FOC is given by

$$x_i = \frac{\alpha_i}{1+\theta} + \frac{\theta}{1+\theta}\overline{x}_i. \tag{25}$$

Plugging the definition (2) of the social norm $\overline{x}_i$ into the (33) implies (4). Restating (4) in matrix form, we get (5). By solving (5) for $\mathbf{x}$, we verify that the Nash equilibrium $\mathbf{x}^*$ is indeed defined by (6). Existence and uniqueness of the Nash equilibrium $\mathbf{x}^*$ is guaranteed by the fact that, for any $\theta > 0$, we have:

$$\frac{\theta}{1+\theta} < \lambda_1(\widehat{\mathbf{G}}) = 1,$$

where $\lambda_1(\widehat{\mathbf{G}})$ stands for the largest Eigenvalue of $\widehat{\mathbf{G}}$, while $\lambda_1(\widehat{\mathbf{G}}) = 1$ holds true because $\widehat{\mathbf{G}}$ is a row-normalized matrix with non-negative entries. This proves part (i). ■

Proof of part (ii). Using (6) and observing that $\overline{\mathbf{x}}^* = \widehat{\mathbf{G}}\mathbf{x}^*$, we obtain equation (7) for the equilibrium social norms. This proves part (ii). ■

Proof of part (iii). We now prove that the equilibrium utility levels are given by (8). To do this, we use the first-order condition (33) to express agent i's equilibrium social norm $\overline{x}_i^*$ as follows:

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

Plugging this expression into the utility function (3), we obtain after simplifications:

$$U_i(x_i^*, \mathbf{x}_{-i}^*, \mathbf{g}) = \left(\frac{1+\theta}{\theta}\right)\alpha_i x_i^* - \frac{1}{2}\left(\frac{1+\theta}{\theta}\right)x_i^{*2} - \frac{1}{2\theta}\alpha_i^2.$$

This, in turn, can be written as:

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

which immediately implies

$$U_i(x_i^*, \mathbf{x}_{-i}^*, \mathbf{g}) = \frac{1}{2}\left[\alpha_i^2 - \left(\frac{1+\theta}{\theta}\right)(\alpha_i - x_i^*)^2\right] \tag{26}$$

Plugging agent i's equilibrium effort $x_i^* = \sum_{j=1}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\alpha_j$ into (26) yields (8) and proves part (iii). ■

Proof of part (iv). We first prove the following Lemma, which holds for any vector $\boldsymbol{\alpha}$ of productivities.

Lemma 4 For each $i \in \mathcal{N}$, the following trichotomy always holds: either $\alpha_i > x_i^* > \overline{x}_i^*$, or $\alpha_i = x_i^* = \overline{x}_i^*$, or $\alpha_i < x_i^* < \overline{x}_i^*$.

Proof of Lemma 4. Restating the first-order condition as $\alpha_i - x_i^* = \theta(x_i^* - \overline{x}_i^*)$ proves the result. ■

Assume now that there are only two levels of productivity: $\alpha_i \in \{\alpha^L, \alpha^H\}$ for all $i \in \mathcal{N}$, where $\alpha^H > \alpha^L > 0$. Then, we have for each i:

$$\alpha^H > x_i^* = \left(\sum_{j:\alpha_j=\alpha^H} \frac{\widehat{m}_{ij}(\theta)}{1+\theta}\right)\alpha^H + \left(\sum_{j:\alpha_j=\alpha^L} \frac{\widehat{m}_{ij}(\theta)}{1+\theta}\right)\alpha^L > \alpha^L,$$

combining which with Lemma 4 proves part (iv). ■

The proof of Proposition 1 is now completed. ▌

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

Lemma 5 *The matrices* $\frac{1}{1+\theta}\widehat{\mathbf{M}}$ *and* $\frac{1}{1+\theta}\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ *are row-normalized.*

Proof. Let us first prove that $\frac{1}{1+\theta}\widehat{\mathbf{M}}$ is row-normalized, that is

$$\frac{1}{1+\theta}\widehat{\mathbf{M}} \cdot \mathbf{1} = \mathbf{1},$$

where $\mathbf{1} \equiv (1, ..., 1)^T$. Using the series decomposition of $\widehat{\mathbf{M}}$, we obtain:

$$\frac{1}{1+\theta}\widehat{\mathbf{M}} \cdot \mathbf{1} = \frac{1}{1+\theta}\sum_{k=0}^{\infty}\left(\frac{\theta}{1+\theta}\right)^k \widehat{\mathbf{G}}^k \cdot \mathbf{1}.$$

Because $\widehat{\mathbf{G}}$ is row-normalized, $\widehat{\mathbf{G}}^k$ for any integer k is also row-normalized, hence

$$\frac{1}{1+\theta}\sum_{k=0}^{\infty}\left(\frac{\theta}{1+\theta}\right)^{k}\widehat{\mathbf{G}}^{k}\cdot\mathbf{1}=\left(\frac{1}{1+\theta}\sum_{k=0}^{\infty}\left(\frac{\theta}{1+\theta}\right)^{k}\right)\cdot\mathbf{1}=\mathbf{1}.$$

This proves that $\frac{1}{1+\theta}\widehat{\mathbf{M}}$ is row-normalized. Now since the product of two row-normalized matrices is a row-normalized matrix, then this shows that $\frac{1}{1+\theta}\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ is a row-normalized matrix. This completes the proof of Lemma 5. ■

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

$$\mathbf{x}^*=\frac{1}{1+\theta}\widehat{\mathbf{M}}\boldsymbol{\alpha} \tag{27}$$

where

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

Hence,

$$\frac{\partial\mathbf{x}^*}{\partial\boldsymbol{\alpha}}=\frac{1}{1+\theta}\widehat{\mathbf{M}}$$

which is strictly positive. Since, by Lemma 5, $\frac{1}{1+\theta}\widehat{\mathbf{M}}$ is a row-normalized matrix with non-negative entries, it must be that $\frac{\partial x_i^*}{\partial\alpha_i}<1$ for any connected network.

Let us now prove that $0<\frac{\partial\overline{x}_i^*}{\partial\alpha_i}<1$. By definition of the social norm, we have $\overline{\mathbf{x}}=\widehat{\mathbf{G}}\mathbf{x}$, whence the equilibrium vector of local averages is given by

$$\overline{\mathbf{x}}^*=\frac{1}{1+\theta}\widehat{\mathbf{G}}\widehat{\mathbf{M}}\boldsymbol{\alpha}=\widehat{\mathbf{G}}\mathbf{x}^* \tag{28}$$

As seen from (28), $\frac{\partial\overline{x}_i^*}{\partial\alpha_i}$ is the iith entry of $\frac{1}{1+\theta}\widehat{\mathbf{G}}\widehat{\mathbf{M}}$. 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 5, $\frac{1}{1+\theta}\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ 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:

$$U_i(x_i,\mathbf{x}_{-i},\mathbf{g})=\alpha_i x_i-\frac{1}{2}x_i^2-\frac{\theta}{2}(x_i-\overline{x}_i)^2. \tag{29}$$

Let us first determine $\frac{\partial U_i^*}{\partial\alpha_i}$. In equilibrium, x_i is determined as the maximizer of (29) under $\mathbf{x}_{-i}=\mathbf{x}_{-i}^*$. Applying the envelope theorem yields:

$$\frac{\partial U_i^*}{\partial\alpha_i}=x_i^*+\theta(x_i^*-\overline{x}_i^*)\frac{\partial\overline{x}_i^*}{\partial\alpha_i}. \tag{30}$$

