

# Overconfidence and Team Coordination\*

Simon Gervais

*Fuqua School of Business*

*Duke University*

Itay Goldstein

*Fuqua School of Business*

*Duke University*

February 2004

---

\*We would like to thank Roland Bénabou, Alon Brav, Ken Kavajecz, Leonid Kogan, Sendhil Mullainathan, John Payne, Steve Ross, Steve Slezak, Eric Van den Steen and S. Viswanathan for their comments and suggestions. Also providing helpful comments and suggestions were seminar participants at MIT, the University of Cincinnati, Duke University, the Wharton School, and the 2003 Workshop on New Ideas and Open Issues in Corporate Finance held in Amsterdam. All remaining errors are of course the authors' responsibility.

**Correspondence address:** Simon Gervais, Fuqua School of Business, Duke University, Box 90120, Durham, NC 27708, (919) 660-7683. Itay Goldstein, Fuqua School of Business, Duke University, Box 90120, Durham, NC 27708, (919) 660-7858.

**email address:** sgervais@duke.edu and itayg@mail.duke.edu.

# Overconfidence and Team Coordination

## Abstract

We model a team in which the marginal productivity of a player increases with the effort of other players on the team. Because the effort of each player is not observable to any other player, the performance of the team is negatively affected by a free-rider problem and by a lack of effort coordination across players. In this context, an overconfident player who overestimates her own marginal productivity works harder, thereby increasing the marginal productivity of her teammates who then work harder as well. This not only enhances team performance but may also create a Pareto improvement at the individual level. Indeed, although the overconfident player overworks, she benefits from the positive externality that other players working harder generates. Interestingly, the benefits of overconfidence may be long-lived even if players learn from team performance, as the overconfident player attributes the team's success to her own ability, and not to the better coordination of the team. Because overconfidence naturally makes players work harder, monitoring, even when it is costless, may hurt the team by causing an overinvestment in effort.

## 1. Introduction

It is well known that moral hazard problems are prevalent in teams when the effort decisions of the teams' agents are unobservable. Because agents make decisions that are in their best self-interest, their unmonitored actions often fail to conform to their organization's objectives, unless proper incentives are provided to them. As pointed out by Groves (1973) and by Holmström (1982), the absence of such incentives leads to lost value through mis-communication, free-riding behavior, and general lack of coordination across team members. These problems are exacerbated when externalities exist across the team's agents, as any one agent does not fully internalize the impact that her decisions have on the decisions of others. Starting with Groves and Holmström, several contracting solutions have been proposed for properly motivating individuals in team contexts. For example, Rasmusen (1987), Itoh (1991), McAfee and McMillan (1991), Vander Veen (1995), Faulí-Oller and Giralt (1995), and Andolfatto and Nosal (1997) study variations of the original solution developed by Holmström that account for risk aversion, monitoring, and various types of externalities between the team's agents. Common to all these papers is the search for the link between compensation and joint output that best fosters effort.

In this paper, we approach team scenarios from a different perspective, namely that of psychology. A large body of the psychology literature shows that individuals tend to overestimate their own skills. For example, Langer and Roth (1975), Weinstein (1980), and Taylor and Brown (1988) document that individuals tend to perceive themselves as having more ability than is warranted. Similarly, Fischhoff, Slovic and Lichtenstein (1977), and Alpert and Raiffa (1982) find that individuals tend to overestimate the precision of their information. We incorporate such behavioral biases, which we collectively refer to as *overconfidence*, into the team problem by assuming that some players overestimate the marginal product of their effort. We show that this overconfidence not only overcomes the free-riding and coordination problems in teams, but can also make all team members, including the overconfident ones, better off.

The idea is that agents who overestimate their own marginal product tend to work harder. In particular, an overconfident agent can sometimes justify making a costly effort when an otherwise identical but rational agent would not. This extra effort reduces the free-rider problem quite naturally, but it does more than that when complementarities exist across team members. Since the effort of one agent increases the marginal productivity of other agents, they too find themselves facing a situation in which their effort is more valuable. As a result, these other agents also exert more effort, making the team even more productive. When her overconfidence is not too

extreme, even the biased agent ends up benefitting from her overinvestment in effort, as she shares the benefits of her teammates' increased effort (but still suffers the cost of her overinvestment in effort).

Other authors have also imported behavioral considerations into team contexts. For example, Rotemberg (1994) analyzes the effect of altruism on coordination in teams.<sup>1</sup> He shows that when complementarities between the team's agents exist, the presence of some altruistic agents can generate Pareto improvements, just like altruism can benefit all members of a family (not just the selfish ones), as argued by Becker (1974). Eshel, Samuelson and Shaked (1998) further show that altruistic teams are more likely to survive in the long run.<sup>2</sup> Another example of behavioral considerations is found in the work of Kandel and Lazear (1992), who show that team coordination problems can be overcome when there is peer pressure among members of the team. In effect, peer pressure imposes an extra cost on agents that do not make the appropriate effort. These authors also discuss how peer pressure can emerge endogenously. Ferreira (2002) combines the effects of peer pressure and altruism to study group loyalty.

Interestingly, it is not the concerns for others or of others that solve coordination problems in our model. Instead, it is the extreme self-perception of some agents that does. Overconfident agents simply think that their contribution is large enough to justify their costly effort, without any consideration for their teammates. The externalities associated with their effort matters little to overconfident agents but does foster cooperation within the team. That is, their flattering views of themselves combine with their self-interest to generate externalities on others. So agents cooperate not because they want to, but because cooperation comes with being skilled (as they think they are) and working.

Because overconfident agents think that their contribution to team output is larger than it really is, they also misinterpret the eventual larger output of the team. As we show, they attribute it to their own skill more than to the effort of their teammates. If agents learn their abilities through the realized performance of their team, this self-attribution bias slows down the learning of their true ability, making the benefits of overconfidence longer-lasting. In other words, we expect these agents' teams to survive the market test, and we expect the individuals on these teams,

---

<sup>1</sup>The idea that altruism affects people's economic decisions dates back to at least Adam Smith (1759; 1976), who suggests that human nature is such that individuals care about others' happiness even if they don't directly benefit from it.

<sup>2</sup>In fact, Dawes and Thaler (1988) argue that the presence of cooperation within teams and organizations is a sign that individuals are not as selfish as they are usually made out to be in economic models.

including the overconfident ones, to thrive in the long run. This possibility that individuals can be made better off in the long run by their biased perceptions through the effect they have on the actions of others has also been demonstrated by Heifetz and Spiegel (2001), and by Heifetz, Shannon and Spiegel (2002). In these papers, however, it is not possible for individuals to learn their biases away. Our results on learning, namely the fact that overconfidence is either slowly or never learned away, further guarantees the survival of individuals with biased self-perceptions. Indeed, individuals with overconfidence will tend to survive in the early rounds and, in the process, will not learn their overconfidence, making their long-run survival possible. Van den Steen (2002) also studies situations in which agents who have a self-serving bias tend to learn slowly. In his model, agents with differing priors endogenously attribute success to their own skills and failure to bad luck. What goes on in our model is different in that learning is slowed down by the fact that the bias does increase team output (through better coordination), but not for the reason the overconfident agent thinks (her high ability).

In many team contexts, perfect free monitoring restores first-best. This is the case in our benchmark model without overconfident agents. However, as we show, the seemingly obvious benefit of free monitoring can disappear when some agents have a biased view of their own skills. In particular, when monitored, these agents tend to overwork, thereby reducing their welfare and the value of their team. Thus we conclude that a team or firm whose output results from the interactions of several agents will want to correctly balance the extent of its monitoring with the characteristics of these agents. Also, because monitoring and overconfidence can be substituted for each other, picking individuals with useful behavioral biases, like overconfidence, becomes quite valuable for the firm when monitoring is costly.

