école d'économie d'aix-marseille
aix-marseille school of economics

Working Papers / Documents de travail

The Dark Side of Peers: Demotivation through Social Comparison in Networks

Frédéric Deroïan
Mohamed Belhaj

WP 2025 - Nr 11

The Dark Side of Peers: Demotivation through Social Comparison in Networks*

Frédéric Deroïan[†,1] and Mohamed Belhaj[2]

[1]Aix-Marseille Univ., CNRS, AMSE, Marseille, France. Email: frederic.deroian@univ-amu.fr

[2]Aix-Marseille Univ., CNRS, Centrale Méditerranée, AMSE, Marseille, France. Email: mohamed.belhaj@centrale-med.fr

May 15, 2025

Abstract

This paper introduces demotivation in the context of social comparison in networks. Social comparison is modeled as a status effect rewarding or penalizing agents according to their relative performance with respect to local peers. A demotivated agent faces both a reduced marginal return to effort

*The authors particularly thank Yann Bramoullé, Marco Fongoni, and Alastair Langtry for insightful discussions. We acknowledge financial support from the French government under the "France 2030" investment plan managed by the French National Research Agency Grant ANR-17-EURE-0020, and by the Excellence Initiative of Aix-Marseille University - A*MIDEX.

†Corresponding author.

and a psychological cost. In the absence of demotivation, social comparison leads to higher effort levels but reduces equilibrium welfare. Introducing demotivation leads to two main findings. First, it generates a network game of strategic substitutes. Second, despite the individual psychological costs incurred by demotivated agents, it can enhance overall welfare—by alleviating social pressure to exert effort and by generating positive externalities for peers.

Keywords. Social Comparison; Demotivation; Networks; Strategic Substitutes, Equilibrium Welfare.

JEL Classification. C72; D83; D85

1 Introduction

Comparing oneself with others is a core aspect of human experience. For instance, social comparison is prevalent for people at school, at work, and in fact in nearly all areas of social life, and across sport and culture. While upward comparison can motivate individuals to increase their effort, it can also produce the opposite effect when the performance gap appears too large or improvement seems out of reach. In such cases, individuals may become demotivated and reduce the importance they attach to social comparison. This idea is rooted in Festinger's seminal theory of social comparison, which posits that the propensity to compare oneself to others declines as the perceived gap in ability or opinion increases:

"The tendency to compare oneself with some other specific person decreases as the difference between his opinion or ability and one's own increases." (Festinger (1954), p. 120)

Applied to the context of status-based incentives, this implies that when individuals suffer from large unfavorable comparisons, they may reduce the weight they place on social status. In environments where status is a key motivator of effort, this psychological response may in turn reduce the perceived return to exerting effort.[1] Moreover, demotivation can entail significant psychological costs,

[1] In the field of economics, Bénabou and Tirole (2003) show how external social comparison can undermine intrinsic motivation and reduce individuals' effort and performance. Lazear and Rosen (1981) examine how individuals' motivation to exert effort depends on their relative rank in a competitive environment. Similarly, Murphy and Weinhardt (2020) and Dening et al. (2023) establish that an individual's ordinal position within a group significantly impacts later objective outcomes, even when controlling for cardinal achievement. Such evidence strongly suggests a causal link between negative social comparison —such as being ranked low— and a decline in motivation and effort. The experimental literature on worker performance in organizations has

such as diminished self-esteem or a weakened sense of self-worth.[2] By significantly lowering the returns to effort and inducing psychological distress, demotivation can have far-reaching consequences for economic behavior and outcomes.

In general, people tend to compare themselves to the people they interact with most frequently, i.e. social comparisons are localized to close social contacts; The collection of those social contacts constitutes a social network. The structure of the social network can play a crucial role in the emergence of demotivation. An individual's position within the network, as well as the social comparisons facilitated by their connections, can significantly influence their perception of effort and success, potentially leading to demotivation. On the other hand, when someone experiences demotivation, it can affect not only their own social status but also the status of others within the network. By improving the social status of immediate peers, demotivation can create indirect effects that ripple through the network, influencing the motivation of more distant individuals. In this way, social status and motivation form a dynamic feedback loop, where both local comparisons and net-

---

also stressed how payments schemes can lead to diminished returns to effort linked to demotivation under peer pressure. See for instance Eriksson et al. (2009), or Bellemare et al. (2010).

[2]The literature in social psychology stresses that upward comparison can undermine self-esteem. For instance, Tesser et al. (1988) provide empirical evidence that when another outperforms the self on a task high in relevance to the self, the closer the other the greater the threat to self-evaluation. Wheeler and Miyake (1992) find that students reported feeling depressed and discouraged when they compared themselves with superior people. Exploring the impact of upward social comparison on self-evaluations, Collins (1996) underlines that 'expecting to be different from an upward target should lead to a contrast effect, feelings of inferiority, and more negative self-appraisals'. Lockwood and Kunda (2003) and Rogers and Feller (2016) show that exposure to exemplary peer performances can undermine motivation and success by causing people to perceive that they cannot attain their peers' high levels of performance.

work structure interact. Understanding how demotivation emerges in the network is therefore a complex issue.

This paper incorporates demotivation into a simple model of social comparison in networks. In this work, agents exert costly effort and derive utility not only from their own effort but also from their relative standing within their social network. To capture social comparison, we introduce in the utility function a status component which is a linear function of the difference between own effort and neighbors' average. We introduce demotivation[3] by incorporating two additional features. First, demotivation induces a lower return to effort. We model this variation of the return to effort through a kink in the social status function: when own effort is sufficiently far below local peers' effort, the marginal return to effort drops. Second, demotivation generates psychological costs that we model as a utility loss at the kink, aimed to reflect undermined self-esteem or self-worth. Our goal is to understand how the network structure influences the emergence of demotivation, and their further consequences on economic outcomes like effort and social welfare.

In the benchmark case with social comparison but without demotivation (that we call the no-demotivation scenario), there is a unique equilibrium, in which all agents exert a high effort level. This leads to lower welfare as compared to the case in which there is no social comparison at all (that we call the no-status scenario). This is the standard arm race result known in the literature on conspicuous goods. Introducing demotivation, the best-response of an agent to local peers' effort consists in choosing a high effort level (and be motivated) when others exert a

[3] In the perspective of the present paper, demotivation does not totally annihilate incentives (which may otherwise confine to even sharper consequences such that renouncement).

low effort level, and choosing a low effort level (and be demotivated) when others exert a high effort level. That is, an agent is demotivated when neighbors' average effort exceeds a critical value. Then, an agent is demotivated at equilibrium when the proportion of motivated neighbors exceeds a threshold, that depends on the severity of the kink and the psychological costs of demotivation, but not on the network structure. This equilibrium characterization corresponds to *a network game of strategic substitutes with binary effort choice.* This sharply contrasts with the complementarity-driven incentives commonly found in the status games studied in the literature. In our model, the emergence of strategic substitutes is a direct consequence of the kink in the status function. Absent this discontinuity, the presence of psychological costs alone does not generate such strategic behavior.

We identify a potential function ensuring equilibrium existence.[4] The network structure matters in shaping equilibria in many respects. In this world of strategic substitutes, equilibrium multiplicity can be huge under complex network structures, which raises the issue of finding all equilibria. We show that *the set of equilibria is in general a NP-complete problem*, by establishing a correspondence with the so-called MaxCut problem in the simplest version of the game. Moreover, the impact of network on demotivation can be very strong: in certain networks, an agent may be demotivated (or motivated) across all equilibria due to their sole position.

We show that, in this game, no equilibrium Pareto-dominates another, reinforcing the non-trivial welfare implications induced by demotivation. We then

[4]The formal structure of equilibria, as well as the existence of a potential function, echoes the literature on anti-coordination games. E.g., Blume (1993), Young (1998), or more recently Bramoullé (2007).

consider a utilitarian welfare approach. Contrary to the conventional view that status concerns drive excessive effort and reduce welfare, the introduction of demotivation can alter these conclusions. Indeed, despite psychological costs, the presence of demotivated agents contribute to reduce the social pressure on effort and to improve the social status of their neighbors.[5]

We undertake several comparative statics. Our main messages can be given by comparing the welfare of an equilibrium to respectively the no-demotivation scenario and the no-status scenario. Addressing comparison with the no-demotivation scenario, the reduction of social pressure always benefits motivated agents through improved social status, and this can even benefit demotivated agents through reduced effort. In total, the psychological costs borne by demotivated agents are decisive. Under sufficiently low costs, for any network, the welfare of any equilibrium is higher than the equilibrium welfare in no-demotivation scenario. More generally, the threshold level of psychological costs that reverses this welfare comparison is equilibrium-specific.

When demotivation leads to a sufficiently large drop in the return to effort while psychological costs remain low, equilibrium welfare can even exceed that of the no-status benchmark—reversing standard predictions regarding the welfare effects of status effects. Again, the threshold is equilibrium-dependent (and thus network-dependent). Again, the presence of high psychological costs qualifies that conclusion.

Interestingly, we also identify a countervailing effect of psychological costs

[5] This aspect echoes the so-called scapegoat mechanism known in social sciences, as the influential works of Allport in social psychology, or René Girard for instance. The present paper could potentially bring an economic perspective to the debate.

on equilibrium welfare. Specifically, we show that increasing these costs can raise the welfare of the second-best equilibrium—i.e., the equilibrium that yields the highest welfare. The intuition is that higher psychological costs strengthen the incentives to remain motivated, potentially leading to a reconfiguration of the set of motivated agents that enhances overall welfare.

Finally, we extend our model in several directions. We introduce heterogeneous agents' characteristics, we incorporate more general utility functions, and we introduce local synergies in the network beyond status effects. Across these extensions, the emergence of strategic substitutes remains a robust outcome.

Relationship to the literature. The literature on status goods has a long tradition in economics.[6] Our paper inserts more specifically in the literature on status games played on networks. Ghiglino and Goyal (2010) introduce a networked positional good leading the Bonacich centrality to predict the consumption levels of the positional good. Immorlica et al. (2017) examine status concern, that is, a situation in which agents care about those neighbors with higher action only. In their model, there is only a disutility of being below others, whereas we also model utility gain from being above. Langtry (2023) assumes that agents form a social reference point based on the (weighted) sum of their neighbors consumption, and examines network formation.[7] Bramoullé and Ghiglino (2024) incorporate loss aversion into the framework of Ghiglino and Goyal (2010) and find a continuum of equilibria in which all consumers consume the same quantity of the status good on the network when agents' incomes are sufficiently close to each other. With respect to that literature, our paper contributes by showing that the presence of

[6] Within economics, see for instance Veblen (1899), Duesenberry (1949), Frank (1985), Clark and Oswald (1996), Hopkins and Kornienko (2004), Luttmer (2005), Frank (2005).

[7] Staab (2024) examines the formation of social groups under status concern.

demotivation in social comparison induces a game of strategic substitutes, where the whole literature obtains that status generates games of strategic complements. To our knowledge, our paper is the first to obtain strategic substitutes in status games played on networks.[8]

Our paper is also related to the literature modeling discouragement of workers in organizations. Gil and Prowse (2012) structurally estimate a model of disappointment aversion in a two-agent real-effort tournament, where only the winner receives a prize. Modeling disappointment-aversion through choice-acclimating reference point, they find that effort can be strategic substitutes (as an agent may reduce effort following an increase of the other agent when their chances of winning are low), which is interpreted as a discouragement effect. Our paper complements these findings by proposing a different source of discouragement, stemming from unfavorable social comparison. In addition, the tractability of our model enables to undertake a network analysis.

The paper is organized as follows. The networked game of social comparison is presented in Section 2. Section 3 studies equilibria of the game, and Section 4 analyzes the welfare properties of these equilibria. Section 5 examines the robustness of the emergence of strategic substitutes under several extensions. Section 6 concludes. All proofs are relegated in Appendix A.

[8] Our equilibria with binary substitutes echo the literature on anti-coordination games played on networks. In particular, Bramoullé (2007) examines a binary anti-coordination game played on a fixed network. Anti-coordination also arise in congestion games (Rosenthal (2017)), or for instance in fashion games (Cao et al. (2013)). Our model can be seen as a providing a possible micro-economic foundations to anti-coordination.