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

Lemma 6 *The following inequalities hold for all $i = 1, ..., n$:*

$$\max_j \alpha_j \geq \max\{x_i^*, \alpha_i\} \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq \min\{x_i^*, \alpha_i\} \geq \min_j\{\alpha_j\} \tag{31}$$

Proof of Lemma 6: Let us, first, establish the second and the third inequalities in (31):

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

If $x_i^* \geq \overline{x}_i^*$, then it follows from equation (30) and Proposition 2 that

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

If, on the contrary, $x_i^* < \overline{x}_i^*$, then equation (30) and Proposition 2 imply

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

In sum, we have:

$$\max\{x_i^*,\ x_i^* + \theta(x_i^* - \overline{x}_i^*)\} \geq \frac{\partial U_i^*}{\partial \alpha_i} \geq \min\{x_i^*,\ x_i^* + \theta(x_i^* - \overline{x}_i^*)\}. \tag{32}$$

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

$$x_i^* + \theta(x_i^* - \overline{x}_i^*) = \alpha_i. \tag{33}$$

Combining (32) with (33) 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 5, $\frac{1}{1+\theta}\widehat{\mathbf{M}}$ is a row-normalized matrix. Since

$$\mathbf{x}^* = \frac{1}{1+\theta}\widehat{\mathbf{M}}\ \boldsymbol{\alpha},$$

or, equivalently,

$$x_i^* = \sum_j \frac{\widehat{m}_{ij}}{1+\theta}\alpha_j$$

This implies that:

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

This comletes the proof of Lemma 6. ■

It remains to observe that Lemma 6 immediately implies our result, that is: $\frac{\partial U_i^*}{\partial \alpha_i} > 0$.

Proof of part (iii): As implied by (8), U_i^* is a strictly concave quadratic function of α_j for any $j \in \mathcal{N} \setminus \{i\}$. 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 takes place depends on the sign of the partial derivative $\partial U_i^* / \partial \alpha_j$ evaluated at $\alpha_j = 0$. Computing this derivative yields:

$$\left.\frac{\partial U_i^*}{\partial \alpha_j}\right|_{\alpha_j=0} = \frac{1+\theta}{\theta}\left(\alpha_i - \sum_{k\in\mathcal{N}\setminus\{j\}} \frac{\widehat{m}_{ik}(\theta)}{1+\theta}\alpha_k\right).$$

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

$$\alpha_i > \sum_{k\in\mathcal{N}\setminus\{j\}} \frac{\widehat{m}_{ik}(\theta)}{1+\theta}\alpha_k.$$

Otherwise, U_i^* strictly decreases in α_j. This proves part (iii). ■

Proof of part (iv): We obtain, first, 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.* We have:

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

where

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

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

$$\text{sign}\left[\frac{\partial U_i^*}{\partial \alpha_j}\right] = \text{sign}\left(x_i^* - \overline{x}_i^*\right). \tag{34}$$

Assume now that there are only two levels of productivity: $\alpha_i \in \{\alpha^L, \alpha^H\}$ for all $i \in \mathcal{N}$, where $\alpha^H > \alpha^L > 0$. Then, combining (34) with part (iv) of Proposition 1 completes the proof. ■

Proof of Proposition 3.

Proof of part (i): Fix $j \in \mathcal{N}$ and write the welfare as follows:

$$\mathcal{W} = U_j(x_j, \mathbf{x}_{-j}, \mathbf{g}) + \sum_{i \in \mathcal{N} \diagdown \{j\}} U_i(x_i, \mathbf{x}_{-i}, \mathbf{g})$$

The social planner solves the welfare maximization problem:

$$\max_{x_1,\ldots,x_n} \mathcal{W}.$$

Differentiating $\mathcal{W}$ with respect to x_j yields:

$$\begin{aligned}\frac{\partial \mathcal{W}}{\partial x_j} &= \frac{\partial U_j}{\partial x_j} + \sum_{i \in \mathcal{N} \diagdown \{j\}} \frac{\partial U_i}{\partial \overline{x}_i} \frac{\partial \overline{x}_i}{\partial x_j} \\ &= \alpha_j - x_j - \theta \left(x_j - \overline{x}_j \right) + \theta \sum_{i \in \mathcal{N} \diagdown \{j\}} \widehat{g}_{ij} \left(x_i - \overline{x}_i \right)\end{aligned}$$

Therefore, the first-order conditions for this problem are given by:[8]

$$x_j = \frac{\alpha_j}{1+\theta} + \frac{\theta}{1+\theta} \overline{x}_j + \frac{\theta}{1+\theta} \sum_{i \in \mathcal{N} \diagdown \{j\}} \widehat{g}_{ij} \left(x_i - \overline{x}_i \right).$$

Restating these conditions in matrix form yields (12) and proves part (i). ■
Proof of part (ii). The first-order condition of the planner problem is given by:

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

To have $\mathbf{x}^* = \mathbf{x}^O$, we need:

$$\widehat{\mathbf{G}}^T \left(\mathbf{I} - \widehat{\mathbf{G}} \right) \mathbf{x}^* = 0. \tag{35}$$

Using (6), this is equivalent to:

$$\widehat{\mathbf{G}}^T \left(\mathbf{I} - \widehat{\mathbf{G}} \right) \frac{1}{1+\theta} \widehat{\mathbf{M}} \boldsymbol{\alpha} = \mathbf{0}$$

When $\boldsymbol{\alpha} = \alpha \mathbf{1}$, then this is always true, since both $\widehat{\mathbf{G}}$ and $\frac{1}{1+\theta} \widehat{\mathbf{M}}$ are row-normalized.

It remains to prove that the converse is true when $\det(\widehat{\mathbf{G}}) \neq 0$. To do this, we observe that the best replies (4) are given by

[8] It is easily checked that there is a unique maximum for each x_j.

$$\mathbf{x}^* = \frac{1}{1+\theta}\boldsymbol{\alpha} + \frac{\theta}{1+\theta}\widehat{\mathbf{G}}\mathbf{x}^*.$$

This, in turn, implies

$$\theta\widehat{\mathbf{G}}\mathbf{x}^* = (1+\theta)\mathbf{x}^* - \boldsymbol{\alpha},$$

combining which with (35) yields:

$$\widehat{\mathbf{G}}^T(\boldsymbol{\alpha} - \mathbf{x}^*) = \widehat{\mathbf{G}}^T\left(\mathbf{I} - \frac{1}{1+\theta}\widehat{\mathbf{M}}\right)\boldsymbol{\alpha} = \mathbf{0}. \tag{36}$$

If $\det(\widehat{\mathbf{G}}) \neq 0$, $\widehat{\mathbf{G}}^T$ is an invertible matrix, which means:

$$\widehat{\mathbf{G}}^T\left(\mathbf{I} - \frac{1}{1+\theta}\widehat{\mathbf{M}}\right)\boldsymbol{\alpha} = \mathbf{0} \Longleftrightarrow \left(\mathbf{I} - \frac{1}{1+\theta}\widehat{\mathbf{M}}\right)\boldsymbol{\alpha} = \mathbf{0} \Longleftrightarrow \boldsymbol{\alpha} = \alpha\mathbf{1}.$$

This proves the results. ▮

Proof of Proposition 4 Omitted. ▮

Proof of Corollary 1

(i) Let us derive (13) for complete bipartite graphs: $g = K_{m,n}$. The Nash equilibrium $\mathbf{x}^*$ is the solution to (5). A necessary and sufficient condition (13) for the Nash equilibrium $\mathbf{x}^*$ to deliver a first best is given by

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

Using (5), this condition can be equivalently restated as follows:

$$\widehat{\mathbf{G}}^T\boldsymbol{\alpha} = \widehat{\mathbf{G}}^T\overline{\mathbf{x}}^*, \tag{37}$$

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

When $\mathbf{g} = K_{m,n}$, the row-normalized adjacency matrix $\widehat{\mathbf{G}}$ is given by

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

where $\mathbf{0}_{p\times q}$ and $\mathbf{1}_{p\times q}$ stand for $(p\times q)$-matrices of, respectively, zeros and ones.