The rest of the paper is organized as follows. In section 2, we set up the two-agent team framework that is used throughout the paper, and highlight the coordination problems that arise in it. Section 3 introduces the concept of overconfidence, and shows how it can naturally help solve the team's coordination problems by facilitating effort. The same section goes on to show that, in the presence of complementarities across agents, the overconfident agent's overinvestment in effort may not only benefit her team and teammates, but also herself. The possibility that an agent's overconfidence changes as she learns her skills through the team's output is considered in section 4. Section 5 looks at the joint roles of overconfidence and monitoring in the team context, and shows that they are substitutes in the sense that the presence of one may render the other detrimental. Alternative interpretations and applications of our model are discussed in section 6. Section 7 offers

some final remarks and concludes. All proofs are contained in the appendix.

## 2. The Basic Framework

### A. A Partnership Model

Our model has one firm owned by two agents, each of which has a claim to half of the firm's value. We also refer to this arrangement as a *team* or *partnership*. The value of the firm comes from a single one-period project, which can either succeed or fail with probabilities  $\pi$  and  $1 - \pi$  respectively. The project generates 2 dollars at the end of the period if it succeeds, and it generates 0 if it fails. Thus the firm's end-of-period cash flow is given by

$$\tilde{v} = \begin{cases} 2 & \text{prob. } \pi \\ 0 & \text{prob. } 1 - \pi. \end{cases} \quad (1)$$

The probability of success  $\pi$  is endogenous; it depends on the choice of effort made by both agents. Each agent  $i$  can choose to work ( $e_i = 1$ ) or not ( $e_i = 0$ ). We assume that

$$\pi = ae_1 + ae_2 + be_1e_2, \quad (2)$$

where  $a$  and  $b$  are non-negative constants. Parameter  $a$  measures the direct effect of an agent's effort on the probability of success. It can be interpreted as the ability level of the agents. Parameter  $b$  captures the effect of the interaction between the two agents on the probability of success. In assuming that  $b \geq 0$ , we are considering a situation in which the interaction is synergistic, that is, the two agents create positive externalities on each other. Indeed, when one agent works, the marginal product of the other agent's effort (i.e., the impact her effort has on the probability of success) increases: it goes from  $a$  to  $a + b$ . The assumption that  $b$  is positive is consistent with Alchian and Demsetz's (1972) views that teams (or firms) form to take advantage of positive externalities or complementarities. Of course, since  $\pi$  is a probability, we need to ensure that it is between zero and one, and so we impose the following restriction on  $a$  and  $b$ :

$$0 \leq 2a + b \leq 1. \quad (3)$$

Agents choose their effort to maximize their expected utility. We assume that both agents are risk-neutral, and that they each bear a private cost of effort. We denote the effort cost of agent  $i$  by  $\tilde{c}_i$ , so that the utility of agent  $i$  at the end of the period is

$$\tilde{U}_i = \frac{1}{2}\tilde{v} - \tilde{c}_ie_i. \quad (4)$$

Effort costs are not known by anyone at the outset, but are known to be uniformly distributed between 0 and 1, and independent across agents.<sup>3</sup> Instead, each agent privately observes her own cost, without observing the other's, before making her effort decision. This describes, for example, a situation in which agents learn the constraints they face (e.g., time, other commitments, etc.) after committing to the partnership, while not being able to infer the constraints of others. Effort decisions are made simultaneously by the two agents, and each agent's decision is unobservable to the other agent, making effort decisions non-contractible.

### *B. Equilibrium in a Benchmark Model*

At the time each agent makes her effort decision, she does not know whether the other agent will exert effort or even the cost of that effort. Instead, **in equilibrium, she must anticipate the expected level of effort from the other agent.** In equilibrium, because utility is decreasing in effort cost, it will be the case that agent  $i$  works if and only if her cost of effort does not exceed some threshold that we denote by  $k_i \in [0, 1]$ . That is, if it is optimal for an agent to work when the cost of effort is  $\tilde{c}_i = k_i$ , then she will also find it optimal to work when  $\tilde{c}_i < k_i$ . Solving for the equilibrium involves finding the equilibrium  $k_i$  for each agent.

Let us take the position of the first agent, after she observes that her effort will cost  $\tilde{c}_1 = c_1$ . She anticipates the second agent to work if  $\tilde{c}_2 \leq k_2$ , and so she anticipates her to work with probability  $k_2$ . Thus agent 1 seeks to solve the following maximization problem:

$$\begin{aligned} \max_{e_1 \in \{0,1\}} \mathbb{E} \left[ \tilde{U}_1 \mid \tilde{c}_1 = c_1 \right] &= \mathbb{E} [\pi] - c_1 e_1 \\ &= a e_1 + (a + b e_1) \mathbb{E} [e_2] - c_1 e_1 \\ &= a e_1 + (a + b e_1) k_2 - c_1 e_1. \end{aligned} \tag{5}$$

From this, it is easy to show that agent 1 works ( $e_1 = 1$ ) if and only if  $\tilde{c}_1 \leq a + b k_2$ . Similarly, taking the position of the second agent, we find that  $e_2 = 1$  if and only if  $\tilde{c}_2 \leq a + b k_1$ . Thus the thresholds in this benchmark equilibrium must satisfy

$$\begin{aligned} k_1 &= a + b k_2, \quad \text{and} \\ k_2 &= a + b k_1. \end{aligned}$$

---

<sup>3</sup>These distributional assumptions about  $\tilde{c}_i$  are made purely for convenience. The only required assumption is that effort costs are not perfectly correlated.

Solving for  $k_1$  and  $k_2$  in these equations, we find

$$k_1 = k_2 = \frac{a}{1-b} \equiv k_{\text{BM}}. \quad (6)$$

We are ultimately interested in the welfare of the team's agents, that is their expected utility at the time the partnership is formed (i.e., before effort costs are observed and effort choices are made). Given effort cost thresholds of  $k_1$  and  $k_2$ , one can use (5) (and the similar maximization problem for agent 2) to find

$$\bar{U}_1 \equiv \mathbb{E} [\tilde{U}_1] = ak_1 + ak_2 + bk_1k_2 - \int_0^{k_1} c_1 dc_1 = ak_1 + ak_2 + bk_1k_2 - \frac{k_1^2}{2} \quad (7)$$

and

$$\bar{U}_2 \equiv \mathbb{E} [\tilde{U}_2] = ak_1 + ak_2 + bk_1k_2 - \frac{k_2^2}{2}. \quad (8)$$

So each agent expects  $a$  when she works,  $a$  when the other works, and  $b$  when they both work. The effort cost that each expects to incur is the last term in the above expressions: agent  $i$  works with probability  $k_i$  and incurs an average cost of  $\frac{k_i}{2}$  when that is the case, for an expected cost of  $\frac{k_i^2}{2}$ . In the benchmark equilibrium,  $k_1 = k_2 = k_{\text{BM}}$  and so both agents' expected utility is given by

$$\bar{U}_{\text{BM}} \equiv 2ak_{\text{BM}} + \left(b - \frac{1}{2}\right)k_{\text{BM}}^2 = \frac{a^2 \left(\frac{3}{2} - b\right)}{(1-b)^2}. \quad (9)$$

From this expression, it is easy to show that both agents are better off when  $a$  and  $b$  are larger. This makes sense, as increasing both of these parameters increases the impact of effort.