2 Model

Let $\mathcal{N} = \{1, 2, \cdots, n\}$ be a finite set of agents organized in a network of social contacts $\mathbf{G} = (g_{ij})_{(i,j)\in\mathcal{N}^2}$, with $g_{ij} \in \{0, 1\}$ for all i, j, $g_{ii} = 0$ by convention, and $\mathbf{G}^T = \mathbf{G}$; the network is therefore binary and undirected. When $g_{ij} = 1$, agents i and j are called neighbors. Let $\mathbf{1}$ represent the n-dimensional vector of ones, let $\mathbf{d} = (d_i)_{i\in\mathcal{N}} = \mathbf{G1}$ be the profile of degrees in network $\mathbf{G}$. To avoid trivialities, we assume that no agent is isolated, so that $\mathbf{d} \geq \mathbf{1}$. Let $\tilde{\mathbf{G}} = (\tilde{g}_{ij})$, with $\tilde{g}_{ij} = \frac{g_{ij}}{d_i}$ be the normalized adjacency matrix in which all entries are divided by agent's degree. Let $x_i \in \mathbb{R}^+$ represent agent i's effort level, and $\mathbf{x} = (x_i)_{i\in\mathcal{N}}$ a profile of effort; let $\overline{x}_i = \sum_{j\in\mathcal{N}} \tilde{g}_{ij} x_j$ the average effort level of agent i's social contacts.

We specify the following utility function for an agent i:

$$u_i(x_i, x_{-i}) = ax_i - \frac{1}{2}x_i^2 + \underbrace{A(x_i - \overline{x}_i)}_{\text{status function}} \tag{1}$$

where parameter a represents agent i's private return of effort (See Section 5 for heterogeneous private returns and for more general utilities). Agent i's utility is separable in a private returns to costly effort and a status effect reflecting the utility of social comparison. The status effect is a function of the difference between own effort and average neighbors' effort. Importantly, status can be positive or negative, depending on whether agent's effort is above or below peers' effort.[9]

We now describe the status function, which is parameterized by four parame-

[9] Immorlica et al. (2017) use same utility specification, but focus on a very different status function, in which agents are exclusively (negatively) impacted by those neighbors whose effort is larger than theirs. Their game can be viewed as a game of loss aversion, of which it shares the strategic complementarities of agents' actions.

ters. Let $\gamma_H, \gamma_L, \beta, V$ be four real numbers such that $0 \leq \gamma_L \leq \gamma_H$, $0 < \beta \leq 1$, and $V \geq 0$. We consider a stylized piecewise-linear status function $A(.)$ given by

$$\begin{cases} A(x_i - \overline{x}_i) = \gamma_H(x_i - \overline{x}_i) & \text{if} \quad x_i \geq \beta\overline{x}_i \\ A(x_i - \overline{x}_i) = \gamma_L(x_i - \overline{x}_i) - (1-\beta)(\gamma_H - \gamma_L)\overline{x}_i - V & \text{if} \quad x_i < \beta\overline{x}_i \end{cases}$$

Figure 2 illustrates the status function, which has several features. First, an agent

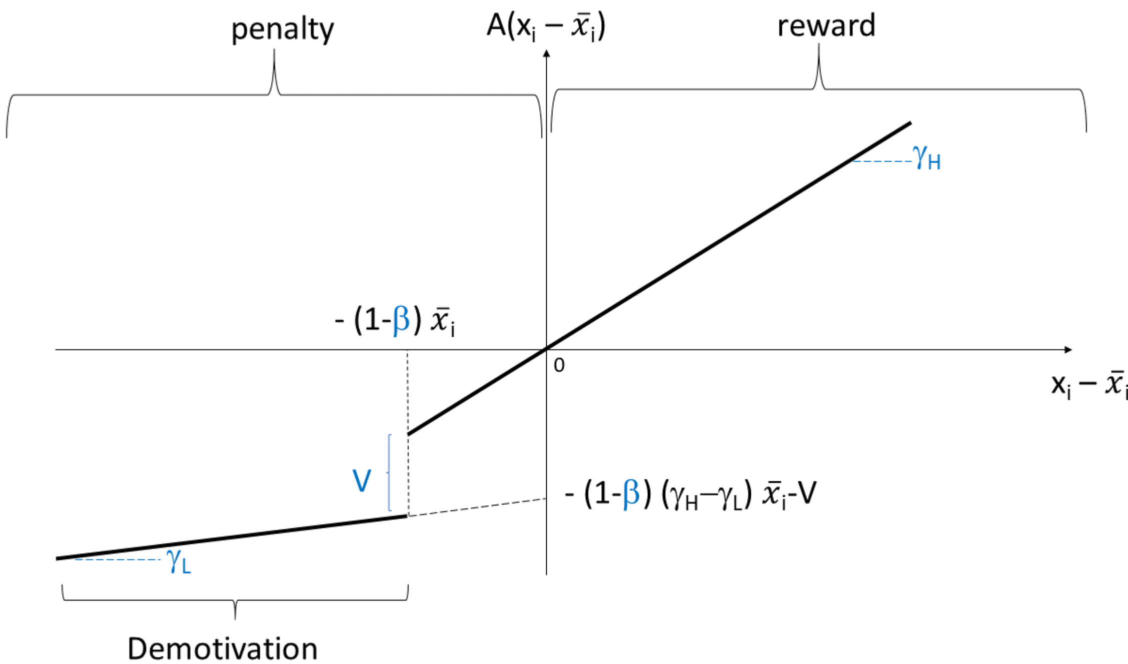


Figure 1: The function A, with demotivation effect, for $\beta < 1$.

suffers a utility loss when effort is below the average of neighbors' effort and experiences a utility gain when effort is higher. Second, the status function incorporates demotivation through the following additional features. First, above a percentage β of neighbors' effort, the marginal return of effort is equal to γ_H, while below that percentage, the return to effort is lowered to γ_L as a result of demotivation. Furthermore, demotivation induces a utility loss V at the kink, aimed to capture possible psychological costs, like for instance lowered self-esteem, self-worth.

To summarize, in this model demotivation means both a smaller focus on social comparision, and a psychological cost.[10] It will be useful to define two benchmark cases. We call the situation where $\gamma_L = \gamma_H = V = 0$ the no-status scenario, and the situation where $\gamma_L = \gamma_H > 0, V = 0$ the no-demotivation scenario.

Throughout the paper, it will be useful to introduce the following notation. For a given effort profile $\mathbf{x}$, we define $\mathbf{e}(\mathbf{x}) \in \{0,1\}^n$, such that $e_i = 1 \Leftrightarrow x_i \geq \beta \overline{x}_i$. Hence, an agent i such that $e_i = 1$ is said *motivated*, and an agent i such that $e_i = 0$ is said *demotivated*. Given profile $\mathbf{e}$, let $d_i^1(\mathbf{e}) = [\mathbf{G}\mathbf{e}]_i$ represent the number of agent i's motivated neighbors, and $d_i^0(\mathbf{e}) = [\mathbf{G}(\mathbf{1} - \mathbf{e})]_i$ represent the number of agent i's demotivated neighbors. Define the index $\rho_i^1(\mathbf{e}) = \frac{d_i^1(\mathbf{e})}{d_i}$ $\big($resp. $\rho_i^0(\mathbf{e}) = \frac{d_i^0(\mathbf{e})}{d_i}\big)$, the proportion of agent i's motivated (resp. demotivated) neighbors.

3 Demotivation brings anti-coordination

In this section, we analyze the equilibria of the game. We establish our main result for general parameters of the status function, and then we put emphasis on three specific cases: the simplest version $\beta = 1, V = 0$, then the case $\beta < 1, V = 0$, and finally the case $\beta = 1, V > 0$. The two latter cases allow to study

[10] We could also consider demotivation impacting private return or effort cost. However, without introducing a kink in the status function, it is readily checked that there is no equilibrium with demotivated agent.

the impact of parameters β and V separately. The main insight of this section is that incorporating demotivation into social comparison generates a binary network game of strategic substitutes sort between motivated versus demotivated agents. This message is robust to several generalizations presented thereafter (see Section 5).

3.1 A main result

We analyze the best-responses of the game, and then we characterize equilibria.

Best-responses. This model generates simple best-responses.

Proposition 1 *Let* $\varphi = \frac{1}{\beta}\left(a + \frac{\gamma_H+\gamma_L}{2} + \frac{V}{\gamma_H-\gamma_L}\right)$. *Agent* i*'s best-response to* $\overline{x}_i$ *is given by:*

$$\begin{cases} x_i^{BR}(\overline{x}_i) = a + \gamma_H & \text{if} \quad \overline{x}_i \leq \varphi \\ x_i^{BR}(\overline{x}_i) = a + \gamma_L & \text{if} \quad \overline{x}_i \geq \varphi \end{cases}$$

Figure 2 illustrates the shape of best-responses, which rests on the tradeoff between effort cost and utility gain on status. Motivated agents have to exert a high effort level to gain status, but when the average of neighbors' effort is too high, the reward in utility in terms of status the agent is low, meaning that the agent is better off reducing effort (thus effort cost) at the expense of a loss in status. Importantly, the best-response play of a motivated (resp. demotivated) agent is always strictly greater (resp. strictly lower) than average of neighbors' effort. When $\overline{x}_i = \varphi$, the agent has two best-responses, as playing either motivated or demotivated generate same payoff.

Equilibria. As a preliminary remark, we observe that, in the absence of demotivation, i.e. for $\gamma_L = \gamma_H$ and $V = 0$, there is a unique equilibrium in which there

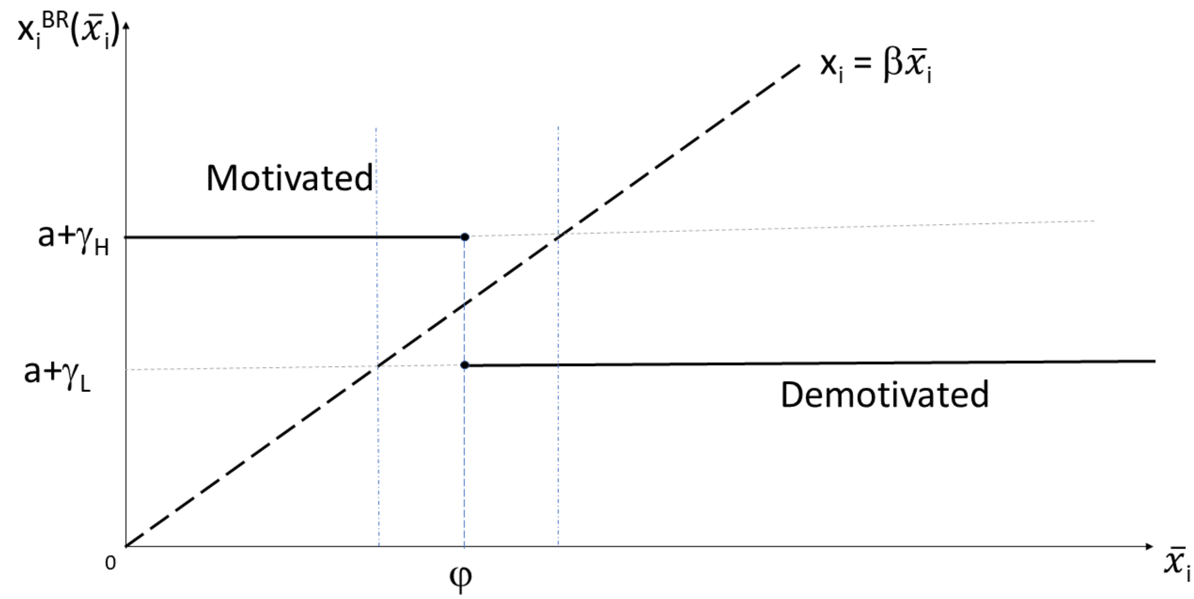


Figure 2: The best-response of an agent i under status effect with demotivation.

is no strategic interaction in decisions. Agent i exerts effort $x_i^* = a + \gamma_H$, and reaches an equilibrium utility $u_i^* = \frac{a^2 - \gamma_H^2}{2}$. Indeed, the search for status pushes agents to increase effort, and since all efforts are identical at equilibrium, there is no gain in status for each agent. This generates a clear-cut message about the impact of status effect on equilibria in the absence of demotivation:

Proposition 2 *When $\gamma_L = \gamma_H$ and $V = 0$, there is a unique equilibrium, in which $x_i^* = a + \gamma_H$ for all i. At equilibrium, effort is increasing in γ_H, and all individual utilities are decreasing in γ_H.*

An immediate implication of Proposition 2 is that, in the absence of demotivation, status effect entails higher effort and decreased utilities for all with respect to the no-status scenario.