Then, the best reply functions (4) take the form:

$$x_i = \frac{1}{1+\theta}\alpha_i + \begin{cases} \frac{\theta}{1+\theta}\frac{1}{n}\sum_{k\in V_2} x_k, & i \in V_1, \\ \frac{\theta}{1+\theta}\frac{1}{m}\sum_{k\in V_1} x_k, & i \in V_2. \end{cases} \tag{38}$$

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

$$\frac{1}{m}\sum_{k\in V_1} x_k = \frac{1}{1+\theta}\frac{1}{m}\sum_{k\in V_1}\alpha_k + \frac{\theta}{1+\theta}\frac{1}{n}\sum_{k\in V_2} x_k, \tag{39}$$

$$\frac{1}{n}\sum_{k\in V_2} x_k = \frac{1}{1+\theta}\frac{1}{n}\sum_{k\in V_2}\alpha_k + \frac{\theta}{1+\theta}\frac{1}{m}\sum_{k\in V_1} x_k. \tag{40}$$

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 $j \in V_2$. Assume without loss of generality that agent $i = 1$ belongs to V_1, while agent $i = 2$ belongs to V_2. Then, we have

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

Plugging (39) and (40) into (38), we get for all $i \in V_1$:

$$x_i = \frac{1}{1+\theta}\alpha_i + \frac{\theta}{(1+\theta)^2}\overline{\alpha}_2 + \left(\frac{\theta}{1+\theta}\right)^2 \overline{x}_2;$$

for all $i \in V_2$:

$$x_i = \frac{1}{1+\theta}\alpha_i + \frac{\theta}{(1+\theta)^2}\overline{\alpha}_1 + \left(\frac{\theta}{1+\theta}\right)^2 \overline{x}_1.$$

Computing the means yields:

$$\overline{x}_r = \frac{1}{1+\theta}\overline{\alpha}_s + \frac{\theta}{(1+\theta)^2}\overline{\alpha}_r + \left(\frac{\theta}{1+\theta}\right)^2 \overline{x}_r,$$

where $r, s = 1, 2$, $r \neq s$. Solving for $\overline{x}_r$, we get

$$\overline{x}_r^* = \frac{\theta}{1+2\theta}\overline{\alpha}_r + \frac{1+\theta}{1+2\theta}\overline{\alpha}_s. \tag{41}$$

So, if $i \in V_r$, we get

$$\alpha_i - \overline{x}_i^* = \alpha_i - \frac{\theta}{1+2\theta}\overline{\alpha}_r - \frac{1+\theta}{1+2\theta}\overline{\alpha}_s.$$

It remains to notice that

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

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

Hence, the condition (37) holds if and only if $\overline{\alpha}_r = \overline{x}_r^*$ for $r = 1, 2$. Using (41), we find that this is equivalent to $\overline{\alpha}_1 = \overline{\alpha}_2$. This completes the proof.

(ii) Let us not derive (13) for complete chains: As $\mathbf{g}$ is a chain, the row-normalized adjacency matrix $\widehat{\mathbf{G}}$ is given by

$$\widehat{\mathbf{G}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \cdots & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Then, (4) can be restated in coordinate form as follows:

$$x_i = \frac{1}{1+\theta}\alpha_i + \begin{cases} \frac{\theta}{1+\theta}x_2, & i = 1, \\ \frac{\theta}{1+\theta}x_{n-1}, & i = n, \\ \frac{\theta}{1+\theta}\frac{x_{i-1}+x_{i+1}}{2}, & \text{otherwise.} \end{cases} \tag{42}$$

The condition (37) becomes:

$$\left(\frac{\alpha_2}{2}, \alpha_1 + \frac{\alpha_3}{2}, \frac{\alpha_2+\alpha_4}{2}, \ldots, \frac{\alpha_{n-3}+\alpha_{n-1}}{2}, \frac{\alpha_{n-2}}{2} + \alpha_n, \frac{\alpha_{n-1}}{2} \right) =$$

$$\left(\frac{\overline{x}_2^*}{2} \quad , \overline{x}_1^* + \frac{\overline{x}_3^*}{2} \quad , \frac{\overline{x}_2^* + \overline{x}_4^*}{2} \quad , \ldots , \quad \frac{\overline{x}_{n-3}^* + \overline{x}_{n-1}^*}{2} \quad , \frac{\overline{x}_{n-2}^*}{2} + \overline{x}_n^* \quad , \frac{\overline{x}_{n-1}^*}{2} \right),$$

combining which with (42), we get

$$x_2^* = \alpha_2, \qquad x_{n-1}^* = \alpha_{n-1},$$

$$x_1^* = \frac{1}{1+\theta}\alpha_1 + \frac{\theta}{1+\theta}\alpha_2, \qquad x_n^* = \frac{1}{1+\theta}\alpha_n + \frac{\theta}{1+\theta}\alpha_{n-1},$$

while for all i such that $3 \le i \le n-2$, we have:

$$\frac{x_{i-1}^* + x_{i+1}^*}{2} = \frac{\alpha_{i-1} + \alpha_{i+1}}{2},$$

which implies (i) that $x_i^* = \alpha_i$ for all even values of i, (ii) that $x_i^* = \alpha_i$ for all i such that $n-i$ are odd, and (iii) that, for all i such that $3 \le i \le n-2$, we have:

$$x_i^* = \frac{1}{1+\theta}\alpha_i + \frac{\theta}{1+\theta}\frac{\alpha_{i-1} + \alpha_{i+1}}{2}, \quad 3 \le i \le n-2.$$

In addition, we need

$$\alpha_1 + \frac{\alpha_3}{2} = \overline{x}_1^* + \frac{\overline{x}_3^*}{2} = x_2^* + \frac{1}{2}\frac{x_2^* + x_4^*}{2} = \frac{5}{4}\alpha_2 + \frac{1}{4}\alpha_4.$$

$$\alpha_n + \frac{\alpha_{n-2}}{2} = \overline{x}_n^* + \frac{\overline{x}_{n-2}^*}{2} = x_{n-1}^* + \frac{1}{2}\frac{x_{n-1}^* + x_{n-3}^*}{2} = \frac{5}{4}\alpha_{n-1} + \frac{1}{4}\alpha_{n-3}.$$

If n is even, then $x_i^* = \alpha_i$ is also true for odd values of i, as implied by $x_{n-1}^* = \alpha_{n-1}$. Thus, we have $\mathbf{x}^* = \boldsymbol{\alpha}$, which is only possible when $\alpha_i = \alpha_j$ for all $i, j \in \mathcal{N}$.