### C. First-Best Allocation

Before proceeding further, we are interested in the first-best allocation of effort, that is, the effort allocation that a social planner would pick in order to maximize the welfare of the team's agents. More specifically, we are interested in determining the cost thresholds that this social planner would impose on the two agents, assuming that these thresholds are chosen ex ante, before agents observe their effort costs. Because the two agents are identical, the first-best thresholds will satisfy  $k_1 = k_2 = k_{\text{FB}}$  and will maximize (7) and (8). We find that the interior solution to this problem is given by

$$k_{\text{FB}} = \frac{a}{\frac{1}{2} - b} \quad (10)$$

as long as  $a + b < \frac{1}{2}$ , which we assume from now on for convenience.<sup>4</sup>

---

<sup>4</sup>Otherwise, the corner solution given by  $k_{\text{FB}} = 1$  unnecessarily complicates the analysis.



Clearly,  $k_{\text{FB}} > k_{\text{BM}}$ . That is, agents do not exert enough effort in the equilibrium of the benchmark model.<sup>5</sup> This happens for two reasons. First, because agents receive only half of the product of their effort but have to bear the full cost of that effort, they tend to free-ride on the effort of others. This is a standard problem in teams, as pointed out and studied by Holmström (1982). Second, in our model, agents do not fully internalize the complementarity effect that their effort has on the effort of others. This effect gets stronger as  $b$  increases and, indeed, one can verify that the difference between  $k_{\text{FB}}$  and  $k_{\text{BM}}$  is increasing in  $b$ .

In this partnership, therefore, both agents would benefit from committing to higher levels of effort (i.e., higher  $k_1$  and  $k_2$ ). Because effort cost and effort are unobservable and non-contractible, however, it seems a priori impossible for the agents to resolve their coordination problem without involving a third-party.

### 3. Overconfidence

In this section, we propose an avenue for mitigating the coordination problem faced by the team's agents. Our solution centers around the argument that the types of agents forming teams in the first place will naturally face a lesser coordination problem.

#### A. Introducing Overconfidence

Suppose that agent 2 is overconfident about her ability. Specifically, she thinks her ability is  $A > a$ , although it is truly only  $a$ . For now, we do not discuss how and why this agent became biased about her ability, except that a large body of psychology literature documents that individuals tend to perceive themselves as better than they actually are.<sup>6</sup> Later in the paper, when we introduce a learning component to the model, we also argue that such biases will naturally persist at both the firm and the individual levels. For now, all the other details of the model remain the same as in the previous section.<sup>7</sup> Since it is the departure from real ability to perceived ability that represents agent 2's bias, we denote her level of overconfidence by  $d \equiv A - a$ .

<sup>5</sup>Of course, because effort is zero or one, this is the same as saying that agents do not exert effort often enough in the equilibrium of the benchmark model.

<sup>6</sup>In addition to the literature we mention in the introduction, Camerer and Lovo (1999) perform an experiment showing that overconfidence naturally causes individuals to enter games and markets.

<sup>7</sup>We do need to impose the restriction that  $0 \leq A + a + b \leq 1$ , so that the probability of a success, as perceived by an overconfident agent, is between zero and one.

We assume that agent 1 knows that agent 2 is overconfident. As we point out later, this assumption is important for some, but not all, of our results. In particular, it does affect our welfare analysis as it pertains to agent 2. This is because our welfare results depend on whether other agents change their behavior when teamed with an overconfident agent. We also assume that agent 2 does not think that agent 1 recognizes her superior ability, and instead thinks that agent 1 perceives her ability to be  $a$ . This latter assumption implies that the overconfident agent is convinced that her ability is higher, but believes that no one else realizes that. This assumption is more harmless than the previous one as only the version of the model that incorporates learning, studied in section 4, is affected by it. Until that section, all our results hold under the alternative assumption that agent 2 thinks that agent 1 recognizes her superior ability. We do not like this alternative assumption as much, though, as it seems unlikely that such a team would be able to negotiate a 50-50 split of the firm's value; indeed, the overconfident agent would argue that she should get more than half the value, as she (thinks she) contributes more to it than the other agent. Under our current assumption, the overconfident agent knows that she simply cannot convince others that she is more skilled than they are, and so agrees to join the team for an equal share of its value. Because we do not model this ex ante negotiation (this would be an interesting problem of its own), the use of either assumption is probably equally valid.

### B. Equilibrium

To find the equilibrium, we proceed as in section 2. However, the equilibrium strategies of the two agents are now slightly more complex, because an agent's true strategy is not necessarily the same as that perceived by her teammate. We use  $k_{ij}$  to denote the threshold used by agent  $j$  as perceived by agent  $i$ . So the actual thresholds used by agents 1 and 2 are  $k_{11}$  and  $k_{22}$  respectively, but they may be perceived to be  $k_{21}$  and  $k_{12}$  by agents 2 and 1.<sup>8</sup> Using the same reasoning as in the benchmark model of section 2, we can derive the equilibrium strategies for the two agents, when agent 2 is overconfident.

**Lemma 1** *Suppose that agent 2 is overconfident, but not agent 1. In equilibrium,*

- (i) *agent 1 makes an effort if and only if her cost of effort does not exceed*

$$k_{11} = k_{\text{BM}} + bd;$$

---

<sup>8</sup>In section 2, we have  $k_1 = k_{11} = k_{21}$  and  $k_2 = k_{22} = k_{12}$ .

(ii) *agent 2 makes an effort if and only if her cost of effort does not exceed*

$$k_{22} = k_{\text{BM}} + d.$$

Notice that, when  $d > 0$ , both agents work harder than in the benchmark scenario. In fact, their effort is strictly increasing in  $d$ . This effect is rather intuitive for agent 2. As the perception of her own ability increases, her own perceived productivity increases. From her perspective, this increased productivity is enough to warrant an effort, that is, her effort does not require as much of an effort on the part of agent 1 as before.

More interesting is the fact that agent 1 also works harder as  $d$  increases. This is due to the fact that, because agent 1 knows that agent 2 works harder, she knows that the potential synergistic gains, through  $b$ , from their combined effort is likely larger than before. This makes her effort more valuable, and so she is more willing to pay its cost. In other words, when the efforts of the teammates are complementary, the marginal productivity of one increases in the other one's effort, and so the higher effort of one increases the effort of the other. Of course, if  $b$  were negative or even zero, this result would disappear. This may be an avenue for potential tests of our model. Indeed, later in the paper, we argue that the increase in effort due to overconfidence should make it more likely for the firm to succeed and for overconfidence to persist. If complementarities are necessary for this to occur, then we should observe more overconfident individuals working in industries that are naturally more prone to synergies among workers.

An interesting aspect of Lemma 1 is the fact that agent 2 believes that agent 1's equilibrium strategy is characterized by a threshold of  $k_{21} < k_{11}$ . That is, she does not know that agent 1 works as hard as she does. This misperception does not have much of an impact here, but will have an important effect on learning later in the paper. Indeed, because of this misperception, agent 2 will tend to attribute the success of her team to her own skills, and so will tend to remain overconfident. We will come back to this issue in section 4.