We turn now to demotivation, i.e. $0 \leq \gamma_L < \gamma_H$ and $V \geq 0$. A key aspect is that individuals can decide of being in either of two states: motivated or demotivated.[11] At any equilibrium $\mathbf{x}^*$, not only vector $\mathbf{e}^*(\mathbf{x}^*)$ keeps track of the status of agents in terms of demotivation, but also it fully reveals the (only) two equilibrium effort levels $x_i^* = a + \gamma_L + (\gamma_H - \gamma_L)e_i^*$. For notational convenience, we shall write equilibrium $\mathbf{e}^*$ and omit the reference to $\mathbf{x}^*$. We define

$$\kappa(a, \gamma_L, \gamma_H, \beta, V) = \frac{2(1-\beta)a + \gamma_H + (1-2\beta)\gamma_L}{2\beta(\gamma_H - \gamma_L)} + \frac{V}{\beta(\gamma_H - \gamma_L)^2} \qquad (2)$$

For convenience, we omit reference to the parameters in what follows and speak about threshold κ. We observe that κ is increasing in a, V, decreasing in β, and $\kappa \geq \frac{1}{2}$ for all parameters values (that minimal bound is attained when $\beta = 1, V = 0$). The kink induced by demotivation in the status function brings strategic interaction. In particular, an equilibrium $\mathbf{e}^*$ satisfies the following first-order conditions (as shown in the proof of Theorem 1 thereafter):

$$\begin{aligned} e_i^* = 1 \quad &\Rightarrow \quad d_i^1(\mathbf{e}^*) \leq \kappa d_i \\ e_i^* = 0 \quad &\Rightarrow \quad d_i^0(\mathbf{e}^*) \leq (1-\kappa) d_i \end{aligned}$$

An agent plays motivated if the proportion of motivated neighbors is less than κ (which implies high reward in status), and similarly an agent plays demotivated if the proportion of demotivated neighbors is less than $1-\kappa$ (which implies high loss in status). Thus agents anti-coordinate with those neighbors of same cate-

[11] Referring to demotivation as a 'decision' is merely a notational convenience to describe optimal actions in the game. It does not necessarily imply that individuals consciously choose to be demotivated in real-life settings.

gory, motivated and demotivated, and that two-group partition is endogenous at equilibrium.[12]

Actually, equilibria are solutions to a maximization problem with concave objective function

$$P(\mathbf{e}) = (\kappa \mathbf{1} - \frac{1}{2}\mathbf{e})^T \mathbf{G}\mathbf{e} \tag{3}$$

This function is called a potential of the game[13] since $\mathbf{G}^T = \mathbf{G}$. Indeed, setting $\mathbf{e}_{-i} = (e_j)_{j \neq i}$,

$$P(1, \mathbf{e}_{-i}) - P(0, \mathbf{e}_{-i}) = \kappa d_i - d_i^1(\mathbf{e})$$

meaning that, when agent i becomes motivated, this improves the potential function whenever the first-order conditions of the game hold. The potential function guarantees equilibrium existence, and we obtain:

Theorem 1 *Let the status function be such that $0 \leq \gamma_L < \gamma_H, \beta \leq 1, 0 \leq V$. There is always an equilibrium. Agents play a binary network game of strategic substitutes. A profile $\mathbf{e}^*$ is a Nash equilibrium if and only if*

$$\begin{aligned} e_i = 1 \quad &\Rightarrow \quad \rho_i^1(\mathbf{e}) \leq \kappa \\ e_i = 0 \quad &\Rightarrow \quad \rho_i^0(\mathbf{e}) \leq 1 - \kappa \end{aligned}$$

Theorem 1 gives a powerful message about existence. Due to the potential function, there exists a Nash equilibrium on any network and for any parameter values.

[12] In graph theory, a k-dependent set is a subset of vertices such that no vertex in the subset is adjacent to more than k vertices of the subset. f-dependent sets generalize k-dependent sets to heterogeneous thresholds. Hence, the set of equilibria is an f-dependent set with heterogeneous thresholds, where, for vertex v_i, the threshold $f(v_i) = \kappa d_i$ (see Diks et al. (1994)).

[13] See Monderer and Shapley (1996) or Voorneveld (2000).

From the shape of the potential, and since the support of actions is compact, a maximum of the potential exists and is then a Nash equilibrium. Theorem 1 also provides a powerful characterization of Nash equilibria, expressing that agents are demotivated as soon as the share of motivated neighbors exceeds κ. Nash stability thus boils down to a simple graph-related criterion.[14]

Equilibria have the following general property. *An equilibrium* $\mathbf{e}^*$*, and a distinct configuration* $\mathbf{e}^{'}$*. If either* $\mathbf{e}^* \leq \mathbf{e}^{'}$ *or* $\mathbf{e}^* \geq \mathbf{e}^{'}$*, then* $\mathbf{e}^{'}$ *is not an equilibrium.*[15] That two equilibria are not nested implies a clear distinction between the groups of demotivated agents.

We explore now some useful polar cases.

Polar case 1: $\beta = 1, V = 0$**.** In this situation, demotivation occurs as soon as effort is below neighbors' effort, and a demotivated agent does not incur any psychological cost. Figure 3 illustrates the status function with $\beta = 1, V = 0$. From equation (2), we deduce $\kappa = \frac{1}{2}$, and thus, by Theorem 1, a configuration $\mathbf{e}^*$ is a Nash equilibrium if and only if

$$e_i = 1 \Rightarrow \rho_i^1(\mathbf{e}) \leq \frac{1}{2}$$
$$e_i = 0 \Rightarrow \rho_i^0(\mathbf{e}) \leq \frac{1}{2}$$

Note that the set of equilibria does not depend on the slopes γ_L, γ_H of the piecewise-linear status function.

[14] The characterization given in Theorem 1 is formally equivalent to Bramoullé (2007), Proposition 1. However, in our model with continuous actions, effort selection is binary only at equilibrium; Furthermore, Theorem 1 relates the anti-coordination threshold κ to the primitives of our model of status effect with demotivation.

[15] Consider otherwise two distinct and nested equilibria $\mathbf{e}^* \leq \mathbf{e}^{*\prime}$. Then, there is an agent i such that $e_i^* = 0$ and $e_i^{*\prime} = 1$. That is, exploiting equilibrium conditions for agent i, $d_i^1(\mathbf{e}^{*\prime}) \leq \kappa d_i < d_i^1(\mathbf{e}^*)$. But since $\mathbf{e}^* \leq \mathbf{e}'^*$, $\mathbf{G}\mathbf{e}^* \leq \mathbf{G}\mathbf{e}^{*\prime}$, which contradicts that $d_i^1(\mathbf{e}^{*\prime}) < d_i^1(\mathbf{e}^*)$.

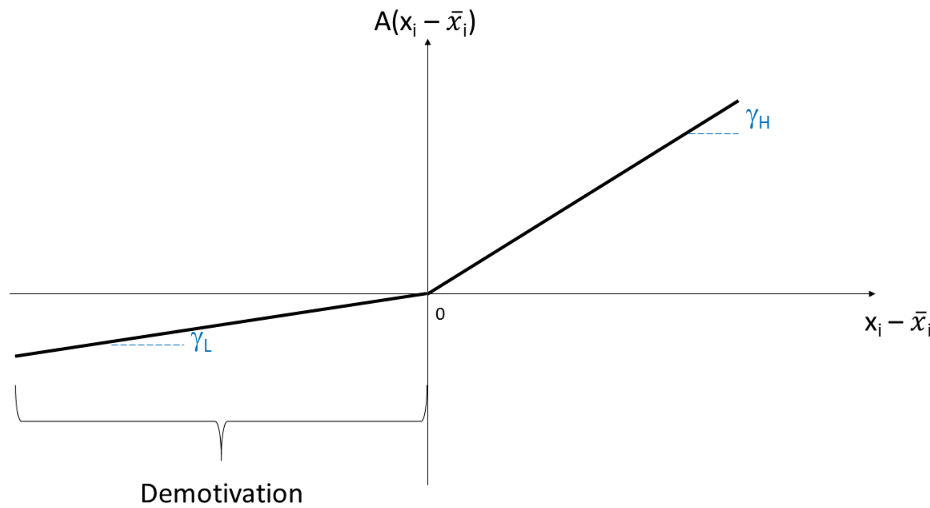


Figure 3: The function A for $\beta = 1$ and $V = 0$.

We give now general insights about how networks affect equilibrium characterization. It is important to stress that demotivation can emerge in regular networks. For instance, for the pair network with two agents, an equilibrium necessarily contains a single demotivated agent: the status effect demotivates agents when the other agent is motivated, whose status effect is enhanced by the demotivation of the other agent. On the complete network of even size, there is unique equilibrium in which the society is shared between two groups of equal size; with an odd number of agents, multiplicity emerges without further refinement on stability solution.[16] There are two equilibria in complete bipartite networks. This

[16]For instance, we might consider a slight refinement to Nash equilibrium by imposing that, in case of indifference, an agent always plays in the motivated region; This could be rationalized through the introduction of a small cost to choosing demotivation. Under that refinement, there is always a unique equilibrium in the complete network.

is because agents on the same side have the same neighborhood. For general network structures, there can be a high number of Nash equilibria in this game. Figure 4 shows equilibria on various network structures.

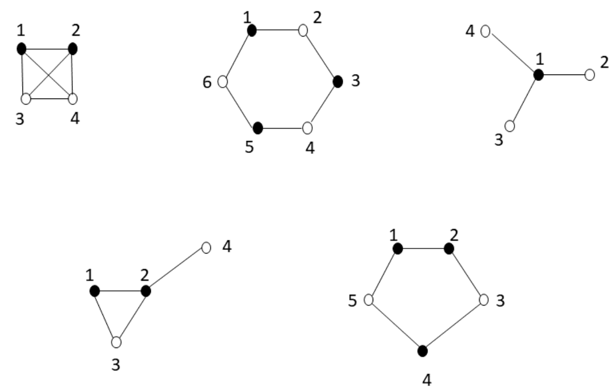


Figure 4: Nash equilibrium on some networks for $\kappa = 0.5$. Black (resp. white) nodes represent motivated (resp. demotivated) agents.

We now complement equilibrium characterization by presenting two properties that hold when $\beta = 1, V = 0$:

For any equilibrium $\mathbf{e}^*$*, the profile* $\mathbf{1} - \mathbf{e}^*$ *is also an equilibrium.*[17] Interestingly, this property rules out the possibility of an agent being locked into demotivation (or motivation) solely based on their position in the network (as we will see thereafter, that possibility emerges when $\beta < 1$ or $V > 0$).

Furthermore, *finding the set of equilibria is a complex problem.* With $\kappa =$

[17]This follows directly from the fact that both motivated and demotivated agents have play an anti-coordination game with the same threshold at equilibrium.

$\frac{1}{2}$, the potential function given in equation (3) counts the number of cross links between motivated and demotivated agents. There is therefore a correspondence with the MaxCut problem. A solution to the MaxCut is, among all two-group partitions, a partition maximizing the number of links between the two groups.[18] By the shape of the potential function, any two-group partition of society that is a solution to the MaxCut problem induces two possible equilibria, in which the set of motivated agents coincides with one of the two groups. The MaxCut problem being NP-complete, we deduce:

Proposition 3 *Assume $\beta = 1$, $V = 0$ and $0 \leq \gamma_L < \gamma_H$. The problem of finding the set of equilibria is NP-complete.*

Polar case 2 (A kink far below neighbors' effort): $\beta < 1, V = 0$. We assume now $\beta < 1$, i.e. demotivation occurs when own effort is sufficiently far below the average of neighbors' effort. Still, we let $V = 0$ meaning no psychologigal cost for demotivation agents, as illustrated in Figure 5. By equation (2), we get

$$\kappa = \frac{2(1-\beta)a + \gamma_H + (1-2\beta)\gamma_L}{2\beta(\gamma_H - \gamma_L)}$$

With $\beta < 1$, agents still anti-coordinate with those of same category, but the tolerance thresholds is now differentiated across categories: motivated agents tolerate more motivated neighbors, while demotivated neighbors tolerate less demotivated neighbors. This asymmetry increases with parameter β. In particular, for β sufficiently low, κ becomes larger than unity, inducing a single equilibrium with no demotivation; This is in sharp contrast with the case $\beta = 1$, which exhibits a huge multiplicity in general.

[18]This is a classical problem of combinatorial optimization, see e.g. Garey and Johnson (1979), Goemans and Williamson (1995).