If $4 \le i \le n-3$, then it must be that:

$$\frac{1}{1+\theta}\alpha_i + \frac{\theta}{1+\theta}\frac{\alpha_{i-1} + \alpha_{i+1}}{2} = \frac{1}{1+\theta}\alpha_i + \frac{\theta}{1+\theta}\frac{\frac{1}{1+\theta}\alpha_{i-1} + \frac{\theta}{1+\theta}\frac{\alpha_{i-2}+\alpha_i}{2} + \frac{1}{1+\theta}\alpha_{i+1} + \frac{\theta}{1+\theta}\frac{\alpha_i + \alpha_{i+2}}{2}}{2}$$

$$\frac{\alpha_{i-1} + \alpha_{i+1}}{2} = \frac{1}{1+\theta}\frac{\alpha_{i-1} + \alpha_{i+1}}{2} + \frac{\theta}{1+\theta}\frac{\alpha_{i-2} + 2\alpha_i + \alpha_{i+2}}{4}$$

$$\frac{\alpha_{i-1} + \alpha_{i+1}}{2} = \frac{\alpha_{i-2} + 2\alpha_i + \alpha_{i+2}}{4}$$

$$\alpha_{i-2} - 2\alpha_{i-1} + 2\alpha_i - 2\alpha_{i+1} + \alpha_{i+2} = 0$$

If $i = 3$ or $i = n-2$:

$$\frac{1}{1+\theta}\alpha_3 + \frac{\theta}{1+\theta}\frac{\alpha_2+\alpha_4}{2} = \frac{1}{1+\theta}\alpha_3 + \frac{\theta}{1+\theta}\frac{\alpha_2+x_4^*}{2}$$

$$x_4^* = \alpha_4 = \frac{1}{1+\theta}\alpha_4 + \frac{\theta}{1+\theta}\frac{\alpha_3+\alpha_5}{2}$$

$$\alpha_4 = \frac{\alpha_3+\alpha_5}{2}, \qquad \alpha_{n-3} = \frac{\alpha_{n-2}+\alpha_{n-4}}{2}.$$

If $i = 2$ or $i = n-1$:

$$\alpha_2 = \frac{1}{1+\theta}\alpha_2 + \frac{\theta}{1+\theta}\frac{x_1^*+x_3^*}{2}$$

$$\alpha_2 = \frac{1}{1+\theta}\alpha_2 + \frac{\theta}{1+\theta}\frac{\frac{1}{1+\theta}\alpha_1 + \frac{\theta}{1+\theta}\alpha_2 + \frac{1}{1+\theta}\alpha_3 + \frac{\theta}{1+\theta}\frac{\alpha_2+\alpha_4}{2}}{2}$$

$$\alpha_2 = \frac{\frac{1}{1+\theta}\alpha_1 + \frac{\theta}{1+\theta}\alpha_2 + \frac{1}{1+\theta}\alpha_3 + \frac{\theta}{1+\theta}\frac{\alpha_2+\alpha_4}{2}}{2}$$

$$\alpha_2 = \frac{1}{2(1+\theta)}\alpha_1 + \frac{3\theta}{4(1+\theta)}\alpha_2 + \frac{1}{2(1+\theta)}\alpha_3 + \frac{\theta}{4(1+\theta)}\alpha_4$$

$$\frac{4+\theta}{4(1+\theta)}\alpha_2 = \frac{1}{2(1+\theta)}\alpha_1 + \frac{1}{2(1+\theta)}\alpha_3 + \frac{\theta}{4(1+\theta)}\alpha_4$$

$$\alpha_2 = \frac{2}{4+\theta}\alpha_1 + \frac{2}{4+\theta}\alpha_3 + \frac{\theta}{4+\theta}\alpha_4,$$

$$\alpha_{n-1} = \frac{2}{4+\theta}\alpha_n + \frac{2}{4+\theta}\alpha_{n-2} + \frac{\theta}{4+\theta}\alpha_{n-3}.$$

To sum up, $\boldsymbol{\alpha}$ must satisfy the following linear relationships:

$$\frac{5}{6}\alpha_2 + \frac{1}{6}\alpha_4 = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_3,$$

$$\alpha_2 = \frac{2}{4+\theta}\alpha_1 + \frac{2}{4+\theta}\alpha_3 + \frac{\theta}{4+\theta}\alpha_4,$$

$$\alpha_4 = \frac{\alpha_3+\alpha_5}{2},$$

$$\frac{\alpha_{i-1}+\alpha_{i+1}}{2}=\frac{\alpha_{i-2}+2\alpha_i+\alpha_{i+2}}{4}, \qquad i=4,...,n-3,$$

$$\alpha_{n-3}=\frac{\alpha_{n-2}+\alpha_{n-4}}{2},$$

$$\frac{5}{6}\alpha_{n-1}+\frac{1}{6}\alpha_{n-3}=\frac{2}{3}\alpha_n+\frac{1}{3}\alpha_{n-2},$$

$$\alpha_{n-1}=\frac{2}{4+\theta}\alpha_n+\frac{2}{4+\theta}\alpha_{n-2}+\frac{\theta}{4+\theta}\alpha_{n-3}.$$

For example, when $n=7$, the above system of linear constraints boils down to:

$$\frac{5}{6}\alpha_2+\frac{1}{6}\alpha_4=\frac{2}{3}\alpha_1+\frac{1}{3}\alpha_3,$$

$$\alpha_2=\frac{2}{4+\theta}\alpha_1+\frac{2}{4+\theta}\alpha_3+\frac{\theta}{4+\theta}\alpha_4,$$

$$\alpha_4=\frac{\alpha_3+\alpha_5}{2},$$

$$\frac{\alpha_3+\alpha_5}{2}=\frac{\alpha_2+2\alpha_4+\alpha_6}{4},$$

$$\frac{5}{6}\alpha_6+\frac{1}{6}\alpha_4=\frac{2}{3}\alpha_7+\frac{1}{3}\alpha_5,$$

$$\alpha_6=\frac{2}{4+\theta}\alpha_7+\frac{2}{4+\theta}\alpha_5+\frac{\theta}{4+\theta}\alpha_4.$$

It is straightforward to check directly that a positive vector of productivities is a solution to this system if and only if it satisfies the constraints given in Example 3 of Section 2.4.1. ■

Proof of Proposition 5.

Proof of part (i): Using the Nash equilibrium conditions (4) and the definition of a social norm $\overline{x}_i$, it can be shown that the equilibrium welfare is given by

$$\mathcal{W}:=\frac{1}{1+\theta}\left[\frac{1}{2}\sum_{j\in\mathcal{N}}\alpha_j^2+\theta\sum_{j\in\mathcal{N}}\alpha_j\overline{x}_j^*-\frac{\theta}{2}\sum_{j\in\mathcal{N}}\left(\overline{x}_j^*\right)^2\right]. \tag{43}$$

Because $\frac{\partial \mathcal{W}}{\partial \alpha_i}|_{\theta=0} = \alpha_i > 0$, and because $\frac{\partial \mathcal{W}}{\partial \alpha_i}$ is continuous in θ, there must exist a strictly positive threshold value, $\widetilde{\theta} > 0$, of θ, such that, for any $\theta < \widetilde{\theta}$ and for any $i \in \mathcal{N}$, we have:

$$\frac{\partial \mathcal{W}}{\partial \alpha_i} > 0.$$

This proves part (i). ■

Proof of part (ii). The results of this part follow immediately from part (iii) of Proposition 2. ■

Proof of part (iii). The Rawlsian welfare critetion given by

$$\mathcal{W}^{\mathrm{R}} := \min_{j \in \mathcal{N}} \left\{ U_j^* \right\} = \frac{1}{1+\theta} \min_{i \in \mathcal{N}} \left\{ \frac{1}{2}\alpha_i^2 + \theta \alpha_i \overline{x}_i^* - \frac{\theta}{2} \left(\overline{x}_i^* \right)^2 \right\}, \tag{44}$$