### *C. Overconfidence and Individual Welfare*

Because, as discussed in section 2, both agents would benefit from committing to working harder in the benchmark scenario, it immediately follows from Lemma 1 that the presence of some overconfidence is always welfare increasing. Indeed, as  $d$  increases from zero to a small but positive value, both agents work slightly harder, and so both enjoy higher expected utility. This is described in more details in the following proposition. Before we turn to this result, however, note that the

welfare of agent 2 can be assessed from two perspectives. First, we could calculate her expected utility as she perceives it ex ante, that is, assuming that  $A$  is really her marginal contribution to the project's success. This, we think, is uninteresting as agent 2 will not experience this utility on average ex post. A more useful perspective is using  $a$  as her correct ability, but taking into account the fact that she and her teammate pick effort thresholds that are different from those in the benchmark scenario. This is a better measure of how agent 2 will feel, on average, at the end of the period. We also think that this measure of “average ex post utility” is more likely to drive the individuals' decisions as to whether they stay or leave a firm, although we do not consider these issues per se in our model. As such, in the following proposition and in the rest of the paper, when referring to the expected utility of the overconfident agent, we refer to this measure unless we mention otherwise.

**Proposition 1** *Suppose that agent 2 is overconfident, but not agent 1. For the equilibrium described in Lemma 1,*

- (i) *the expected utility of agent 1 is always increasing in  $d$ ;*
- (ii) *the expected utility of agent 2 is increasing in  $d$  if and only if*

$$d < \frac{ab}{(1-b)(1-2b^2)}. \quad (11)$$

The first part of the proposition shows that an increase in the level of overconfidence of agent 2 always improves the welfare of agent 1. This is not surprising. When agent 2 thinks she has a higher ability, she works harder. Since agent 1 shares the output of agent 2's effort but does not share the cost, her ex ante welfare is always increased when agent 2 works harder.

The second part of the proposition shows that agent 2 is made better off by an increase in her own overconfidence, as long as her overconfidence level is not too large, i.e., as long as (11) is satisfied. This result is more interesting. It means that agent 2 ends up benefitting from her own misperceptions, as long as these misperceptions are not too extreme. Intuitively, this result comes from the tradeoff between agent 2's overinvestment in effort and the synergistic feedback effect of agent 1's increased effort. More precisely, even though agent 2 does not properly choose her own effort given her true ability, her cost of effort, and the level of effort of agent 1 (this is the cost), she benefits from the fact that agent 1 works harder as a response to her increased effort (this is the benefit). A marginal increase in  $k_{22}$  when it is small (i.e., close to  $k_{BM}$ ) creates a synergistic gain that more than outweighs the increased cost of effort. When  $d$  (and  $k_{22}$ ) gets

larger however, the marginal cost of effort becomes larger,<sup>9</sup> and agent 2 ends up hurting herself through her effort decisions. Notice also that the right-hand side of (11) is increasing in both  $a$  and  $b$ . As the (actual) marginal productivity and complementarities of the two agents increase, the larger effort cost associated with the overconfidence of agent 2 becomes more worthwhile.

Taken together, the two parts of Proposition 1 imply that the overconfidence of agent 2 creates a Pareto improvement for the team when (11) holds. This Pareto dominance result is where our paper differs from most of the behavioral economics literature, which describes how biased agents fail to realize the full value of their opportunities while others benefit from their mistakes. Behavioral biases not only affect the decisions of the biased agents, but they also affect the decisions of the agents they interact with. Our model shows that, in the presence of positive externalities across agents (synergistic product of effort, in this case), these changes in their decision-making (choice of effort, in this case) may have a positive effect on everyone.<sup>10</sup>

A different perspective on the result is offered in Figure 1, which shows the iso-utility curves for each of the two agents in this model. The two curved lines show the set of effort cost threshold combinations,  $k_{11}$  and  $k_{22}$ , that leave the two agents as well off as in the benchmark scenario (which has  $k_{11} = k_{22} = k_{\text{BM}}$ ). Every point above (to the right of) the dashed (continuous) curve makes agent 1 (agent 2) better off, and so the shaded region represents all the Pareto-improving sets of thresholds; in particular,  $k_{11} = k_{22} = k_{\text{FB}}$  is included in that region. It is straightforward to show that, as  $d$  is increased from zero, the two agents change their equilibrium thresholds along the straight line with the arrow. Clearly, because this straight line lies above the dashed curve, agent 1 is better off. More interesting is the fact that the line starts inside the Pareto-improving region but eventually comes out of it: agent 2 is better off when she is slightly overconfident, but not better off when her overconfidence is too extreme.

The key ingredient for the result is the presence of complementarities across agents. Mathematically, this can be seen from (11), whose right-hand side is strictly positive when  $b > 0$ , implying that the condition is always satisfied for small values of  $d$ . In fact, the right-hand side of condition (11) is strictly increasing in  $b$ , implying that overconfidence is more likely to help when complementar-

---

<sup>9</sup>To be precise, the marginal effect on average effort cost from an increase in  $k_{22}$  is  $\frac{\partial}{\partial k_{22}} \left( \frac{k_{22}^2}{2} \right) = k_{22}$ .

<sup>10</sup>In fact, it is straightforward to show that the expected end-of-period payoff of the firm is also larger, so that firm value is also improved with overconfidence. So it seems possible, or even likely, that even passive shareholders would benefit from agent overconfidence if we were to include them in the economy/model. This last result is similar to those derived by Heaton (2002) and by Gervais, Heaton and Odean (2003). In these models, however, biased managers are not made better off by their biases; only the firm and its principals (owners) benefit.

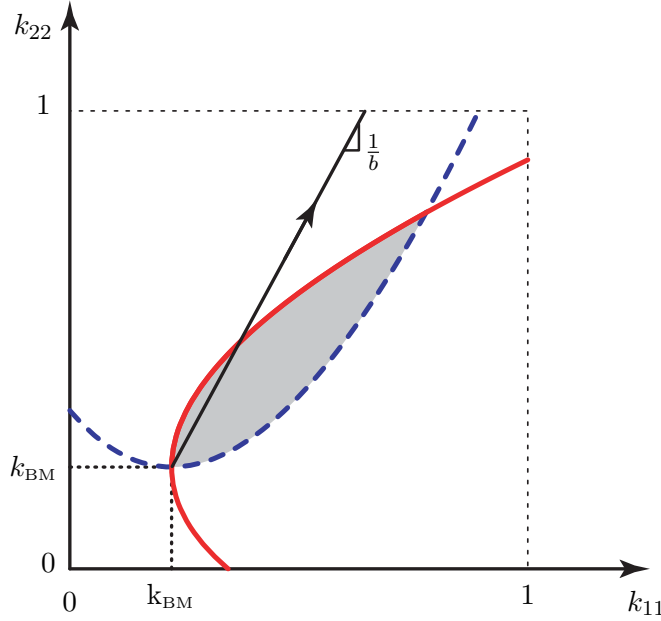


Figure 1: Iso-utility curves and agent welfare.

In the benchmark scenario, both agents use an effort cost threshold of  $k_{\text{BM}}$ . The curved dashed (continuous) line shows the set of thresholds  $(k_{11}, k_{22})$  for the two agents that keep agent 1 (agent 2) equally well off. The shaded region shows the set of thresholds that make both agents better off. Making agent 2 more and more overconfident by increasing  $d$  increases the equilibrium thresholds used by the two agents along the line with the arrow. For small levels of  $d$ , the two agents are better off.

ities are stronger. The fact that complementarities are essential for our Pareto-dominance result can also be seen graphically. In Figure 1, the slope of the straight line can be shown to be  $\frac{1}{b}$ . Because the continuous curve representing the iso-utility points of agent 2 has an infinite slope at  $k_{11} = k_{22} = k_{\text{BM}}$  (this is a simple application of the envelope theorem), it will always be the case that the equilibrium thresholds resulting from small increases in  $d$  will lie inside the figure's shaded region. Intuitively speaking, the presence of complementarities is necessary because the behavior of agent 1 has to be affected by the overconfidence of agent 2. In particular, it has to be the case that agent 1 is induced to work harder as a result of agent 2's bias. This happens precisely when  $b$  is greater than zero. In fact, although we do not consider the possibility that the agents' efforts are substitutes (i.e.,  $b < 0$ ) in this paper, it is easy to see that overconfidence would only improve the utility of agent 1 in that case.