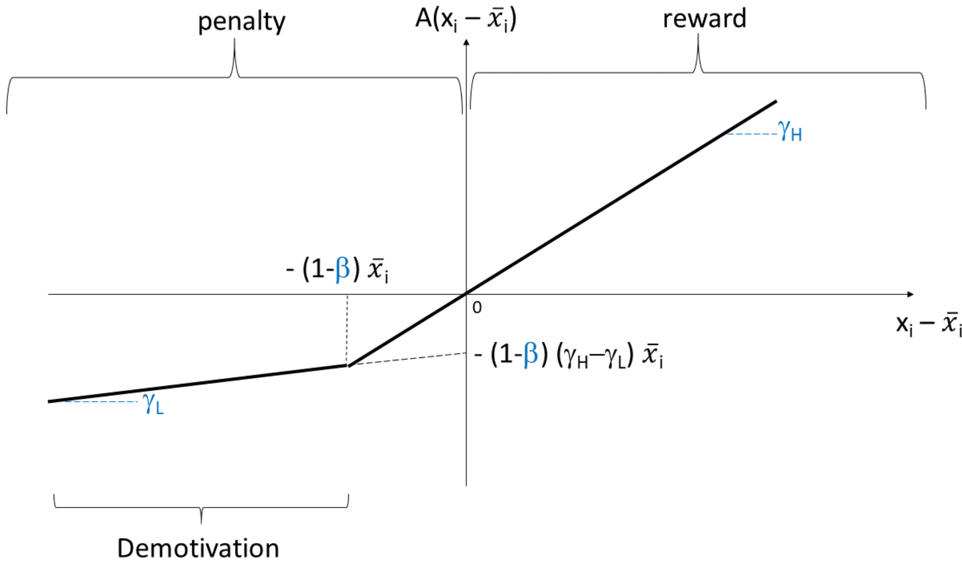
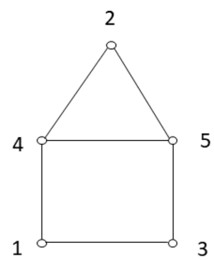


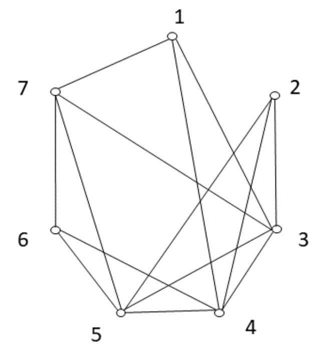
Figure 5: The function A for $\beta < 1$ and $V = 0$.

As said before, no agent can be locked-in to demotivation when $\beta = 1$. Things are different for $\beta < 1$, i.e. *by her position on the network, an agent can be demotivated in all equilibria*. As well, an agent can be motivated in all equilibria. To illustrate, the Left-panel of Figure 6 depicts a five-agent network, with $a = 1, \gamma_H = 1, \gamma_L = 0, V = 0, \beta = 0.92$ (which induces $\kappa = 0.63$). There are 3 equilibria, and agent 2 is never demotivated. The Right-panel of Figure 6 depicts a seven-agent network with same parameters. There are 6 equilibria, and agent 2 is always demotivated.[19] Interestingly, demotivation traps are sensitive to parameter κ, suggesting that targeted public interventions —through modifications of the underlying parameters shaping κ— could help agents escape demotivation traps.

Polar case 3 (Utility loss at the kink): $V > 0, \beta = 1$**.** We assume now that

[19]Finding a general network property ensuring that at least one agent is demotivated in all equilibria as a function of κ is an open issue.





Agent	1	2	3	4	5
Eq 1	1	1	1	0	0
Eq 2	1	1	0	0	1
Eq 3	0	1	1	1	0

Agent	1	2	3	4	5	6	7
Eq 1	0	0	1	1	1	0	1
Eq 2	0	0	1	1	1	1	0
Eq 3	1	0	1	1	1	0	0
Eq 4	1	0	1	0	1	1	0
Eq 5	0	0	1	1	0	1	1
Eq 6	1	0	0	1	1	0	1

Figure 6: Two networks containing locked agents for $a = 1, \gamma_H = 1, \gamma_L = 0, V = 0, \beta = 0.92$. The state of agents in Nash equilibria are given in respective tables below; 1 means motivated, 0 means demotivated.

demotivated agents suffer a loss in utilities that captures some psychological costs, like for instance lowered self-esteem or self-worth, associated with demotivation, i.e. we allow $V > 0$. To isolate the specific effect of utility loss on equilibria, we also assume $\beta = 1$, entailing a change of slope and a discontinuity at 0, as illustrated in Figure 7.

From equation (2), we find

$$\kappa = \frac{1}{2} + \frac{V}{(\gamma_H - \gamma_L)^2}$$

Hence, incorporating a utility loss affects equilibria. The threshold number of mo-

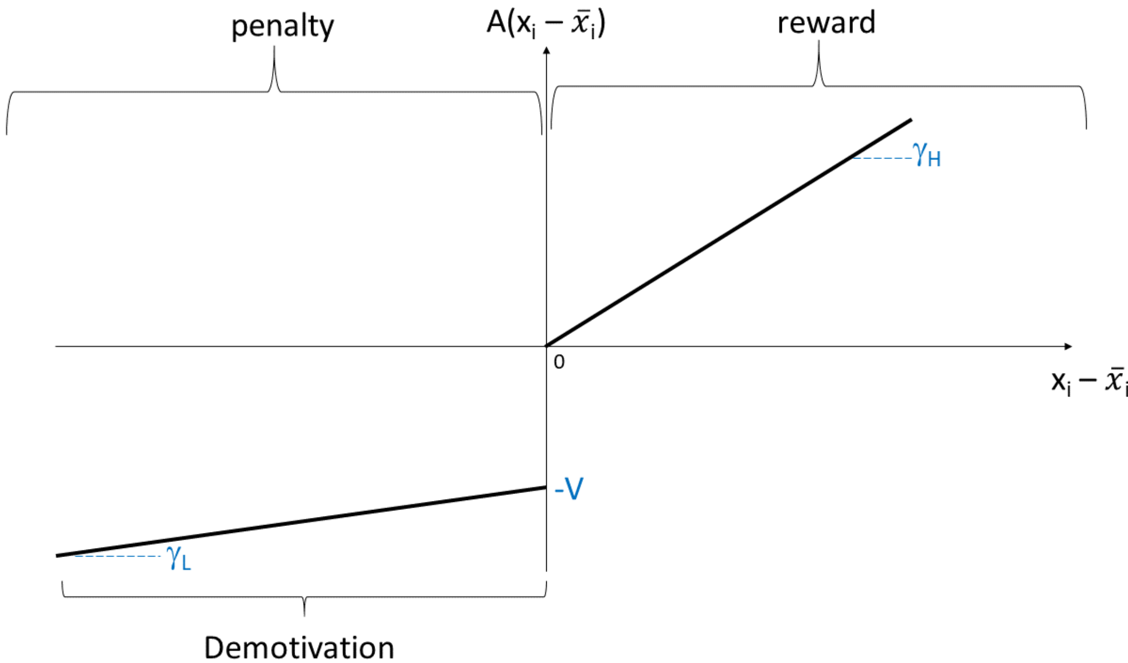


Figure 7: The function A for $\beta = 1$ and $V > 0$.

tivated neighbors above which an agent becomes demotivated is higher, reflecting that the agent prefers to incur a higher cost to exerting high effort and avoiding the utility loss arising under demotivation. Motivation is thus enhanced by the utility loss V. Note that the threshold is decreasing in the gap $\gamma_H - \gamma_L$ whereas it does not depend on the gap when $V = 0$. One consequence of increased threshold is that, as for $\beta < 1$, for some values of parameter V, agents can be locked to demotivation (or motivation) in all equilibria.

4 Demotivation benefits equilibrium welfare

In this section, we examine equilibrium utilities and the welfare of equilibria. Since parameter β plays no pivotal role on result, we assume $\beta = 1$ for simplicity throughout the section.

4.1 A utilitarian approach

We consider a standard utilitarian approach, in which the social welfare for a profile of effort $\mathbf{x}$ is

$$W(\mathbf{x}; \mathbf{G}) = \sum_{i \in \mathcal{N}} u_i(\mathbf{x}; \mathbf{G})$$

We start here by analyzing the no-demotivation scenario $\gamma_L = \gamma_H > 0$ and $V = 0$. In that situation, there is a single Nash equilibrium $\mathbf{x}^*$ in which every agent exerts same effort level $x_i^* = a + \gamma_H$. This implies, for any network $\mathbf{G}$,

$$W(\mathbf{x}^*; \mathbf{G}) = \frac{n}{2}(a^2 - \gamma_H^2) \tag{4}$$

Equilibrium welfare does not depend on the network when there is no kink. Furthermore, equilibrium welfare is lowered compared to the no-status scenario. This is because higher effort means higher effort cost but supplementary effort with respect to the no-status scenario generates no status-related utility gain as everyone does same effort. These results are in line with the conclusions of the economic literature on the impact of status effect on effort and welfare. Next proposition summarizes the main messages from the equilibrium welfare analysis in the absence of demotivation:

Proposition 4 *Consider the no-demotivation scenario, i.e., assume* $0 < \gamma_L = \gamma_H$ *and* $V = 0$*. For any network* $\mathbf{G}$*, the equilibrium welfare under status effect* *without demotivation is lower than the equilibrium welfare in the absence of status effect (for which $\gamma_L = \gamma_H = 0$ and $V = 0$).*

We examine now the impact of demotivation on any equilibrium. Network effects induce heterogeneous externalities on social status along three dimensions. To see this, consider agent i switching from demotivation to motivation, thus generating negative status-related externalities to others. First, there is a composition effect across agent i's neighborhood as the negative impact of the switch is larger on the utility of a motivated neighbor. Moreover, the impact of agent i's switch on a neighbor j is larger when agent j's degree is lower, and the impact is larger when agent i's degree is higher.

At equilibrium $\mathbf{e}^*$, the first-order conditions defining effort are given by $x_i^* = a + \gamma_L + (\gamma_H - \gamma_L)e_i^*$ for agent i. Define $h_i^1(\mathbf{e}) = \sum\limits_{j \in N} \frac{g_{ij}}{dj} e_j$, and $h_i^0(\mathbf{e}) = \sum\limits_{j \in N} \frac{g_{ij}}{dj}(1 - e_j)$. Define $e^* = \mathbf{1}^T \mathbf{e}^*$ for convenience. A few computations provides a characterization of equilibrium welfare:

$$W(\mathbf{e}^*; \mathbf{G}) = \frac{n(a^2 - \gamma_L^2)}{2} + \left(\frac{(\gamma_H - \gamma_L)^2}{2} + V\right)e^* - nV$$
$$- (\gamma_H - \gamma_L)\left(\gamma_L \sum_i e_i^* h_i^0(\mathbf{e}^*) + \gamma_H \sum_i e_i^* h_i^1(\mathbf{e}^*)\right) \tag{5}$$

The welfare of an equilibrium depends three factors: it is increasing in the number of motivated agents e^*, it is decreasing in $\sum_i e_i^* h_i^0$, that captures the aggregate status-related externality generated by motivated agents on demotivated neighbors, and it is decreasing in $\sum_i e_i^* h_i^1$, that captures the aggregate status-related externality generated by motivated agents on motivated neighbors. The impact on motivated neighbors is larger than the impact on demotivated neighbors. Given the conflicting forces shaping equilibrium welfare, the second-best (i.e. the equilibrium with highest welfare) may not contain the largest set of motivated agents, as illustrated on the five-agent network depicted in Figure 8. The figure depicts

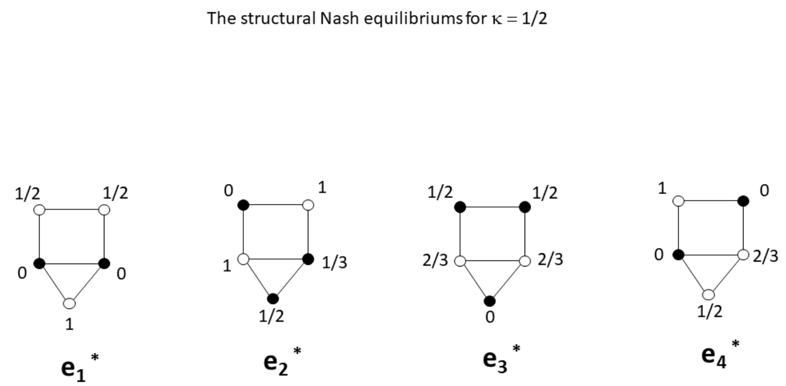


Figure 8: All (structural) Nash equilibria on a 5-agent network, labeled $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ for $\kappa = 0.5$. Black (resp. white) nodes represent motivated (resp. demotivated) agents. Numbers represent the proportion of motivated neighbors.

the four structural Nash equilibria (up to permutation in agents labeling). Table 9 ranks the four Nash equilibria depicted in Figure 8 according to their welfare level.