As implied by (44), we have $\lim_{\alpha_j \to \infty} \mathcal{W}^{\mathrm{R}} = -\infty$ for any $j \in \mathcal{N}$ and for any $\theta > 0$. To sum up, *the welfare criteria given by (43) and (44) generate the opposite welfare consequences of someone's individual productivity unbounded growth.*
■

Proof of Proposition 6: In the first stage of the game, the strategy space of each agent i is given by $s_i \in \{0, 1\}$, where $s_i = 0$ means that agent i chooses *not* to adopt the new technology while $s_i = 1$ means that the agent adopts the technology. Assume that individual i chooses to deviate from a given strategic profile $\mathbf{s} = (s_1, ..., s_n)$. Denote the resulting change in her payoff by ΔU_i. If agent i individually deviates from $\mathbf{s}$, her productivity α_i changes either from α to $\alpha - t$, or from $\alpha - t$ to α. In both cases, we have $|\Delta \alpha_i| = t$. Let us restate (8) as follows:

$$2U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}) = \alpha_i^2 - \left(\frac{1+\theta}{\theta} \right) \left[\left(1 - \frac{\widehat{m}_{ii}}{(1+\theta)} \right) \alpha_i - \sum_{j \neq i}^{n} \frac{\widehat{m}_{ij}}{(1+\theta)} \alpha_j \right]^2, \tag{45}$$

and define individual gains from a unilateral deviation by

$$\Delta U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}) := U_i^*(\alpha_i + \Delta\alpha_i, \boldsymbol{\alpha}_{-i}) - U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}). \tag{46}$$

By combining (45) with (46), we get:

$$\begin{aligned} &2\Delta U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}) = \Delta\alpha_i \left(2\alpha_i + \Delta\alpha_i \right) \qquad (47) \\ &- \left(\frac{1+\theta}{\theta} \right) \left(1 - \frac{\widehat{m}_{ii}}{1+\theta} \right) \Delta\alpha_i \left[\left(1 - \frac{\widehat{m}_{ii}}{1+\theta} \right) \left(2\alpha_i + \Delta\alpha_i \right) - 2 \sum_{j=1, j \neq i}^{n} \frac{\widehat{m}_{ij}}{1+\theta} \alpha_j \right] \end{aligned}$$

We first derive a necessary and sufficient condition for the Adopting Equilibrium to exist, in which $\mathcal{A} = \mathcal{N}$. In that case, a unilateral deviation of agent i means switching from adoption to non-adoption. More precisely, for all $i \in \mathcal{N}$, we have: $\alpha_i = \alpha$, $\Delta\alpha_i = -t$. The necessary and sufficient condition for such deviations to be unprofitable for all agents is given by

$$2c \leq -2\Delta U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}) \tag{48}$$

for all $i \in \mathcal{N}$. By combining (48) with (46), we obtain:

$$2c \leq 2\alpha t - t^2 - \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) t \left[\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)(2\alpha - t) - 2\alpha \sum_{j=1, j\neq i}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\right].$$

Furthermore, by Lemma 5, the matrix $\frac{1}{1+\theta}\widehat{\mathbf{M}}$ is row-normalized (i.e. $\sum_{j=1}^{n} \frac{\widehat{m}_{ij}}{1+\theta} = 1$), which implies

$$\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)(2\alpha - t) - 2\alpha \sum_{j=1, j\neq i}^{n} \frac{\widehat{m}_{ij}}{1+\theta} = 2\alpha\left(1 - \sum_{j=1}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\right) = 0.$$

Therefore, the condition (48) is equivalent to:

$$c \leq \alpha t - \frac{t^2}{2}\left[1 - \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)^2\right]$$

for all $i \in \mathcal{N}$. Observe that, for all $i \in \mathcal{N}$, the definition of $\overline{c}(\theta, \mathbf{g})$, given by (17), can restated as

$$\overline{c}(\theta, \mathbf{g}) = \min_{i \in \mathcal{N}} \left\{\alpha t - \frac{t^2}{2}\left[1 - \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)^2\right]\right\}$$

Hence, a necessary and sufficient condition for the Adopting Equilibrium to exist is given by

$$c \leq \overline{c}(\theta, \mathbf{g}). \tag{49}$$

We now derive a necessary and sufficient condition for the Non-Adopting Equilibrium to exist, in which $\mathcal{A} = \varnothing$. In that case, a unilateral deviation of agent i means switching from non-adoption to adoption: $\Delta\alpha_i = -t$. The necessary and sufficient condition for such deviations to be unprofitable for all agents is given by

$$2c > 2\Delta U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}) \tag{50}$$

for all $i \in \mathcal{N}$. Using (50) and (46) yields

$$2c > t\,(2\,(\alpha - t) + t)$$

$$-\left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)t\left[\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)(2\,(\alpha - t) + t) - 2\,(\alpha - t)\sum_{j=1, j\neq i}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\right].$$

Since, by Lemma 5, the matrix $\frac{1}{1+\theta}\widehat{\mathbf{M}}$ is row-normalized, (50) is equivalent to the following system of inequalities:

$$c > \alpha t - \frac{t^2}{2}\left[1 + \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)^2\right],$$

for all $i \in \mathcal{N}$. Observe that, for all $i \in \mathcal{N}$, the definition of $\underline{c}(\theta, \mathbf{g})$, given by (18), can restated as

$$\underline{c}(\theta, \mathbf{g}) = \max_{i \in \mathcal{N}}\left\{\alpha t - \frac{t^2}{2}\left[1 + \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)^2\right]\right\}$$

Hence, a necessary and sufficient condition for the Not Adopting Equilibrium to exist is given by

$$c > \underline{c}(\theta, \mathbf{g}). \tag{51}$$

Combining (49) with (51) completes the proof. ▮

Proof of Corollary 3: Using (17)−(19), it is readily verified that a necessary and sufficient condition $\underline{c}(\theta, \mathbf{g}) < c < \overline{c}(\theta, \mathbf{g})$ for multiple symmetric SPNE to exist can be equivalently reformulated by (20). ▮

Proof of Proposition 7: Let us start with the following Lemma:

Lemma 7 *For all $i \in \mathcal{N}$, we have:*

$$\alpha t - t^2 \leq |\Delta U_i^*| \leq \alpha t \tag{52}$$

Proof of Lemma 7: To prove (52), we prove a more general result, that is, when $\alpha_i \in [\alpha - t, \alpha]$ (instead of taking only two values: $\alpha - t$ and α), the following inequality holds:

$$\min_{j=1,\ldots n} \{\alpha_j\} \, t \leq |\Delta U_i| \leq \max_{j=1,\ldots n} \{\alpha_j\} \, t. \tag{53}$$

By the mean value theorem, there exists $\lambda \in (0,1)$ such that

$$\Delta U_i = \left.\frac{\partial U_i}{\partial \alpha_i}\right|_{\alpha_i = \alpha - \lambda t} \cdot \Delta \alpha_i. \tag{54}$$

Combining (54) with the fact that $\frac{\partial U_i}{\partial \alpha_i} > 0$ and using $|\Delta \alpha_i| = t$, by taking the absolute values of both parts of (54), we obtain:

$$\left(\min_{\alpha_i \in [\alpha - t, \alpha]} \frac{\partial U_i}{\partial \alpha_i} \right) \cdot t \leq |\Delta U_i| \leq \left(\max_{\alpha_i \in [\alpha - t, \alpha]} \frac{\partial U_i}{\partial \alpha_i} \right) \cdot t. \tag{55}$$

By Lemma 6, equation (31), we have for all $i \in \mathcal{N}$ and all $\alpha_i \in [\alpha - t, \alpha]$:

$$\min_{j=1,\ldots n} \{\alpha_j\} \leq \frac{\partial U_i^*}{\partial \alpha_i} \leq \max_{j=1,\ldots n} \{\alpha_j\}. \tag{56}$$

Combining (55) with (56) and the fact that $|\Delta \alpha_i| = t$, we obtain (53). Finally, since $\alpha - t \leq \min_{j=1,\ldots n} \{\alpha_j\}$ and $\alpha \geq \max_{j=1,\ldots n} \{\alpha_j\}$, we obtain (52). ▮

When $c \leq \alpha t - t^2$, the Lemma 7 implies that $c < |\Delta U_i^*|$ and therefore no profile in which at least one agent does not adopt is an equilibrium, because, as implied by (52), gains from adopting are always positive. This proves that the Adopting Equilibrium is the unique SPNE, whence asymmetric SPNE do not exist.