Bénabou and Tirole (2002) show how some behavioral biases can enhance personal motivation and welfare. In their work, the individual is studied in isolation: self-deception improves welfare when the motivation gains from ignoring negative signals outweigh the losses from ignoring positive ones. In contrast, our model revolves around the interactions of biased individuals with others. In particular, the gains from the biased decisions of some individuals (their mis-allocation of effort) are not the result of improved self-motivation. Instead, they come from the effect they have on the motivation of others. In a related paper, Bénabou and Tirole (2003) study the role of motivation in a decision setting involving two individuals. However, the emphasis of their work is different from ours, as they concentrate on the role played by ego-bashing when private benefits are associated with the adoption of one's idea.

#### *D. Overconfidence and Team Welfare*

In our model, the sharing rule between the two teammates is prescribed: they each get half the team's output. For this reason, the firm's output is very much like a common good whose value depends on the combined effort of two individuals. In fact, some of the model's applications that we discuss in section 6 rely on this interpretation. An alternative interpretation is possible if we view the team's output as the profit of a stand-alone firm, whose labor input consists of the effort of two individuals hired by the firm's owner. In this context, if we assume that the firm captures all of the surplus resulting from the contractual relationship with its employees (i.e., if we assume that labor markets are competitive and employees receive their reservation salary), then, because both employees are risk-neutral, firm value will be given by  $\bar{U}_1 + \bar{U}_2$ . This quantity, which we refer to as *team welfare*, is studied in the following proposition.

**Proposition 2** *Suppose that agent 2 is overconfident, but not agent 1. For the equilibrium described in Lemma 1, team welfare is increasing in  $d$  if and only if*

$$d < \frac{a(1+b)}{(1-b)(1-3b^2)}. \quad (12)$$

Just like the overconfident agent's welfare in Proposition 1, team welfare is increasing in the level of overconfidence as long as overconfidence is not too extreme, i.e., as long as (12) is satisfied. Of course, condition (12) is implied by (11): if overconfidence increases the welfare of each agent, it trivially increases team welfare. However, it is possible for team welfare to increase with  $d$  even though the overconfident agent is made worse off by his greater bias. This is illustrated in Figure 2

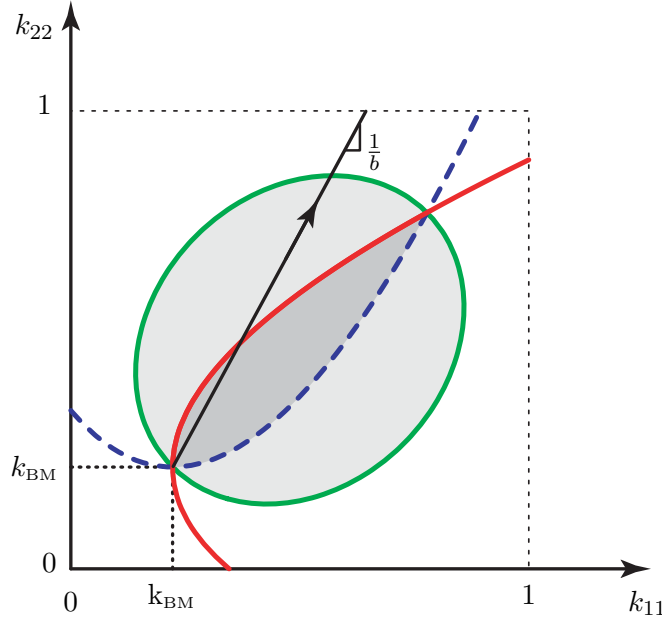


Figure 2: Iso-utility curves and team welfare.

In the benchmark scenario, both agents use an effort cost threshold of  $k_{\text{BM}}$ . The curved dashed (continuous) line shows the set of thresholds  $(k_{11}, k_{22})$  for the two agents that keep agent 1 (agent 2) equally well off. The dark shaded region shows the set of thresholds that make both agents better off. The light shaded region shows the set of thresholds that increase team welfare, the sum of the two agents' expected utilities.

which adds the set of effort thresholds that increase team welfare from its benchmark value to Figure 1. As before, overconfidence is more likely to be beneficial when complementarities between the agents are stronger (i.e., when  $b$  is large). However, the presence of complementarities is no longer necessary for overconfidence to be beneficial. Indeed, even when  $b$  is zero (or negative), the right-hand side of (12) is positive and the condition is satisfied for small values of  $d$ . This is because, even without complementarities, the overconfidence of agent 2 helps mitigate the free-rider problem by making agent 2 work harder. Since a more sustained effort is optimal from a social planner's perspective, this increases team welfare but, without surplus redistribution, reduces the welfare of agent 2.

To sum up, overconfidence plays two roles in our model. The first is to mitigate the free-rider problem: this increases the welfare of the rational agent and that of the team, but decreases the overconfident agent's. The second role of overconfidence is to induce internalization of effort externalities when  $b > 0$ : this makes both agents better off, and thus can generate Pareto improvements.



## 4. Learning

As shown in section 3, the presence of overconfident agents within a team makes that team more productive and its members better off. Thus it is likely that teams that include some overconfident agents will be better equipped to compete with other teams: the agents work harder, produce more, and make the team more valuable. Furthermore, because the presence of overconfidence has a Pareto-improving effect on its members, any individual member should be less tempted to leave and look for better opportunities elsewhere. As such, the team's composition is likely to remain intact, making the effect of overconfidence potentially long-lasting. Although we do not explicitly tackle the long-run survival prospects of the team in this paper, it is reasonable to expect that teams with some overconfident members are more likely to prosper over time.<sup>11</sup>

We think that a more challenging question is whether, as the team and its members prosper, overconfidence can sustain itself? In other words, if biased agents eventually learn their true skills, the benefits of overconfidence disappear, and the team starts suffering from the same coordination problems described in section 2.<sup>12</sup> In this section, we confront this issue by incorporating learning into the model. In particular, we now consider a situation where agents do not know their true ability when joining a team, but learn it based on the decisions they make and the outcomes of these decisions.

### *A. Unknown Ability and Updating*

We assume that neither agent knows her own skill at the outset. In particular, we assume that agent  $i$ 's skill,  $\tilde{s}_i$ , is uniformly and independently distributed over  $[0, 2a]$ . As before, agent 2 is biased when it comes to her own skill. More precisely, she believes that her own skill is uniformly distributed over  $[0, 2(a + d)]$ .<sup>13</sup> Thus, although each agent's average skill is really  $a$ , agent 2 thinks that her average skill is  $a + d$  where, as in section 3,  $d$  denotes the extent of her overconfidence. Because the rest of the model is unchanged and because everyone is risk-neutral, we can use these quantities in Lemma 1 to obtain the equilibrium strategies of the two agents. The only difference

---

<sup>11</sup>Intuitively, the analysis would show that teams with some overconfidence, and so better coordination, are more likely to come out as winners in industry tournaments.