4.2 Comparison with no-demotivation scenario

Demotivation entails both positive and negative effect. One the one hand, demotivation is good for motivated agents: because demotivated agents reduce their effort level, this enhances their status gain. On the other hand, demotivation has mixed effect on demotivated agents: by reducing their effort level, they trade ef-

$\mathbf{e}^*$	$\mathbf{e}_1^*$	$\mathbf{e}_2^*$	$\mathbf{e}_3^*$	$\mathbf{e}_4^*$
$W(\mathbf{e}^*; \mathbf{G})$	2.83	3.16	3.00	3.50
e^*	2	3	3	2
$\sum_i e_i^* h_i^0(\mathbf{e}^*)$	2	2	1.33	2.16
$\sum_i e_i^* h_i^1(\mathbf{e}^*)$	0.66	0.83	1	0

Figure 9: The welfare of the four Nash equilibria depicted in Figure 8, for $a = 1, \gamma_H = 1, \gamma_L = 0, V = 0$. The three factors shaping the welfare are also presented.

fort cost against status, and they suffer a utility loss. Next Theorem summarizes the implications in terms of welfare.

Theorem 2 *Consider equilibrium* $\mathbf{e}^*$ *on any network* $\mathbf{G}$. *The equilibrium welfare is higher compared to the no-demotivation scenario if and only if*

$$V \leq \frac{\gamma_H - \gamma_L}{2(n - e^*)}\left(n(\gamma_L+\gamma_H)+(\gamma_H-\gamma_L)e^*-2\Big(\gamma_H \sum_i e_i^* h_i^1(\mathbf{e}^*)+\gamma_L \sum_i e_i^* h_i^0(\mathbf{e}^*)\Big)\right)$$

In particular, when $V = 0$, *both motivated and demotivated agents are better off compared to the no-demotivation scenario, meaning that the equilibrium welfare is higher compared to the no-demotivation scenario.*

By Theorem 2, the impact of demotivation, through relaxation of social pressure, is globally positive when psychological costs associated with demotivation are sufficiently low. The threshold on V is equilibrium-dependent (and thus network-dependent). In particular, when psychological costs are negligible, demotivation is welfare-enhancing on any equilibrium. In opposite, significant psychological costs have heavy consequence on equilibrium welfare.

4.3 Comparison with no-status scenario

Status effect, in absence of demotivation, is detrimental to all agents. Demotivation modifies the pictures. It can be good for motivated agents with respect to no-status scenario because, by reducing the social pressure on effort of demotivated agents, it improves the status of motivated agents. However, it is always bad for demotivated agents, who suffer both penalty on status and psychological cost.

Recall that $\rho_i^1(\mathbf{e}^*)$ denotes the share of agent i's neighbors who are motivated in equilibrium $\mathbf{e}^*$. In terms of individual utilities, we obtain:

Proposition 5 *A demotivated agent is strictly worse off compared to the no-status scenario. A motivated agent is strictly better off compared to the no-status scenario if and only if*

$$\rho_i^1(\mathbf{e}^*) \leq \frac{1}{2} - \frac{\gamma_L}{2(\gamma_H - \gamma_L)}$$

Hence, comparing to the no-status scenario, demotivated agents always experience unfavorable social comparisons, incurring both higher effort costs and a status penalty. In opposite, motivated agents can benefit from status effect when the proportion of motivated neighbors is sufficiently low (ensuring high status reward). As intuition suggests, the condition is less demanding when the gap $\gamma_H - \gamma_L$ is larger. Note in particular that, for $\gamma_L = 0$, a motivated agent always benefits from status effect, because the inequality $\rho_i^1(\mathbf{e}^*) \leq \frac{1}{2}$ holds at equilibrium. And in opposite, a motivated agent is always penalized by status effect when γ_L is sufficiently close from γ_H, in particular when $\gamma_H < 2\gamma_L$ (as this means $\frac{1}{2} - \frac{\gamma_L}{2(\gamma_H - \gamma_L)} < 0$).

We then compare equilibrium welfare to the no-status scenario. Overall, the impact on equilibrium welfare depends on both the severity of the kink and the magnitude of psychological cost:

Theorem 3 *Consider equilibrium* $\mathbf{e}^*$ *on any network* $\mathbf{G}$*. The equilibrium welfare is higher compared to the no-status scenario if and only if*

$$V \leq \frac{(\gamma_H - \gamma_L)^2 e^* - (\gamma_H - \gamma_L)\Big(\gamma_H \sum_i e_i^* h_i^1(\mathbf{e}^*) + \gamma_L \sum_i e_i^* h_i^0(\mathbf{e}^*)\Big) - n\gamma_L^2}{2(n - e^*)}$$

By Theorem 3, demotivation can be welfare-improving with respect to the no-status scenario when demotivated agents suffer sufficiently low pychological costs, and when the kink is sufficiently pronounced. The critical bound on psychological cost is equilibrium-dependent.

Under sufficiently low psychological costs, the key driver is the ratio $\frac{\gamma_L}{\gamma_H}$. For convenience, we denote $\phi^* = \sum_i e_i^*(1 - 2\rho_i^1(\mathbf{e}^*)) \geq 0$ (as, in any equilibrium, for all $i : e_i^* = 1$, $\rho_i^1(\mathbf{e}^*) \leq \frac{1}{2}$) and $\psi^* = \sum_i \rho_i^1(\mathbf{e}^*)$. Recall that in any equilibrium, for all $i : e_i^* = 0$, $\rho_i^1(\mathbf{e}^*) \geq \frac{1}{2}$. We obtain:

Proposition 6 *Assume* $\beta = 1$*,* $V = 0$ *and* $0 \leq \gamma_L < \gamma_H$*. For any network* $\mathbf{G}$*, the welfare at equilibrium* $\mathbf{e}^*$ *is larger than the equilibrium welfare in the absence of status effect (i.e.,* $\gamma_L = \gamma_H = 0$*) if and only if the ratio* $\frac{\gamma_L}{\gamma_H}$ *is lower than the following threshold* $\tau_c(\mathbf{e}^*)$*:*

If $\forall i : e_i^* = 0$*,* $\rho_i^1(\mathbf{e}^*) = \frac{1}{2}$*, then* $\tau_c(\mathbf{e}^*) = \frac{\phi^*}{2(\phi^* + \psi^*)}$*. Otherwise,*

$$\tau_c(\mathbf{e}^*) = \frac{\phi^* + \psi^* - \sqrt{(\psi^*)^2 + n\phi^*}}{\phi^* - n + 2\psi^*}$$

Note that many equilibria are such that, for all demotivated agents, the proportion of motivated neighbors is equal to $\frac{1}{2}$ (so that the relevant condition in the above proposition is the first one).[20] By Proposition 6, the presence of demotivated agents induces a complete reversal in the qualitative effect of status on welfare. While status effects reduce equilibrium welfare in the absence of the demotivation kink, a sufficiently pronounced kink can lead to welfare gains, as demotivated agents strongly enhance the status, and thus the utility, of motivated agents. Again, the threshold ratio for which welfare with status effect dominates welfare without status effect is equilibrium-dependent (and thus also network-dependent).

In the extreme case where $\gamma_L = 0$, demotivated agents are not affected by status effect in absence of psychological cost, meaning that, for them, both effort and utilities are equal to those in the absence of status. Then, only motivated agents are affected by status. By the presence of demotivated agent, motivated agents get status-related reward, and that reward dominates the extra effort cost necessary to be motivated by construction of the equilibrium. Therefore:[21]

Corollary 1 *Assume $\beta = 1, V = 0, \gamma_L = 0$ and $\gamma_H > 0$. The welfare of an equilibrium $\mathbf{e}^*$ on a given network $\mathbf{G}$ is given by*

$$W(\mathbf{e}^*;\mathbf{G}) = \frac{na^2}{2} + \frac{\gamma_H^2}{2}\sum_i e_i^*(1 - 2\rho_i^1(\mathbf{e}^*))$$

Hence, for all networks, the welfare of any equilibrium is larger than the welfare of the equilibrium in the no-status scenario.

[20] For instance, consider a three-agent complete network, with two demotivated agents, or consider a eight-agent circle, in which agents $1, 2, 5, 6$ are motivated, and agents $3, 4, 7, 8$ are demotivated.

[21] The proof of Corollary 1 is immediate, recalling that $\rho_i^1(\mathbf{e}^*) \leq \frac{1}{2}$ for all $i : e_i^* = 1$ on any equilibrium.

4.4 Statics on psychological cost V

The impact of the psychological costs of demotivated agents on the welfare of equilibria is subtle. To see this, we consider an increase of V. When the increase does not affect the set of equilibria, a higher utility cost V lowers the welfare of all equilibria containing demotivated agents. However, increasing V can affect the set of equilibria by increasing incentives to be motivated, and this can result in higher equilibrium welfare.

We illustrate how this countervailing effect operates by focusing on the second-best equilibrium, assuming $\gamma_L = 0$ for simplicity. Few computation gives the welfare of an equilibrium:

$$W(\mathbf{e}^*) = \frac{na^2}{2} + \frac{\gamma_H^2}{2} e^* - \gamma_H^2 \sum_i e_i^* \rho_i^1(\mathbf{e}^*) - (n - e^*)V$$

Hence, the equilibrium welfare takes into account the aggregate utility loss $(n - e^*)V$ generated by demotivated agents, and it also takes into account the negative aggregate impact of status among motivated agents, as measured by the sum over all motivated agents of the shares of their motivated neighbors. In the example given by the 11-agent network depicted in Figure 10, increased V enhances the welfare at the second-best equilibrium for the following parameter values. We fix $a = 2, \gamma_H = 1, \gamma_L = 0, \beta = 1$, and we consider $V = 0.16$ and $V = 0.17$. Setting $V = 0.16$, we get $\kappa = 0.66$. With these parameter values, the network depicted in the figure has ten equilibria. The second-best equilibrium, presented in the Left-panel, reaches a welfare of 22.26. We note that the sum of the shares of motivated neighbors over all motivated agents is equal to 2.6. For $V = 0.17$, we obtain $\kappa = 0.67$. Again, there are ten equilibria, but the second-best equilibrium is modified. The second-best equilibrium for $V = 0.17$ is depicted in the Right-

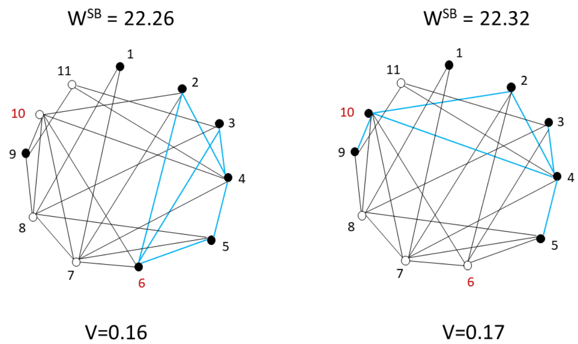


Figure 10: Increased V can improve the welfare of the second-best equilibrium; $n = 11$, $a = 2, \gamma_H = 1, \gamma_L = 0, \beta = 1$. Black nodes (resp. white nodes) are motivated (resp. demotivated). The links among motivated agents are in blue.

panel. This slight increase in V modifies the incentives of agents 10 and 6: Agent 10 becomes motivated by the increase of κ, but that change makes agent 6 become demotivated. The consequence on the welfare at second-best is as follows: given that the number of motivated agents is unchanged between the two second-bests, the increase in V is neutral with respect to the aggregate utility losses. The only difference is the sum of the shares of motivated neighbors over motivated agents, that is now equal to 2.5 (note that the share of motivated neighbors is modified for many agents). This reduction of the aggregate (negative) impact of motivated agents on others is good for the welfare.

4.5 Statics on the kink

We examine how the sharpness of the kink in the status function affects equilibrium welfare through separate comparative statics with respect to parameters

γ_L, γ_H. As observed earlier, the equilibria do not depend on γ_L, γ_H when $V = 0$, but do depend on those parameters for $V > 0$. For simplicity we undertake these two statics assuming that the set of equilibria is unaffected by a change in these parameters. Interestingly, the two statics do not have a symmetric effect on equilibrium welfare.[22]

Statics on γ_L. To start with, we examine how a marginal decrease in parameter γ_L affects agents' utilities:

Proposition 7 *Assume $\beta = 1, V \geq 0$ and $0 \leq \gamma_L < \gamma_H$. Let $\mathbf{G}$ be any network, and $\mathbf{e}^*$ any corresponding equilibrium robust to a marginal decrease of γ_L. When γ_L decreases marginally, this improves the equilibrium utilities of both motivated and demotivated agents.*

By Proposition 7, sharpening the kink through a reduction of γ_L induces a reduction of the pressure of social comparison, which benefits both demotivated and motivated agents. Hence, decreased γ_L improves equilibrium welfare.