When $c > \alpha t$, the Lemma 7 implies $|\Delta U_i^*| < c$ and therefore no profile in which at least one agent adopts is a SPNE because, as implied by (52), gains from non-adopting are always positive. This proves that the Non-Adopting Equilibrium is the unique SPNE. Therefore, asymmetric SPNE do not exist in this case either.

This completes the proof. ▮

Proof of Proposition 8: Let $\mathcal{A} \subseteq \mathcal{N}$ be such that $\mathcal{A} \neq \varnothing, \mathcal{N}$. We now derive a necessary and sufficient condition for an asymmetric SPNE to exist where $\mathcal{A}$ is the set of adopters and $\mathcal{N} \setminus \mathcal{A}$ is the set of non-adopters.

Case (i): $i \in \mathcal{A}$. Let us first check that a unilateral deviation is not profitable for adopters. For all $i \in \mathcal{A}$, we have $\alpha_i = \alpha$ and $\Delta \alpha_i = -t$. Thus, we need to check that

$$2c + 2\Delta U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}) < 0 \tag{57}$$

where

$$
\begin{aligned}
& 2\Delta U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}) \\
= \; & -t\,(2\alpha - t) \\
& + \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) t \left[\begin{array}{c} \left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)(2\alpha - t) - 2\alpha \sum_{j\in\mathcal{A}\setminus\{i\}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} \\ -2\,(\alpha - t) \sum_{j\in\mathcal{N}\setminus\mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} \end{array} \right]
\end{aligned}
$$

Thus condition (57) can be written as:

$$
2c < t\,(2\alpha - t)
$$

$$
-\left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) t \left[\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)(2\alpha - t) - 2\alpha \sum_{j\in\mathcal{A}\setminus\{i\}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} - 2\,(\alpha - t) \sum_{j\in\mathcal{N}\setminus\mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\right].
$$

Because $\frac{1}{1+\theta}\widehat{\mathbf{M}}$ is a row-normalized matrix, we can restate this as

$$
2c < t\,(2\alpha - t) - \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) t \left[-t\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) + 2t \sum_{j\in\mathcal{N}\setminus\mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\right],
$$

or, equivalently, as

$$
2c < 2\alpha t - t^2 - \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) t^2 \left[\frac{\widehat{m}_{ii}}{1+\theta} + 2 \sum_{j\in\mathcal{N}\setminus\mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} - 1\right].
$$

Observe that

$$
\sum_{j\in\mathcal{N}\setminus\mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} = 1 - \sum_{j\in\mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} = 1 - \frac{\widehat{m}_{ii}}{1+\theta} - \sum_{j\in\mathcal{A}\setminus\{i\}}^{n} \frac{\widehat{m}_{ij}}{1+\theta}.
$$

Therefore, (57) can be further rewritten as:

$$
2c < 2\alpha t - t^2
$$

$$
-\left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) t^2 \left[\frac{\widehat{m}_{ii}}{1+\theta} + \sum_{j\in\mathcal{N}\setminus\mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} + \left(1 - \frac{\widehat{m}_{ii}}{1+\theta} - \sum_{j\in\mathcal{A}\setminus\{i\}}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\right) - 1\right]
$$

$$\Leftrightarrow 2c < 2\alpha t - t^2 - \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) t^2 \left(\sum_{j\in\mathcal{N}\setminus\mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} - \sum_{j\in\mathcal{A}\setminus\{i\}}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\right)$$

$$\Leftrightarrow 2c < 2\alpha t - t^2 + t^2 \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \left(\sum_{j\in\mathcal{A}\setminus\{i\}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} - \sum_{j\in\mathcal{N}\setminus\mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\right)$$

$$\Leftrightarrow c < \alpha t - \frac{t^2}{2}\left[1 - \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \sum_{j\in\mathcal{N}\setminus\{i\}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} I_{\mathcal{A}}(j)\right]$$

Define

$$\overline{c}_{\mathcal{A}}(\theta, \mathbf{g}) := \min_{i\in\mathcal{A}} \left\{\alpha t - \frac{t^2}{2}\left[1 - \left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \sum_{j\in\mathcal{N}\setminus\{i\}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} I_{\mathcal{A}}(j)\right]\right\},$$

or, equivalently,

$$\overline{c}_{\mathcal{A}}(\theta, \mathbf{g}) := \alpha t - \frac{t^2}{2}\left[1 - \left(\frac{1+\theta}{\theta}\right) \min_{i\in\mathcal{A}} \left\{\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \sum_{j\in\mathcal{N}\setminus\{i\}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} I_{\mathcal{A}}(j)\right\}\right].$$

Hence, individual deviations are unprofitable for each agent $i \in \mathcal{A}$ if and only if $c < \overline{c}_{\mathcal{A}}(\theta, \mathbf{g})$.

Case 2: $i \in \mathcal{N} \setminus \mathcal{A}$. Let us now check when an individual deviation is not profitable for $i \in \mathcal{N}\setminus\mathcal{A}$ (i.e for a non-adopter). In that case, we have $\alpha_i = \alpha - t$ and $\Delta\alpha_i = t$. For that, we need to check that

$$2\Delta U_i^*(\alpha_i, \boldsymbol{\alpha}_{-i}) < 2c, \tag{58}$$

or, equivalently,

$$2c > t\,(2\alpha - t)$$

$$-\left(\frac{1+\theta}{\theta}\right)\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) t \left[\begin{array}{c} \left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right)(2\alpha - t) \\ -2(\alpha - t)\sum_{j\in(\mathcal{N}\setminus\mathcal{A})\setminus\{i\}} \frac{\widehat{m}_{ij}}{(1+\theta)} - 2\alpha \sum_{j\in\mathcal{A}} \frac{\widehat{m}_{ij}}{(1+\theta)} \end{array}\right]$$

Because $\frac{1}{1+\theta}\widehat{\mathbf{M}}$ is a row-normalized matrix, we can restate (58) as

$$2c > 2\alpha t - t^2 - t^2\left(\frac{1+\theta}{\theta}\right)\left(1-\frac{\widehat{m}_{ii}}{1+\theta}\right)\left[\frac{\widehat{m}_{ii}}{1+\theta} + 2\sum_{j\in(\mathcal{N}\setminus\mathcal{A})\setminus\{i\}}\frac{\widehat{m}_{ij}}{(1+\theta)} - 1\right].$$

Since

$$\sum_{j\in(\mathcal{N}\setminus\mathcal{A})\setminus\{i\}}^{n}\frac{\widehat{m}_{ij}}{1+\theta} = 1 - \frac{\widehat{m}_{ii}}{1+\theta} - \sum_{j\in\mathcal{A}}^{n}\frac{\widehat{m}_{ij}}{1+\theta},$$

we get after simplifications:

$$c > \alpha t - \frac{t^2}{2}\left[1 - \left(\frac{1+\theta}{\theta}\right)\left(1-\frac{\widehat{m}_{ii}}{1+\theta}\right)\sum_{j\in\mathcal{N}\setminus\{i\}}^{n}\frac{\widehat{m}_{ij}}{1+\theta}I_{\mathcal{A}}(j)\right]$$