<sup>12</sup>It is important to note that overconfidence, like risk aversion, cannot be faked. In fact, it is crucial that biased agents are unaware of their biases when making decisions, as it is this unawareness that affects their behavior.

<sup>13</sup>As before, we need to impose some constraints on the model's parameters in order to ensure that perceived probabilities are between zero and one. In this case, we need to assume that  $4a + 2d + b \leq 1$ .

from before is that agent 2's overconfidence will change if her beliefs about  $s_2$  change. In particular, if the team also faces a second-period coordination problem, the fact that agent 2 updates her beliefs about her own skills after observing the outcome of the first period changes her perceived expected ability for the second period.

Since her ability gets impounded into the probability of the project being successful only when she exerts an effort, agent 2 only updates about  $\tilde{s}_2$  when  $e_2 = 1$ . We denote the average belief that she reaches after exerting an effort by  $\alpha \equiv E_B[\tilde{s}_2 | e_2 = 1]$ , where the “B” subscript denotes the fact that agent 2 is biased. The law of iterated expectations tells us that this average belief should correspond exactly to the prior mean for a rational agent; as we show next, this is not the case for an overconfident agent.

**Proposition 3** *On average, the overconfidence of agent 2 decreases but remains positive at the end of the period, that is,  $a < \alpha < a + d$ .*

Agent 2 has two misconceptions in her assessment of the project's probability of success. The first, using  $A = a + d$  instead of  $a$  for her average ability, has a positive impact of  $d$  on this probability. The second, using  $k_{BM}$  instead of  $k_{BM} + bd$  for agent 1's effort cost threshold, has a negative impact of  $(a + b)bd$ . Because  $d > (a + b)bd$  however, it is always the case that agent 2 overestimates the probability that the project will succeed. So, on average, it will be the case that she will revise her beliefs downwards, that is, her overconfidence will decrease. However, her updated beliefs, after a successful or a failed project, are always above what they would otherwise be were she rational, and so some overconfidence always remains. As the following proposition shows, the extent of this ex post overconfidence is more extreme when the complementarities between the two agents are stronger.

**Proposition 4** *The end-of-period overconfidence of agent 2 is increasing in  $b$ , that is,  $\frac{\partial \alpha}{\partial b} > 0$ .*

As discussed above, the fact that  $d$  exceeds  $(a + b)bd$  is what makes agent 2 revise her beliefs towards her true ability. Notice, however, that the difference between these two quantities gets smaller as  $b$  increases. Indeed, when  $b$  is large, the rational agent works hard as a result of the complementarities that exist inside the team. Because agent 2 fails to fully account for this increased effort, she attributes the success of the team to her own skills. This slows down her learning.

It is interesting that the presence of complementarities allows overconfidence to make team members better off by facilitating coordination and, *at the same time*, makes convergence to ratio-

nality more difficult and thus slower. This makes overconfidence a good candidate as an ingredient for long-term team success.

### *B. Convergence of Beliefs*

Proposition 3 shows that the overconfident agent remains overconfident after observing the team's first-period output. Of course, a team will in general be involved in multiple projects, and so learning about others can in general be more precise for any one team member. To capture this possibility, we assume that the team is involved in an infinite number of simultaneous projects for each of which both agents have to choose whether they should exert an effort.<sup>14</sup> We assume that each project has an independent payoff, and that the effort cost of both agents are independent across projects.<sup>15</sup>

With multiple projects to learn from, agent 2 can indirectly learn about her own skill from the outcomes of projects for which she exerts no effort. Indeed, these projects allow agent 2 to learn about the skill of agent 1, which in turn allows for a more precise inference about herself from the projects she works on. As the following proposition shows, however, this is not enough to rid agent 2 of her overconfidence.

**Proposition 5** *With an infinite number of projects, the overconfident agent concludes that her ability is  $\tilde{s}_2 + b^2d$ , where  $\tilde{s}_2$  is her true ability realization.*

An infinite number of projects to learn from still leaves agent 2 with a bias of  $b^2d$  about her own ability. This is because she attributes the success of the team to her own skill and not to the concerted effort of her teammate. More precisely, the overconfident agent expects the team to do well because of her own ability, and the team does indeed perform well. However the team's good performance is the result of a more sustained effort on the part of the rational agent whose marginal product is improved by the effort level of agent 2. So, even though the rational agent

---

<sup>14</sup>A multi-period model in which agents observe the outcome of a project before choosing their effort on the next project would yield insights similar to those of the analysis that follows. However, such a model would require both agents to keep track of every possible path of their teammate's choices of effort in previous periods in order to form "beliefs about their beliefs." This would quickly render the analysis intractable as the number of such paths grows exponentially with each period.

<sup>15</sup>Implicitly, we are assuming that the effort capital of each agent is unlimited and that the firm can be worth infinity. This is harmless, as every project could be made infinitesimal, restoring the bounded nature of effort capital and firm value.

always correctly infers both agents' skills after an infinite number of project payoffs are realized, an infinite amount of data does not make the overconfident agent properly calibrated. Interestingly, as the proof to Proposition 5 shows, the overconfident agent is also biased about the skill of her teammate. Indeed, because the overconfident agent does not expect her teammate to work as hard as she actually does, she concludes that the team's success when only agent 1 works is due to that agent's high skill.

Notice also that  $b$  and  $d$  combine to make the overconfident agent's learning biased. In other words, overconfidence and complementarities again go hand in hand: they improve team performance and welfare and, at the same time, they prolong the positive effects of overconfidence. In fact, as in Proposition 4, it is the case that the ex post overconfidence of agent 2 is increasing in  $b$  and that learning is impaired by larger values of  $b$  (as overconfidence is reduced by a factor of  $1 - b^2$ ).

## 5. Monitoring

As shown in section 3, the presence of an overconfident agent on a team can improve welfare by increasing the equilibrium levels of effort. Of course, overconfidence is not so much a voluntary solution to team problems, but one that evolves from market forces. Indeed, it is unlikely that a team can choose to make some of its members overconfident; instead, as we argue in section 4, the teams that include overconfident members will simply tend to do better than competing teams. Other solutions to team problems have been offered in the literature. These solutions usually revolve around contracting and monitoring mechanisms that restore some of the surplus that the lack of coordination fails to generate. In this section, we explore the role of such a mechanism in the presence of overconfidence.

### *A. A Simple Monitoring Mechanism*

Let us assume that, with probability  $q$ , the effort choices of both agents ( $e_1$  and  $e_2$ ) are observed. This makes it possible for the two agents to share the firm's output unequally, as the two agents can now sometimes tell who is responsible for the team's success. For example, one can easily imagine a situation in which a principal will allocate a larger fraction of the firm's total compensation to one of its agents, either by paying this agent a bonus, by promoting her, or by firing her colleague. To incorporate monitoring into our model, suppose that the two agents agree ex ante that they

keep sharing the project's payoff equally, except when it is revealed that one worked and one shirked. In that scenario, it is agreed that the entire payoff of the project goes to the agent who works.<sup>16</sup> Although the mechanism we consider is rather simplistic, it captures the main features of monitoring: it makes shirking potentially costly for the team's agents and, as we later show, pushes them to work harder. The mechanism's only parameter,  $q$ , measures the intensity of monitoring that is applied to the team: when  $q$  is close to zero, agents are left unmonitored, as before; when  $q$  is close to one, agents are monitored perfectly, that is, their actions can be observed perfectly.