Statics on γ_H. We examine how a marginal increase in parameter γ_H affects agents' utilities:

Proposition 8 *Assume $\beta = 1, V \geq 0$ and $0 \leq \gamma_L < \gamma_H$. Let $\mathbf{G}$ be any network, and $\mathbf{e}^*$ any equilibrium robust to a marginal increase of γ_H. When γ_H increases marginally, this reduces the equilibrium utilities of demotivated agents, and this increases the equilibrium utility of a motivated agent i if and only if*

$$\rho_i^1(\mathbf{e}^*) < \frac{1}{2 + \frac{\gamma_L}{\gamma_H - \gamma_L}}$$

By Proposition 8, sharpening the kink through an increase of γ_H induces more pressure from social comparison, which is detrimental to demotivated agents but

[22] Hence, a rotation of the kink preserving its angle is not neutral for welfare.

leads motivated agents to be better off through a gain in status when the proportion of their motivated neighbors is low enough (which ensures a high status reward). Overall, the qualitative impact on equilibrium welfare depends on the network:

Proposition 9 *Assume* $\beta = 1$*,* $V = 0$*, and* $0 \leq \gamma_L < \gamma_H$*. For any network* $\mathbf{G}$*, any equilibrium* $\mathbf{e}^*$*, the equilibrium welfare is increasing in parameter* γ_H *if and only if*

$$\sum_i e_i^*(1 - 2\rho_i^1(\mathbf{e}^*)) > \frac{\gamma_L}{\gamma_H - \gamma_L} \sum_i \rho_i^1(\mathbf{e}^*) \tag{6}$$

By Proposition 9, the impact of increased γ_H, through a higher social pressure, is equilibrium-dependent. The condition given in equation (6) is more favorable to welfare improvement for lower values of γ_L – It is met for $\gamma_L = 0$ and it fails for γ_L sufficiently close to γ_H.

4.6 Pareto-dominance

We also investigate whether some equilibria Pareto-dominate others. Take the 4-star network, which has two equilibria. It is easily shown that no equilibrium Pareto-dominates the other. That message is confirmed more generally:

Proposition 10 *Assume* $\beta = 1$*, and* $0 \leq \gamma_L < \gamma_H$*. For any network* $\mathbf{G}$*, no equilibrium Pareto-dominates another equilibrium.*

The absence of Pareto-dominance among equilibria stems from the fact that it is always better to be motivated in a given equilibrium than to be demotivated in any equilibrium. Then the point follows from the non-nestedness property of equilibria that stems from the nature of the game (as presented before).

5 Extensions

We present three extensions, introducing heterogeneous private returns, more general utility functions, and local synergies. All formulations preserve the strategic substitute nature of incentives, and a potential function exists in all extensions. For simplicity, we assume $V = 0$ in all three extensions, meaning that in what follows, demotivation is solely reflected in a lower return to effort.

Heterogeneous private returns. We assume now that utilities are as follows:

$$u_i(x_i, x_{-i}) = a_i x_i - \frac{1}{2}x_i^2 + A(x_i - \overline{x}_i)$$

with $a_i > 0$ the private return of agent i. In that extended setting, best-responses are given as follows. Denote by $\overline{a}_i = \frac{1}{d_i}\sum_j g_{ij}a_j$ agent i's average neighbors' private returns. Let $\varphi_i = \frac{a_i + \frac{\gamma_H + \gamma_L}{2}}{\beta}$. Agent i's best-response to $\overline{x}_i$ is given by:

$$\begin{cases} x_i^{BR}(\overline{x}_i) = a_i + \gamma_H & \text{if} \quad \overline{x}_i \leq \varphi_i \\ x_i^{BR}(\overline{x}_i) = a_i + \gamma_L & \text{if} \quad \overline{x}_i > \varphi_i \end{cases}$$

Then, we define

$$\kappa_i = \frac{2(a_i - \beta\overline{a}_i) + \gamma_H + (1 - 2\beta)\gamma_L}{2\beta(\gamma_H - \gamma_L)}$$

Parameter κ_i is increasing in agent i's private return a_i, meaning that agents with higher private returns are less likely to become demotivated.

Proposition 11 *Under heterogeneous private returns, agents play a potential game of anti-coordination, hence there is always an equilibrium. A configuration* $\mathbf{e}^*$ *is a Nash equilibrium if and only if*

$$e_i = 1 \quad \Rightarrow \quad \rho_i^1(\mathbf{e}) \leq \kappa_i$$

$$e_i = 0 \quad \Rightarrow \quad \rho_i^0(\mathbf{e}) \leq 1 - \kappa_i$$

When $\mathbf{G}^T = \mathbf{G}$, the game with $\beta \leq 1$ admits a potential function

$$P(\mathbf{e}) = \sum_{i \in \mathcal{N}} \kappa_i d_i e_i - \frac{1}{2}\mathbf{e}^T \mathbf{G}\mathbf{e}$$

which ensures equilibrium existence.

Generalizing on utility function. The model generates strategic substitutes under more general utility functions. We consider agent i's utility:

$$u_i(x_i, \overline{x}_i) = v_i(x_i) + A(x_i - \overline{x}_i)$$

where function v_i is a concave (and single-peaked function), with $v_i(0) = 0$. For simplicity, we assume $\beta = 1$ in the status function. Agent i's best-response is thus

$$\begin{cases} x_i^{\gamma_H}(\overline{x}_i) = v_i'^{-1}(-\gamma_H) & \text{if} \quad x_i - \overline{x}_i \geq 0 \\ x_i^{\gamma_L}(\overline{x}_i) = v_i'^{-1}(-\gamma_L) & \text{if} \quad x_i - \overline{x}_i < 0 \end{cases}$$

The structure of best-responses is the same as the linear quadratic case. That is, agent i's best-response is a step function (replacing $a_i + \gamma_H$ by $x_i^{\gamma_H}$ and $a_i + \gamma_L$ by $x_i^{\gamma_L}$). By concavity of v_i, function $v_i'^{-1}$ is increasing, so that $x_i^{\gamma_H} > x_i^{\gamma_L}$. Letting $v_i^{\gamma_L} = v_i(v_i'^{-1}(-\gamma_L)), v_i^{\gamma_H} = v_i(v_i'^{-1}(-\gamma_H))$, the threshold φ_i below which agents play demotivated satisfies $u_i(x_i^{\gamma_H}, \overline{x}_i) = u_i(x_i^{\gamma_L}, \overline{x}_i)$, that is,

$$\varphi_i = \frac{v_i^{\gamma_H} - v_i^{\gamma_L} + \gamma_H(x_i^{\gamma_H} - x_i^{\gamma_L})}{\gamma_H - \gamma_L}$$

Let binary profile $\mathbf{e}$ describe the status of agents. At equilibrium $\mathbf{e}^*$,

$$x_i^* = v_i^{\gamma_H} + (v_i^{\gamma_H} - v_i^{\gamma_L})e_i^*$$

Let $\mathbf{H}$ be such that $h_{ij} = g_{ij}(v_j^{\gamma_H} - v_j^{\gamma_L})$ and $\kappa_i = \varphi_i - \frac{1}{d_i}\sum_j g_{ij} v_j^{\gamma_L}$ for convenience. We obtain $\overline{x}_i \leq \varphi_i$ if and only if

$$\sum_j h_{ij} e_j^* \leq \kappa_i d_i$$

When function $v_i = v$ for all i, matrix $\mathbf{H}$ is symmetric. Hence, as with linear quadratic function:

Proposition 12 *When function $v_i = v$ for all i, the game is a potential game with strategic substitutes.*

Incorporating local synergies. We incorporate local synergies in sum in the model. For instance, this can fit with applications related to education, or workers.[23] Like status effects, synergies tend to higher effort as a source of strategic complementarities. The analysis mainly suggests that even in presence of synergies the strategic substitute nature of interactions is a robust mechanism. However the analysis is challenging, and identifying who is demotivated is more complex than examining the sole neighbors' behavior.

We assume the following specification:

$$u_i = x_i - \frac{1}{2}x_i^2 + \delta x_i d_i \overline{x}_i + A(x_i - \overline{x}_i)$$

with parameter $\delta \geq 0$ representing the intensity of synergies among neighbors. For simplicity, we assume $\beta = 1$ in the status function (the proof of Theorem 4 below is presented for $\beta \leq 1$). The network intervenes twice, shaping local synergies and social comparison. Local synergies are the sum of neighbors' bilateral synergies, and agents compare their effort to the average of their peers (the benchmark model studied in the paper corresponds to assuming $\delta = 0$). The equilibrium conditions are as follows (see the proof of Theorem 4 thereafter for more details):

$$[(\mathbf{M} - \mathbf{I})\mathbf{e}]_i \leq \mathbf{k}(\mathbf{G}, \delta)$$

[23]See Calvó-Armengol et al. (2009) for empirical evidence of synergies in school context, or Cornelissen et al. (2017) at the workplace.

where vector $k(\mathbf{G},\delta) = (\kappa_i(\mathbf{G},\delta))_{i\in\mathcal{N}}$ is such that

$$\kappa_i(\mathbf{G},\delta) = \frac{\delta d_i(\gamma_H - \gamma_L) + 2(a+\gamma_L)}{2(\gamma_H - \gamma_L)(1-\delta d_i)} - \frac{a+\gamma_L}{\gamma_H - \gamma_L} b_i \tag{7}$$

Shortly speaking, the first-order conditions indicate that an agent is motivated when her connection to other motivated agents is sufficiently low, where 'connection' is no longer restricted to direct neighbors, but is extended to account for paths of any length with decay. One interest with that specification is that, like the no-synergy case, under symmetry of matrix $\mathbf{G}$, the game still admits a potential function:[24]

$$P(\mathbf{e}) = \mathbf{k}(\mathbf{G},\delta)^T\mathbf{e} - \frac{1}{2}\mathbf{e}^T(\mathbf{M}-\mathbf{I})\mathbf{e}$$

This guarantees equilibrium existence. The general picture is then as follows. For $\gamma_L = \gamma_H$ and $\delta > 0$, the model is the game of local synergies. There is a single equilibrium given by standard Bonacich centralities. For $\delta = 0$, we get the anti-coordination game, exhibiting multiple equilibria. In-between, i.e. for $\gamma_L < \gamma_H$ and $\delta > 0$, the game incorporates both local synergies and anti-coordination. Uniqueness should therefore be confined to a region of the parameter space such that the anti-coordination effect is dominated by the local synergy effect. Let matrix $\mathbf{M} = (\mathbf{I} - \delta\mathbf{G})^{-1}$, vector $\mathbf{b} = \mathbf{M}\mathbf{1}$. We obtain:

Theorem 4 *The game with status effect including demotivation and local synergies in sum is a potential game. Let $\mathbf{e}^*$ be an equilibrium. Then*

- *if $d_i \geq \frac{1}{\delta}$, $e_i^* = 1$.*
- *if $d_i < \frac{1}{\delta}$, $e_i^* = 1$ if and only if*

$$[(\mathbf{M}-\mathbf{I})\mathbf{e}^*]_i \leq \kappa_i(\mathbf{G},\delta) \tag{8}$$

[24] In opposite, there is no potential with linear-in-means synergies, because in that case the interaction stays asymmetric even if $\mathbf{G}$ is symmetric.

with κ_i defined as in (7). The equilibrium effort profile is given by

$$\mathbf{x}^* = (a + \gamma_L)\mathbf{b} + (\gamma_H - \gamma_L)\mathbf{M}\mathbf{e}^*$$

Conform to intuition, more demotivation, through lower γ_L, leads to more anti-coordination and thus favors multiplicity. However, the role of synergies is ambiguous with respect to demotivation because increased effort can induce mixed effect on social status. There are at least two consequences. First, more synergies can increase the number of equilibria. Second, more synergies can reduce the number of motivated agents (simple examples illustrate both claims).

The case of extreme synergies calls for interest. Recall that the intensity of synergies is not larger than $\bar{\delta} = \frac{1}{\mu(\mathbf{G})}$:

Corollary 2 *When δ is sufficiently close to $\bar{\delta}$, there is a single equilibrium. In that equilibrium, agent i is demotivated if and only if $d_i \leq \mu(\boldsymbol{G})$.*

Here the synergies strongly dominate anti-coordination, so that the uniqueness property that holds under synergies and no anti-coordination extends straightforwardly.

6 Conclusion

This paper analyzes demotivation stemming from unfavorable social comparison with local peers in networks. We modeled effort decisions with a status-dependent utility function, and introduced demotivation through two features: a lower return to effort when effort is below a fixed percentage of neighbors' effort, and a utility loss reflecting psychological cost. The introduction of demotivation entails a binary potential game of strategic substitutes played on networks, generating multiple Nash equilibria in general. Our findings also highlight that demotivated agents can become locked into low-effort outcomes due to their placement in the network, and that demotivation can have a positive impact on equilibrium welfare, by improving the social status of peers, when psychological costs are sufficiently low. These results are robust to various model extensions. These insights provide a theoretical foundation for understanding better how network structures influence demotivation and performance in social and economic settings.