Define

$$\underline{c}_{\mathcal{A}}(\theta,\mathbf{g}) := \alpha t - \frac{t^2}{2}\left[1 - \left(\frac{1+\theta}{\theta}\right)\max_{i\in\mathcal{N}\setminus\mathcal{A}}\left\{\left(1-\frac{\widehat{m}_{ii}}{1+\theta}\right)\sum_{j\in\mathcal{N}\setminus\{i\}}^{n}\frac{\widehat{m}_{ij}}{1+\theta}I_{\mathcal{A}}(j)\right\}\right]$$

The necessary and sufficient condition for a non-trivial domain of c to exist in which the partition $\{\mathcal{A}, \mathcal{N}\setminus\mathcal{A}\}$ generates an asymmetric SPNE is given by:

$$\underline{c}_{\mathcal{A}}(\theta,\mathbf{g}) < \bar{c}_{\mathcal{A}}(\theta,\mathbf{g}),$$

which is readily verified to be equivalent to (24). In other words, for all $i \in \mathcal{A}$ and for all $k \in \mathcal{N}\diagdown\mathcal{A}$, we need:

$$\left(1-\frac{\widehat{m}_{ii}}{1+\theta}\right)\sum_{j\in\mathcal{N}\setminus\{i\}}^{n}\frac{\widehat{m}_{ij}}{1+\theta}I_{\mathcal{A}}(j) > \left(1-\frac{\widehat{m}_{kk}}{1+\theta}\right)\sum_{j\neq k}^{n}\frac{\widehat{m}_{kj}}{1+\theta}I_{\mathcal{A}}(j)$$

This completes the proof. ▮

Proof of Proposition 9:
Characterization: There is no asymmetric SPNE
We have

$$
\begin{aligned}
\widehat{\mathbf{M}} &= \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \left(\frac{\theta}{1+\theta} \right) \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix} \right)^{-1} \\
&= \left(\frac{1}{1+2\theta} \right) \begin{pmatrix} \frac{1}{2}\left(\theta^2+4\theta+2\right) & \theta\left(\theta+1\right) & \frac{1}{2}\theta^2 \\ \frac{1}{2}\theta\left(\theta+1\right) & \left(\theta+1\right)^2 & \frac{1}{2}\theta\left(\theta+1\right) \\ \frac{1}{2}\theta^2 & \theta\left(\theta+1\right) & \frac{1}{2}\left(\theta^2+4\theta+2\right) \end{pmatrix}
\end{aligned}
$$

Assume $\mathcal{A} = \{2\}$ and $\mathcal{N} \diagdown \mathcal{A} = \{1,3\}$ without loss of generality. Then, we have:

$$
\begin{aligned}
& \min_{i \in \{1,3\}} \left\{ \left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \left[\left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) - 2 \sum_{j \in \mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta} \right] \right\} \\
= & \left(1 - \frac{\left(\theta^2+4\theta+2\right)}{2\left(1+2\theta\right)\left(1+\theta\right)}\right) \left(\left(1 - \frac{\left(\theta^2+4\theta+2\right)}{2\left(1+2\theta\right)\left(1+\theta\right)}\right) - \frac{4\theta}{\left(1+2\theta\right)} \right) \\
= & -\frac{15\theta^4 + 28\theta^3 + 12\theta^2}{16\theta^4 + 48\theta^3 + 52\theta^2 + 24\theta + 4} < 0
\end{aligned}
$$

and

$$
\begin{aligned}
& \min_{i \in \{2\}} \left\{ \left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \left(1 - \frac{\widehat{m}_{ii}}{1+\theta} - 2 \sum_{j \in \mathcal{N} \diagdown \mathcal{A}}^{n} \frac{\widehat{m}_{ij}}{1+\theta}\right) \right\} \\
= & \left(1 - \frac{\left(1+\theta\right)}{\left(1+2\theta\right)}\right) \left(1 - \frac{1+\theta}{\left(1+2\theta\right)} - \frac{2\theta}{\left(1+2\theta\right)}\right) = -\frac{\theta^2}{\left(2\theta+1\right)^2} < 0
\end{aligned}
$$

Thus condition (24) is never satisfied.

Assume now that $\mathcal{A} = \{1,3\}$ and $\mathcal{N} \diagdown \mathcal{A} = \{2\}$. We have

$$
\min_{i \in \mathcal{A}} \left\{ \left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \sum_{j \neq i}^{n} \frac{\widehat{m}_{ij}}{1+\theta} I_{\mathcal{A}}(j) \right\} = \left(1 - \frac{\widehat{m}_{11}}{1+\theta}\right) \frac{\widehat{m}_{13} - \widehat{m}_{12}}{1+\theta} < 0
$$

$$
\max_{i \in \mathcal{N} \diagdown \mathcal{A}} \left\{ \left(1 - \frac{\widehat{m}_{ii}}{1+\theta}\right) \left(\sum_{j \neq i}^{n} \frac{\widehat{m}_{ij}}{1+\theta} I_{\mathcal{A}}(j) \right) \right\} = \left(1 - \frac{\widehat{m}_{22}}{1+\theta}\right) \sum_{j \neq 2}^{n} \frac{\widehat{m}_{2j}}{1+\theta} > 0.
$$

Assume, finally, that $\mathcal{A} = \{1,2\}$ and $\mathcal{N} \diagdown \mathcal{A} = \{3\}$. Then,

$$\begin{aligned}
&\min_{i\in\mathcal{A}}\left\{\left(1-\frac{\widehat{m}_{ii}}{1+\theta}\right)\sum_{j\neq i}^{n}\frac{\widehat{m}_{ij}}{1+\theta}I_{\mathcal{A}}(j)\right\}\\
=\;&\min\left\{\left(1-\frac{\widehat{m}_{11}}{1+\theta}\right)\frac{\widehat{m}_{12}-\widehat{m}_{13}}{1+\theta},\left(1-\frac{\widehat{m}_{22}}{1+\theta}\right)\frac{\widehat{m}_{21}-\widehat{m}_{23}}{1+\theta}\right\}=0,
\end{aligned}$$

$$\max_{i\in\mathcal{N}\diagdown\mathcal{A}}\left\{\left(1-\frac{\widehat{m}_{ii}}{1+\theta}\right)\left(\sum_{j\neq i}^{n}\frac{\widehat{m}_{ij}}{1+\theta}I_{\mathcal{A}}(j)\right)\right\}=\left(1-\frac{\widehat{m}_{33}}{1+\theta}\right)\frac{\widehat{m}_{31}+\widehat{m}_{32}}{1+\theta}>0.$$

In all these cases, (24) fails to hold. This means that there cannot be an asymmetric equilibrium in a star-shaped network with three players.