As in previous sections, we assume that agent 2 is overconfident and believes that her ability is  $A = a + d$  although it is truly only  $a$ . Whether agent 2 learns her skill or not will not affect the analysis, and so the underlying structure of overconfident beliefs can be that of section 3 or section 4. For most of this section, we assume that  $b = 0$ . Thus, our analysis focuses on the joint role of overconfidence and monitoring in mitigating the free-rider problem. This assumption is made partly for simplicity and partly to make our analysis comparable to the rest of the literature on team effort. Because Pareto improvements require complementarities between agents, our results focus instead on team welfare which, as argued in section 3, maps into firm value when a principal who hires the two agents is able to capture the surplus from the contracting arrangement. In fact, given that monitoring is probably more likely to take place in a principal-agent framework than in a partnership, this is the interpretation we adopt for this section.

We assume that monitoring is costless. We could close the model by assuming a monitoring cost that is increasing and convex in  $q$ . This additional layer of complexity is not important here as our goal is not to describe the tradeoff between the value and cost of monitoring. Instead we are mostly interested in assessing the value that monitoring can create for a team with and without overconfident agents. Adding a monitoring cost to our analysis would have little or no effect on that comparison; it would simply reduce surplus creation for all scenarios.

### *B. The Effect of Monitoring*

The question we seek to address in the rest of this section is whether overconfidence complements the monitoring mechanism or reduces its value. To address it, we characterize the optimal level of monitoring, that is, the level of monitoring that maximizes team value. More precisely, for a given level  $d$  of overconfidence for the second agent, we determine the optimal monitoring intensity  $q$ .

---

<sup>16</sup>As long as the agent who works gets more than the one who doesn't, any sharing rule will lead to the same results.

Because monitoring is free, it is tempting to immediately conclude that perfect monitoring ( $q = 1$ ) is always optimal. As we next show, this is not the case.

The equilibrium can be derived as in section 3, with two additional considerations: an agent who makes an effort expects to receive an extra payoff of one if the project is successful and the other agent is discovered shirking; an agent who chooses not to work foregoes a payoff of one if the project is successful and the other agent's effort is revealed by monitoring. This tilts the tradeoff between working and shirking towards working. In other words, paying the cost of effort becomes more appealing for both agents, and so they both work harder than without monitoring. The equilibrium is summarized in the following lemma.

**Lemma 2** *In equilibrium with monitoring intensity  $q$ ,*

(i) *agent 1 makes an effort if and only if her cost of effort does not exceed*

$$k_{11} = a(1 + q);$$

(ii) *agent 2 makes an effort if and only if her cost of effort does not exceed*

$$k_{22} = [a + d(1 - aq)](1 + q).$$

As expected, both thresholds are increasing in  $q$ , as both agents work harder when they are monitored more closely. To ensure that both effort thresholds stay smaller than one for all  $q \in [0, 1]$ , we assume that  $a + d - ad < \frac{1}{2}$ . Also, notice that  $d$  does not enter the expression for agent 1's threshold. This is because  $b = 0$ . In the absence of complementarities, the additional effort on the part of agent 2 as a result of her overconfidence does not affect the tradeoff of working and shirking for agent 1. The extra effort exerted by agent 2 as a result of her overconfidence can be shown to be equal to  $\kappa \equiv d(1 + q)(1 - aq)$ , as  $k_{22} = k_{11} + \kappa$ .

The increased effort prompted by monitoring restores some of the team value that is lost to the coordination problems described in section 2. With overconfidence, however, this increased effort may be redundant, as the team already benefits from an increased effort from agent 2. In other words, monitoring is not as needed in the presence of overconfidence, even if it is costless. This point is made more precisely in the following proposition, which studies the optimal intensity  $q^*$  of monitoring as a function of overconfidence ( $d$ ).

**Proposition 6** *The intensity  $q^*$  of monitoring that maximizes team value is*

(i) *equal to one when  $d = 0$ ;*

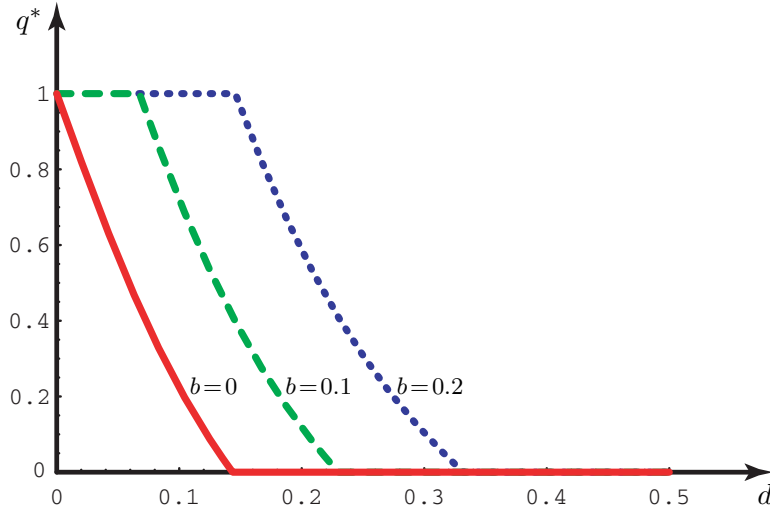


Figure 3: Optimal monitoring as a function of overconfidence.

This figure shows the level of monitoring  $q^*$  that maximizes firm value as a function of agent 2's overconfidence  $d$ . This is done for three different levels of complementarity  $b$  between the two agents. In all three cases, we use  $a = 0.1$  as the actual skill of each agent.

(ii) *decreasing in the level  $d$  of overconfidence.*

It is not surprising that costless monitoring is used as much as possible in the absence of individual biases. More intense monitoring means that the compensation of agents is more sensitive to their effort choices. As a result, agents tend to work harder and this helps solve coordination problems. Less obvious is the result that less monitoring is optimal in the presence of overconfidence. This is because intense monitoring can create an overinvestment in effort on the part of overconfident agents. More precisely, increases in effort by agent 2 resulting from increases in  $q$  above some level  $\bar{q} < q^*$  have a negative impact on team welfare. The optimal monitoring intensity is reached when this effect is exactly offset by the benefit from an increase in effort by agent 1.

Although we could only verify the result of Proposition 6 analytically for  $b = 0$ , numerical calculations show that it also holds when there exist complementarities between the two agents. The only difference is that the optimal monitoring intensity remains at  $q^* = 1$  for small levels of overconfidence. This is intuitive: with complementarities, an increase in effort has a larger impact on firm value, and so overconfidence may prove to be insufficient in generating firm value. This is illustrated in Figure 3, which shows the optimal level of monitoring that the firm should adopt depending on the level of overconfidence and complementarity between the agents.

In sum, monitoring and overconfidence are substitutes rather than complements. Free monitoring can be detrimental for a team when some of its agents are overconfident. If monitoring is costly, then overconfidence might be a more effective way to overcome coordination problems. As mentioned above however, it may not be easy to identify agents that fit this overconfident profile. Thus, as the firm learns about the behavioral characteristics of its agents, it may have to adjust the monitoring that it applies on them.

## 6. Applications

The model we study in this paper is admittedly specialized. However, we think that the lessons that can be learned from it apply to a wide range of economic situations. In this section, we discuss how the model’s main result, the fact that the overconfidence of some agents can generate beneficial outcomes not only to others but to themselves, is transportable to a host of economic problems.

### *A. Technological Innovations*

Many important technological innovations require the concerted effort and dedication of a critical mass of people. For example, the computer industry of the 1970’s would not have flourished this rapidly had it not been for the simultaneous efforts of hardware and software developers. Much the same can be said about the internet in the early 1990’s. In these situations, and surely in several others, the decision for an individual to spend his energy on an innovative activity is likely to depend, at least in part, on the likely contributions of others.