This theoretical work opens up several avenues for future research. First, the model might generate testable predictions linking network position, relative performance, and individual effort levels — in particular, it predicts that individuals who are demotivated relative to their peers exert systematically lower effort, even at equal ability. This suggests empirical strategies exploiting exogenous variation in local peer performance or network structure to identify demotivation effects. Second, social comparison —and its link to demotivation— likely plays a key role in shaping social networks. Introducing endogenous network formation in a setting where effort incentives depend on relative comparisons could shed light on how individuals strategically choose their social ties in order to avoid being demotivated.

References

Bellemare, C., Lepage, P., and Shearer, B. (2010). Peer pressure, incentives, and gender: An experimental analysis of motivation in the workplace. *Labor Economics*, 17(1):276–283.

Blume, L. (1993). The statistical mechanics of strategic interaction. *Games and*

Economic Behavior, 5:387–424.

Bramoullé, Y. (2007). Anti-coordination and social interactions. *Games and Economic Behavior*, 58.

Bramoullé, Y. and Ghiglino, C. (2024). Status consumption in networks: a reference dependent approach. *Unpublished*.

Bénabou, R. and Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(3):489–520.

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

Cao, Z., Gao, H., Qu, X., Yang, M., and Yang, X. (2013). Fashion, cooperation, and social interactions. *PLoS ONE*, 8(e49441).

Clark, A. and Oswald, A. (1996). Satisfaction and comparison income. *Journal of Public Economics*, 61(3):359–381.

Collins, R. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, 119(1):51–69.

Cornelissen, T., Dustmann, C., and Schönberg, U. (2017). Peer effects in the workplace. *American Economic Review*, 107(2):425–456.

Dening, J., Murphy, R., and Weinhardt, F. (2023). Class rank and long-run outcomes. *The Review of Economics and Statistics*, 105(6):1426–1441.

Diks, K., Garrido, A., and Lingas, A. (1994). The maximum k-dependent and f-dependent set problem. *Lecture Notes in Computer Science; Springer*, 855.

Duesenberry, J. (1949). Income, saving, and the theory of consumer behavior. *Harvard University Press*.

Eriksson, T., Poulsen, A., and Villeval, M.-C. (2009). Feedback and incentives: Experimental evidence. *Labor Economics*, 16(6):679–688.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2):117–140.

Frank, R. (1985). The demand for unobservable and other nonpositional goods. *American Economic Review*, 75(1):101–116.

Frank, R. (2005). Positional externalities cause large and preventable welfare losses. *American Economic Review*, 95(2):137–141.

Garey, M. and Johnson, D. (1979). Computers and intractability: A guide to the theory of np-completeness. *Freeman, San Francisco*.

Ghiglino, C. and Goyal, S. (2010). Keeping up with the neighbors: Social interaction in a market economy. *Journal of the European Economic Association*, 8(1):90–119.

Gil, D. and Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, 102(1):469–503.

Goemans, M. and Williamson, D. (1995). Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Machinery*, 42(6):1115–1145.

Hopkins, E. and Kornienko, T. (2004). Running to keep in the same place: Consumer choice as a game of status. *American Economic Review*, 94(4):1085–1107.

Immorlica, N., Kranton, R., Manea, M., and Stoddard, G. (2017). ocial status in networks. *American Economic Journal: Microeconomics*, 9(1):1–30.

Langtry, A. (2023). Keeping up with "the joneses": Reference-dependent choice with social comparisons. *American Economic Journal: Microeconomics*, 15(3):474–500.

Lazear, E. and Rosen, S. (1981). Rank-order tournaments as optimal labor contracts. *Journal of Political Economy*, 89(5):841–864.

Lockwood, P. and Kunda, Z. (2003). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology*, 73:91–103.

Luttmer, E. (2005). Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics*, 120(3):963–1002.

Monderer, D. and Shapley, L. (1996). Potential games. *Games and Economic Behavior*, 14:124–143.

Murphy, R. and Weinhardt, F. (2020). Top of the class: The importance of ordinal rank. *The Review of Economic Studies*, 87(6):2777–2826.

Rogers, T. and Feller, A. (2016). Demotivated by peer excellence: exposure to exemplary peer performance causes quitting. *Psychological Science*, 27(3):365–374.

Rosenthal, R. (2017). A class of games possessing pure-strategy nash equilibria. *Int J of Game Theory*, 2:65–67.

Staab, M. (2024). The formation of social groups under status concern. *Journal of Economic Theory*, 222.

Tesser, A., Millar, M., and Moore, J. (1988). Some affective consequences of social comparison and reflection processes: The pain and pleasure of being close. *Journal of Personality and Social Psychology*, 57:102048.

Veblen, T. (1899). The theory of the leisure class: An economic study of institutions. *New York, NY: Macmillan*.

Voorneveld, M. (2000). Best-response potential games. *Economics Letters*, 66:289–295.

Wheeler, L. and Miyake, K. (1992). Social comparison in everyday life. *Journal of Personality and Social Psychology*, 62:760–773.

Young, P. (1998). Individual strategy and social structure. *Princeton University*.

A Appendix A: Proofs

[Proof of Proposition 1] To define the best-response to others' play, consider the restriction of utilities in each domain (no-demotivation and demotivation).

For convenience, write $x_i^{\gamma_H}$ (resp. $x_i^{\gamma_L}$) the best-response of agent i under no-demotivation (resp. under demotivation). Without demotivation,

$$x_i^{\gamma_H} = a + \gamma_H$$

which holds for $x_i^{\gamma_H} \geq \beta \overline{x}_i$. Feasibility then implies $\overline{x}_i \leq \overline{x}^0 = \frac{a+\gamma_H}{\beta}$. With demotivation,

$$x_i^{\gamma_L} = a + \gamma_L$$

which holds for $x_i^{\gamma_L} \leq \beta \overline{x}_i$. Feasibility then implies $\overline{x}_i \geq \overline{x}^1 = \frac{a+\gamma_L}{\beta}$. Clearly $\overline{x}^1 < \overline{x}^0$ since $\gamma_L < \gamma_H$. We then need to compare utilities for $\overline{x}_i \in [\overline{x}^1, \overline{x}^0]$. We find $U_i(x_i^{\gamma_H}) \geq U_i(x_i^{\gamma_L})$ whenever

$$\frac{1}{2}(a_i + \gamma_H)^2 - \gamma_H \overline{x}_i \geq \frac{1}{2}(a_i + \gamma_L)^2 - ((1-\beta)\gamma_H + \beta\gamma_L)\overline{x}_i - V$$

That is:

$$\overline{x}_i \leq \varphi = \frac{1}{\beta}\left(a + \frac{\gamma_H + \gamma_L}{2} + \frac{V}{\gamma_H - \gamma_L}\right)$$

[Proof of Theorem 1] We start with the characterization, and then we turn to existence.

Characterization. Let $\mathbf{e} \in \{0,1\}^n$ be a binary vector such that $e_i = 1$ if agent i's plays is not in the demotivation region, and $e_i = 0$ if agent i's plays is in the demotivation region. Inverting the system of best-responses, an equilibrium can be written

$$\mathbf{x}^* = (a + \gamma_L)\mathbf{1} + (\gamma_H - \gamma_L)\mathbf{e}^*$$

Noticing that $\tilde{\mathbf{G}}\mathbf{1} = \mathbf{1}$, we get

$$\overline{\mathbf{x}}^* = (a + \gamma_L)\mathbf{1} + (\gamma_H - \gamma_L)\tilde{\mathbf{G}}\mathbf{e}^*$$

Then agent i plays in the motivated region whenever $\overline{x}_i^* \leq \varphi$, i.e.,

$$a + \gamma_L + (\gamma_H - \gamma_L)\rho_i^1(\mathbf{e}) \leq \frac{1}{\beta}\left(a + \frac{\gamma_H + \gamma_L}{2} + \frac{V}{\gamma_H - \gamma_L}\right)$$

That is,

$$\rho_i^1(\mathbf{e}) \leq \kappa = \frac{2\beta a + \gamma_H + (1-2\beta)\gamma_L}{2\beta(\gamma_H - \gamma_L)} + \frac{V}{\beta(\gamma_H - \gamma_L)^2}$$

Existence. Let $d_i^1 = [\mathbf{GE}]_i$ represent the number of agent i's neighbors being in state 1. We have

$$e_i^* = 1 \Rightarrow d_i^1(\mathbf{e}) \leq \kappa d_i$$
$$e_i^* = 0 \Rightarrow d_i^1(\mathbf{e}) \geq \kappa d_i$$

Consider the function

$$P(\mathbf{e}) = \kappa \mathbf{1}^T \mathbf{G}\mathbf{e} - \frac{1}{2}\mathbf{e}^T \mathbf{G}\mathbf{e}$$

with binary matrix $\mathbf{G}$ representing the network. This function is a potential of the demotivation game when $\mathbf{G}^T = \mathbf{G}$. Indeed, the impact of a switch to motivation is as follows. Let $\mathbf{e}$ and $i : e_i = 0$, and let $\mathbf{1}_i = (0, 0, \cdots, 1, 0, \cdots, 0)^T$ the vector of zeros but a one at entry i. Since $\mathbf{G}^T = \mathbf{G}$,

$$P(\mathbf{e} + \mathbf{1}_i) - P(\mathbf{e}) = \kappa d_i - d_i^1(\mathbf{e})$$

Hence, function P increases with the switch whenever $\frac{d_i^1(\mathbf{e})}{d_i} \leq \kappa$, which is equivalent to playing $e_i = 1$ as a best-response. That potential has a maximum (the strategy space is finite). Hence any maximum of the potential over the strategy space is a Nash equilibrium.

[Proof of Theorem 2] The first sentence follows from direct comparison between equation (5) and equation (4). The proof of the second sentence is a direct consequence of Proposition 7.

[Proof of Proposition 5] In the no-status scenario ($\gamma_L = \gamma_H = 0$), there is a single equilibrium in which every agent's effort is a, and equilibrium utility is $\frac{a^2}{2}$.

A demotivated agent i is worse off:

$$u_i^* = \frac{(a + \gamma_L)^2}{2} - \gamma_l \overline{x}_i^* \leq \frac{(a + \gamma_L)^2}{2} - \gamma_l (a + \gamma_L) = \frac{a^2 - \gamma_L^2}{2} < \frac{a^2}{2}$$

Consider now a motivated agent i, whose equilibrium utility can be written

$$u_i^* = \frac{(a+\gamma_H)^2}{2} - \gamma_H\big(a+\gamma_L+(\gamma_H-\gamma_L)\rho_i^1(\mathbf{e}^*)\big)$$

Then, $u_i^* \geq \frac{a^2}{2}$ whenever

$$\rho_i^1(\mathbf{e}^*) \leq \frac{1}{2} - \frac{\gamma_L}{2(\gamma_H-\gamma_L)}$$

[Proof of Proposition 6] At equilibrium $\mathbf{e}^*$, the first-order conditions defining effort are given by $x_i^* = a+\gamma_L+(\gamma_H-\gamma_L)e_i^*$ for agent i. We then get

$$W(\mathbf{e}^*) = \frac{1}{2}\sum_{i\in N}\big(x_i^*\big)^2 - \sum_{i\in N}\big(\gamma_L+(\gamma_H-\gamma_L)e_i^*\big)\overline{x}_i^*$$

That is, plugging effort into equilibrium welfare,

$$W(\mathbf{e}^*) = \frac{1}{2}\sum_i(a+\gamma_L+(\gamma_H-\gamma_L)e_i^*)^2 - \sum_i(\gamma_L+(\gamma_H-\gamma_L)e_i^*)(a+\gamma_L+(\gamma_H-\gamma_L)\rho_i^1(\mathbf{e}^*))$$

That is,

$$W(\mathbf{e}^*) = \frac{na^2}{2} - \frac{n}{2}\gamma_L^2 - (\gamma_H-\gamma_L)\Big(\sum_i\rho_i^1(\mathbf{e}^*)\Big)\gamma_L + \frac{(\gamma_H-\gamma_L)^2}{2}\Big(\sum_i e_i^*(1-2\rho_i^1(\mathbf{e}^*))\Big)$$

That is, denoting $\phi^* = \sum_i e_i^*(1-2\rho_i^1(\mathbf{e}^*)) \geq 0$ and $\psi^* = \sum_i \rho_i^1(\mathbf{e}^*)$ for convenience,

$$W(\mathbf{e}^*) = \frac{na^2}{2} + P_2(\gamma_L)$$

with

$$P_2(\gamma_L) = \underbrace{\left(\frac{\phi^*-n+2\psi^*}{2}\right)}_{a_2}\gamma_L^2 + \underbrace{\gamma_H\Big(-\phi^*-\psi^*\Big)}_{a_1}\gamma_L + \underbrace{\frac{1}{2}\gamma_H^2\phi^*}_{a_0}$$

Clearly $a_0 > 0, a_1 < 0$. Also, $a_2 \geq 0$: indeed

$$a_2 = \sum_i(1-e_i^*)(2\rho_i^1(\mathbf{e}^*)-1)$$

and, for all $e_i^* = 0$, and all $\beta \leq 1$, $\rho_i^1(\mathbf{e}^*) \geq \frac{1}{2}$. Then two cases can arise.