Characterization of symmetric equilibria

Set $\alpha=10$, $t=6$ and $\theta=0.5$. In that case, we have:

$$\widehat{\mathbf{M}}=\begin{pmatrix}1.0625 & 0.375 & 0.0625\\ 0.1875 & 1.125 & 0.1875\\ 0.0625 & 0.375 & 1.0625\end{pmatrix}$$

Then, $\alpha t=60$, $\alpha t-t^2=24$, $\underline{c}(0.5,\mathbf{g})=38.625$ and $\overline{c}(0.5,\mathbf{g})=45.375$. Using Proposition 6, we obtain the characterization result for the symmetric equilibria. ∎

Appendix B: Additional results

Proposition 10 *[Comparative statics of equilibrium efforts] The effect of* θ *on* x_i^* *is ambiguous and given by:*

$$\frac{\partial x_i^*}{\partial \theta} = -\frac{1}{(1+\theta)} \sum_j \widehat{m}_{ij} \left(x_j^* - \overline{x}_j^* \right)$$

while the effect of θ *on* $\overline{x}_i^*$ *is independent of* i *and given by:*

$$\frac{\partial \overline{x}_i^*}{\partial \theta} = -\frac{1}{(1+\theta)} \sum_j \left(\sum_l \widehat{m}_{lj} - \sum_r \sum_l \widehat{g}_{rl} \widehat{m}_{lj} \right) \alpha_j$$

Proof of Proposition 10: By implicitly differentiating (27), we obtain:

$$\begin{aligned} \frac{\partial \mathbf{x}^*}{\partial \theta} &= -\left[(1+\theta)\, \mathbf{I} - \theta \widehat{\mathbf{G}} \right]^{-1} \left(\mathbf{I} - \widehat{\mathbf{G}} \right) \mathbf{x}^* \\ &= -\frac{1}{(1+\theta)} \widehat{\mathbf{M}} \left(\mathbf{I} - \widehat{\mathbf{G}} \right) \mathbf{x}^* \\ &= -\frac{1}{(1+\theta)} \widehat{\mathbf{M}} \left(\mathbf{x}^* - \overline{\mathbf{x}}^* \right) \end{aligned}$$

which implies that

$$\frac{\partial x_i^*}{\partial \theta} = -\frac{1}{(1+\theta)} \sum_j \widehat{m}_{ij} \left(x_j^* - \overline{x}_j^* \right)$$

which sign is indeterminate.

Furthermore, by differentiating (28), we have:

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

This implies that

$$\begin{aligned} \frac{\partial \overline{x}_i^*}{\partial \theta} &= -\frac{1}{(1+\theta)} \sum_l \sum_j \widehat{g}_{il} \widehat{m}_{lj} \left(x_j^* - \overline{x}_j^* \right) \\ &= -\sum_j \left(x_j^* - \overline{x}_j^* \right) \frac{1}{(1+\theta)} \sum_l \widehat{g}_{il} \widehat{m}_{lj} \end{aligned}$$

Since, by Lemma 5, $\frac{1}{1+\theta}\widehat{\mathbf{G}}\widehat{\mathbf{M}}$ is row-normalized, then $\frac{1}{(1+\theta)}\sum_l \widehat{g}_{il}\widehat{m}_{lj} = 1$. Thus,

$$\frac{\partial \overline{x}_i^*}{\partial \theta} = -\sum_j \left(x_j^* - \overline{x}_j^*\right) = \sum_j \overline{x}_j^* - \sum_j x_j^* \tag{59}$$

This derivative does not depend on i and thus a change in θ leads to a change in all social norms in the same direction, i.e.

$$\frac{\partial \overline{x}_i^*}{\partial \theta} = \sum_j \overline{x}_j^* - \sum_j x_j^*, \; \forall i$$

That is,

$$\frac{\partial \overline{x}_i^*}{\partial \theta} = \mathbf{1}^T\overline{\mathbf{x}}^* - \mathbf{1}^T\mathbf{x}^*$$

where $\mathbf{1}$ is the $n \times 1$ vector of 1s. This is equivalent to:

$$\begin{aligned}
\frac{\partial \overline{x}_i^*}{\partial \theta} &= \mathbf{1}^T\widehat{\mathbf{G}}\mathbf{x}^* - \mathbf{1}^T\mathbf{x}^* \\
&= \mathbf{1}^T\widehat{\mathbf{G}}\frac{1}{1+\theta}\widehat{\mathbf{M}}\boldsymbol{\alpha} - \mathbf{1}^T\frac{1}{1+\theta}\widehat{\mathbf{M}}\boldsymbol{\alpha} \\
&= \frac{1}{(1+\theta)}\mathbf{1}^T\left(\widehat{\mathbf{G}} - \mathbf{I}\right)\widehat{\mathbf{M}}\boldsymbol{\alpha} \\
&= -\frac{1}{(1+\theta)}\mathbf{1}^T\left(\mathbf{I}-\widehat{\mathbf{G}}\right)\widehat{\mathbf{M}}\boldsymbol{\alpha} \\
&= -\frac{1}{(1+\theta)}\sum_j\left(\sum_l \widehat{m}_{lj} - \sum_r\sum_l \widehat{g}_{rl}\widehat{m}_{lj}\right)\alpha_j
\end{aligned}$$

This implies that

$$\frac{\partial \overline{x}_i^*}{\partial \theta} \gtreqless 0 \Leftrightarrow \sum_j\left(\sum_l \widehat{m}_{lj}\right)\alpha_j \lesseqgtr \sum_i\left(\sum_r\sum_l \widehat{g}_{rl}\widehat{m}_{lj}\right)\alpha_j$$

This proves the result. ∎

Observe that $\frac{\partial \overline{x}_i^*}{\partial \theta}$ does not depend on i and thus a change in θ leads to a change in all social norms in the same direction. Moreover, the sign of $\sum_l \widehat{m}_{lj} - \sum_j\sum_l \widehat{g}_{il}\widehat{m}_{lj}$ determines the sign of $\frac{\partial \overline{x}_i^*}{\partial \theta}$.

Proposition 11 *[Comparative statics of equilibrium utility and welfare] There exists a $\widetilde{\theta} > 0$, such that, when $\theta < \widetilde{\theta}$, a marginal increase in θ reduces the equilibrium utilities of all individuals, that is, for all $i = 1, ..., n$,*

$$\frac{\partial U_i(x_i^*, \mathbf{x}_{-i}^*, g)}{\partial \theta} < 0$$

Proof of Proposition 11: Let us calculate $\frac{\partial U_i^*}{\partial \theta}$. Using the envelope theorem, we have:

$$\begin{aligned} \frac{\partial U_i^*}{\partial \theta} &= -\frac{1}{2}(x_i^* - \overline{x}_i^*)^2 + \theta \left(x_i^* - \overline{x}_i^*\right) \frac{\partial \overline{x}_i^*}{\partial \theta} \\ &= -\left(x_i^* - \overline{x}_i^*\right) \left[\frac{1}{2}(x_i^* - \overline{x}_i^*) - \theta \frac{\partial \overline{x}_i^*}{\partial \theta}\right] \end{aligned}$$

Using (59), we have

$$\frac{\partial U_i^*}{\partial \theta} = -\left(x_i^* - \overline{x}_i^*\right) \left[\frac{1}{2}(x_i^* - \overline{x}_i^*) - \theta \left(\sum_j \overline{x}_j^* - \sum_j x_j^*\right)\right]$$

That is,

$$\frac{\partial U_i^*}{\partial \theta} = -\frac{1}{2}\left(x_i^* - \overline{x}_i^*\right)^2 + \theta \left(x_i^* - \overline{x}_i^*\right) \sum_j \left(\overline{x}_j^* - x_j^*\right) \tag{60}$$

This implies that

$$\frac{\partial U_i^*}{\partial \theta} \mid_{\theta=0} = -\frac{1}{2}\left(x_i^* - \overline{x}_i^*\right)^2 < 0$$

Therefore, by contuinity, $\exists \widetilde{\theta} > 0$ such that $\frac{\partial U_i^*}{\partial \theta} < 0$. ∎