Case 1: $a_2 = 0$; Then $\frac{\partial W(\mathbf{e}^*)}{\partial \gamma_L} < 0$, and $W(\mathbf{e}^*) \geq \frac{na^2}{2}$ whenever

$$\gamma_L \leq \gamma_H \cdot \frac{\phi^*}{2(\phi^* + \psi^*)}$$

Case 2: $a_2 > 0$; Then $P_2(\gamma_L)$ has two positive roots γ_l', γ_L''. Since $P_2(\gamma_H) < 0 < P_2(0)$, it follows that $0 < \gamma_L' < \gamma_H < \gamma_L''$. Therefore, $W(\mathbf{e}^*) \geq \frac{na^2}{2}$ whenever

$$\gamma_L \leq \gamma_H \cdot \frac{\phi^* + \psi^* - \sqrt{(\psi^*)^2 + n\phi^*}}{\phi^* - n + 2\psi^*}$$

[Proof of Proposition 7] Assume $0 \leq \gamma_L < \gamma_H$ and consider an equilibrium $\mathbf{e}^*$. For a motivated agent i, equilibrium effort is $x_i^* = a + \gamma_H$, and equilibrium utility $u_i^* = \frac{(a+\gamma_H)^2}{2} - \gamma_H \overline{x}_i^*$. For a demotivated agent i, equilibrium effort is $x_i^* = a + \gamma_L$, and equilibrium utility $u_i^* = \frac{(a+\gamma_L)^2}{2} - \gamma_L \overline{x}_i^*$.

Then, for a motivated agent i: $\frac{\partial u_i^*}{\partial \gamma_L} = -\gamma_H(1 - \rho_i^1(\mathbf{e}^*))$. For a demotivated agent i, $\frac{\partial u_i^*}{\partial \gamma_L} = -(\gamma_H - \gamma_L)\rho_i^1(\mathbf{e}^*) - \gamma_L(1 - \rho_i^1(\mathbf{e}^*))$. The proposition follows directly.

[Proof of Proposition 8] Assume $0 \leq \gamma_L < \gamma_H$ and consider an equilibrium $\mathbf{e}^*$. For a motivated agent i, equilibrium effort is $x_i^* = a + \gamma_H$, and equilibrium utility $u_i^* = \frac{(a+\gamma_H)^2}{2} - \gamma_H \overline{x}_i^*$. For a demotivated agent i, equilibrium effort is $x_i^* = a + \gamma_L$, and equilibrium utility $u_i^* = \frac{(a+\gamma_L)^2}{2} - \gamma_L \overline{x}_i^*$.

Then, for a motivated agent i: $\frac{\partial u_i^*}{\partial \gamma_H} = (\gamma_H - \gamma_L)(1 - \rho_i^1(\mathbf{e}^*)) - \gamma_H \rho_i^1(\mathbf{e}^*)$. For a demotivated agent i, $\frac{\partial u_i^*}{\partial \gamma_H} = -\gamma_L \rho_i^1(\mathbf{e}^*)$. The proposition follows directly.

[Proof of Proposition 9] Deriving (5) with respect to γ_H, we find

$$\frac{\partial W(\mathbf{e}^*; \mathbf{G})}{\partial \gamma_H} = (\gamma_H - \gamma_L)e^* - \gamma_L \sum_i e_i^* h_i^0(\mathbf{e}^*) - (2\gamma_H - \gamma_L) \sum_i e_i^* h_i^1(\mathbf{e}^*)$$

Or equivalently, recalling $h_i^0(\mathbf{e}^*) + h_i^1(\mathbf{e}^*) = h_i(\mathbf{e}^*)$,

$$\frac{\partial W(\mathbf{e}^*; \mathbf{G})}{\partial \gamma_H} = (\gamma_H - \gamma_L) \sum_i e_i^*(1 - 2h_i^1(\mathbf{e}^*)) - \gamma_L \sum_i e_i^* h_i(\mathbf{e}^*)$$

meaning

$$\frac{\partial W(\mathbf{e}^*;\mathbf{G})}{\partial \gamma_H} > 0$$

if and only if

$$(\gamma_H - \gamma_L)\sum_i e_i^*(1-2h_i^1(\mathbf{e}^*)) > \gamma_L \sum_i e_i^* h_i(\mathbf{e}^*)$$

I.e., noticing that $\sum_i e_i^* h_i^1(\mathbf{e}^*) = \sum_i e_i^* \rho_i^1(\mathbf{e}^*)$ and $\sum_i e_i^* h_i(\mathbf{e}^*) = \sum_i \rho_i^1(\mathbf{e}^*)$,

$$\sum_i e_i^*(1-2\rho_i^1(\mathbf{e}^*)) > \frac{\gamma_L}{\gamma_H - \gamma_L}\sum_i \rho_i^1(\mathbf{e}^*)$$

[Proof of Proposition 10] By the property that two equilibria cannot be nested, there is no two distinct equilibria being nested. It is then sufficient to show the proposition under $V = 0$ (i.e. the most favorable case for demotivated agents). That none of the two equilibria can Pareto-dominate the other is then a direct implication of the following lemma:

Lemma 1 *Assume $a_i = a$ for all i, and $V = 0$. For any two distinct equilibria $\boldsymbol{e}^*, \boldsymbol{e}^{'*}$, and any pair of agents (i,j) such that $e_i^* = 1, e_j^{'*} = 0$, $u_i(\boldsymbol{e}^*) > u_j(\boldsymbol{e}^{'*})$.*

Proof of Lemma 1. Consider any equilibrium with agent i being motivated, and any equilibrium with agent j being demotivated. The respective equilibrium utilities are written after few computation as

$$\begin{cases} u_i(\mathbf{x}^*) = \frac{(a+\gamma_H)^2}{2} - \gamma_H \overline{x}_i \\ u_j(\mathbf{x}^{'*}) = \frac{(a+\gamma_L)^2}{2} - ((1-\beta)\gamma_H + \beta\gamma_L)\overline{x}_j \end{cases}$$

Hence, $u_i(\mathbf{e}^*) > u_j(\mathbf{e}^{'*})$ whenever

$$(\gamma_H - \gamma_L)\Big(a + \frac{\gamma_H + \gamma_L}{2}\Big) > \gamma_H \overline{x}_i - ((1-\beta)\gamma_H + \beta\gamma_L)\overline{x}_j \tag{9}$$

By the property of equilibrium, we have

$$\overline{x}_i \leq \varphi < \overline{x}_j$$

A sufficient condition to get (9) is when $\overline{x}_i = \overline{x}_j = \varphi$ (since these conditions minimize the RHS of the inequality). This means

$$(\gamma_H - \gamma_L)\Big(a + \frac{\gamma_H + \gamma_L}{2}\Big) > \beta(\gamma_H - \gamma_L)\varphi$$

i.e.,

$$\varphi < \frac{a + \frac{\gamma_H + \gamma_L}{2}}{\beta} = \varphi$$

a contradiction.

[Proof of Theorem 4] This proof is presented for $\beta \leq 1$, including the case $\beta = 1$. This game admits the following best-responses:

$$\begin{cases} x_i^{BR}(\overline{x}_i) = a + \gamma_H + \delta d_i \overline{x}_i \text{ if } x_i \geq \beta \overline{x}_i \\ x_i^{BR}(\overline{x}_i) = a + \gamma_L + \delta d_i \overline{x}_i \text{ if } x_i \leq \beta \overline{x}_i \end{cases}$$

Suppose first that $\beta \leq \delta d_i$. Then $a + \gamma_H + \delta d_i \overline{x}_i \geq (1 - \beta)\overline{x}_i$ whenever

$$-(a + \gamma_H) \leq (\delta d_i - 1 + \beta)\overline{x}_i$$

which is true. Hence, agent i's best-response does not contain the demotivation region.

Second, Suppose that $1 - \beta > \delta d_i$. Then $a + \gamma_H + \delta d_i \overline{x}_i \geq \beta \overline{x}_i$ whenever $\overline{x}_i \leq y^1 = \frac{a+\gamma_H}{\beta - \delta d_i}$. Similarly, $a + \gamma_L + \delta d_i \overline{x}_i \leq \beta \overline{x}_i$ whenever $\overline{x}_i \geq y^0 = \frac{a+\gamma_L}{\beta - \delta d_i}$. Since $y^0 < y^1$, we have to study the best-play in the interval $\overline{x}_i \in (y^0, y^1)$. Let us define

$$\varphi_i = \frac{a + \frac{\gamma_H + \gamma_L}{2}}{\beta - \delta d_i} \tag{10}$$

where the value φ_i stems from equating utilities at best-responses in both regions of non-demotivation and demotivation. Actually: $U_i(x_i^{\gamma_H}) \geq U_i(x_i^{\gamma_L})$ whenever

$$\frac{1}{2}(a+\gamma_H+\delta d_i\overline{x}_i)^2 - \gamma_H\overline{x}_i \geq \frac{1}{2}(a+\gamma_L+\delta d_i\overline{x}_i)^2 - ((1-\beta)\gamma_H+\beta\gamma_L)\overline{x}_i$$

That is, $U_i(x_i^{\gamma_H}) > U_i(x_i^{\gamma_L})$ if and only if $\overline{x}_i < \varphi_i$.

Agent i's best-response is then written as

$$x_i^{BR}(\overline{x}_i) = a+\gamma_L+(\gamma_H-\gamma_L)e_i+\delta d_i\overline{x}_i$$

with $e_i = 1$ if and only if $\overline{x}_i \leq \varphi_i$. Denoting $\mathbf{M} = (\mathbf{I}-\delta\mathbf{G})^{-1}$, we have

$$\mathbf{x} = (a+\gamma_L)\mathbf{M1} + (\gamma_H-\gamma_L)\mathbf{Me}$$

from which we deduce, denoting $\tilde{\mathbf{G}} = (\frac{g_{ij}}{d_i})$,

$$\overline{\mathbf{x}} = (a+\gamma_L)\tilde{\mathbf{G}}\mathbf{M1} + (\gamma_H-\gamma_L)\tilde{\mathbf{G}}\mathbf{Me}$$

That is, for agent i,

$$\overline{x}_i = (a+\gamma_L)\sum_j \frac{g_{ij}b_j}{d_i} + (\gamma_H-\gamma_L)\sum_j \frac{g_{ij}b_{\mathbf{e},j}}{d_i}$$

Now, $e_i = 1 \Leftrightarrow \overline{x}_i \leq \varphi_i$, that is, multiplying by δd_i both RHS and LHS,

$$(a+\gamma_L)\delta\sum_j g_{ij}b_j + (\gamma_H-\gamma_L)\delta\sum_j g_{ij}b_{\mathbf{e},j} \leq \delta d_i\varphi_i$$

Recalling that $\delta\mathbf{GM} = \mathbf{M}-\mathbf{I}$, and thus $\delta\mathbf{GM1} = \mathbf{b}-\mathbf{1}$ and $\delta\mathbf{GMe} = (\mathbf{M}-\mathbf{I})\mathbf{e}$, we find $e_i = 1$ if and only if

$$(a+\gamma_L)(b_i-1) + (\gamma_H-\gamma_L)[(\mathbf{M}-\mathbf{I})\mathbf{e}]_i \leq \delta d_i\varphi_i$$

or $e_i = 1$ if and only if

$$[(\mathbf{M}-\mathbf{I})\mathbf{e}]_i \leq \kappa_i \qquad (11)$$

with

$$\kappa_i = \frac{\delta d_i \varphi_i - (a + \gamma_L)(b_i - 1)}{\gamma_H - \gamma_L}$$

Plugging φ_i from (10), we get

$$\kappa_i = \frac{\delta d_i(\gamma_H - \gamma_L) + 2(a + \gamma_L)\beta}{2(\gamma_H - \gamma_L)(\beta - \delta d_i)} - \frac{a + \gamma_L}{\gamma_H - \gamma_L} b_i \tag{12}$$

Thus, κ_i is increasing in d_i, decreasing in b_i.

Denote $\mathbf{k} = (\kappa_i)_i$. Exploiting (11), the potential function is then:

$$P(\mathbf{e}) = \mathbf{k}^T \mathbf{e} - \frac{1}{2}\mathbf{e}^T(\mathbf{M} - \mathbf{I})\mathbf{e}$$

which guarantees existence.