Because coordinated goals and efforts make technological innovations more likely to occur and to benefit its participants, our model can be viewed as a description of these economic scenarios. In particular, our model would imply that there may be some welfare gains from the presence of some individuals who, perhaps irrationally, feel less of a need for others to contribute when they themselves choose whether or not to embark on a technological movement. Coming back to the computer example, one could argue that assembling the first micro-computer to be sold to the masses in one’s apartment<sup>17</sup> may not be fully rational as an economic decision. Indeed, without the efforts of a panoply of other scientists dedicated to the same technology, it is certainly possible that the micro-computer industry would have taken longer before taking off, and the initial efforts

---

<sup>17</sup>We are referring to Steve Jobs and Steve Wozniak building and assembling the Apple I computer in Wozniak’s apartment. See Butcher (1988) for more details on Apple Computer’s early story.



of early movers would not have paid off, at least not for them. However, counting on the likely presence of such individuals, others can feel more justified in investing their time and effort in related innovative products that are necessary for the success of the whole industry.

In this light, our model is close to the entrepreneurship model of Bernardo and Welch (2001) who show that the presence of stubborn individuals who ignore the public information that is available to them in favor of their own less informative information may foster the development of new ideas or the better aggregation of information. Our model differs from theirs in that it shows that individuals who, because of their irrationality, help technological progress along can in fact be better off themselves. Our model also has the virtue of showing that these individuals will tend to take a long time to learn and correct their biases, making their repeated contribution to technological changes possible.

### *B. Investment in Infrastructure*

Investment in any kind of physical infrastructure often benefits from economies of scale and convexities. That is, the coordinated actions of the parties involved in these projects often benefit them all. For example, it may be worthwhile to invest in the opening of a new art gallery in a relatively unexploited area of a city if it is expected that other businesses will be opening at about the same time in the same area. Our model captures the economics of these decisions, and shows that some overconfidence may be key to the success of large-scale investments in infrastructure. Indeed, being able to count on the overinvestment in effort by overconfident restaurateurs may make the art gallery owner more dedicated to his own investment. That is, simultaneous coordinated effort is a better recipe for success, and the overconfidence of some may facilitate that.

### *C. New Movement*

Ideological movements, almost by definition, require the joint participation of multiple individuals. Indeed, whether it is a new art movement, a new religion or a new academic field or subfield, radical shifts in thoughts can only occur if several individuals simultaneously dedicate themselves to their development. Impressionism, for example, would in all likelihood not have affected so many lives if it were not for the concerted enthusiasm of a group of 19<sup>th</sup>-century painters. The success of the project in our model can be interpreted as the eventual large-scale acceptance of the new ideas promoted by such groups of individuals. With this interpretation, our model suggests

that some overconfident individuals may make ideological breakthroughs more likely as they invest more time and effort than is warranted, making it worthwhile for others to join the movement, and accelerating the speed at which the ideas spread and come to influence people.

#### *D. Co-Authoring a Paper*

Finally, our model has an interpretation that is particularly related to our lives as economists. Indeed, if we interpret the successful outcome of a project as the publication of a joint paper, our model points out that the overconfidence of an economist may, in some cases, improve his chances of publishing his work, improving his co-authors' utility and his own in the process. At the time that two economists form a co-authoring team, they do not readily know how costly their effort will be; for example, they do not perfectly know the demands that students, editors and other co-authors will put on their time in the future. If the likelihood of publishing a co-authored paper is improved by complementarities across co-authors, the overconfidence of one will make the other allocate more of his time and effort to their joint project, and in the process make that project more likely to succeed.<sup>18</sup>

#### *E. Sequentiality Issues*

Although our model is about simultaneous coordination, a slight modification to it would also allow for the sequencing of decisions by agents. In these scenarios, moving first is unappealing as it is often not profitable to do so. For example, the initial absence of other businesses may translate into high costs, low profits and a sizeable risk of early bankruptcy for an art gallery in a new part of town. If the risks and costs of moving first are so high, then no one ever moves, and the gains to investment remain untapped.

However, we would expect overconfident individuals to be willing to move early and, expecting that, more rational individuals should be tempted to imitate them more quickly, thus making the success of all more likely. This might be the case for the art gallery owner: believing in his ability to attract a clientele by himself, he is willing to make the investment without knowing about the intentions of restaurateurs; seeing this, restaurateurs are more likely to follow him quickly with

---

<sup>18</sup>Of course, if a paper's success depends only on a threshold number of hours that need to be spent on it as opposed to the total time invested by all the co-authors, then the sustained effort of one author may reduce the involvement of his co-authors and will not produce the Pareto improvements that we describe in this paper. In fact, this would correspond to a case where the efforts of our model's agents are substitutes rather than complements.

their own investment. In fact, one could argue that there is an element of sequentiality in many coordination problems, and so it is reassuring that our model can be adapted to such problems.

## 7. Conclusion

As shown by Holmström (1982), when players share their team’s output but their contribution to that output is unobservable, these players have a tendency to free-ride. Indeed, because a player pays the full cost of her effort but only gets a fraction of its benefit, she scales back on her own effort and instead tends to rely on the effort of others. In equilibrium, the team fails to realize its full first-best value. This problem is exacerbated by the presence of complementarities within the team: because agents don’t fully account for the positive externalities that their effort creates, the team’s level of cooperation is suboptimal and more value is lost. With both problems, mechanisms that increase the effort exerted by the team’s agents recover some of the lost surplus.

This paper explores a new route for increasing effort, namely overconfidence. When agents overestimate their own skills, and thus overestimate the marginal product of their effort, they naturally tend to work harder as, for them, the extra cost of effort is worth the extra reward that they perceive. This of course reduces the extent of the free-rider problem. Such agents also care less about potential complementarities: their own marginal product warrants the extra cost of effort whether or not synergies are realized. Interestingly, this can make the team and all teammates, including the overconfident ones, better off. On the one hand, the overinvestment in effort by an overconfident agent costs her some utility. On the other hand, her increased effort creates a beneficial feedback effect, as the other agents react to the synergistic increase in their marginal product by working harder, thereby increasing the team’s output and thus the overconfident agent’s share of that payoff.

Because all agents are better off and because the team performs better when some of its agents are overconfident, we expect overconfidence to survive the market test. That is, teams equipped with some overconfidence will tend to outperform those without it, and their well-off agents will remain on those teams. Key to this argument, however, is the survival of overconfidence itself. That is, it is important that agents do not quickly figure out their own biases, leaving their team without the benefit of overconfidence. Interestingly, as we show, the same factor that makes overconfidence valuable, namely the presence of complementarities, also makes learning slow. Indeed, overconfident agents expect their effort to increase team output more than is warranted by their ability. Because they also fail to account for the positive effect that their own effort has on that of their teammates,

they attribute the success of their firm to their own ability. In other words, their bias is sustained or, at least, difficult to learn.

Monitoring, even when it is costless, is not always useful for teams that include overconfident agents. For such teams, it is possible that monitoring pushes agents to work so hard that team welfare is sacrificed. When the team is owned by a third-party principal who captures most of the surplus created by labor, firm value may be destroyed by too much monitoring. In a world where the overconfidence of individuals can only be inferred over time, the ability of the principal to adjust incentives through a combination of compensation contracts and monitoring will be key to the firm's success. This last consideration is not studied explicitly here, but should prove to be a fruitful area for future research.

## 8. References

- Alchian, A. A., and H. Demsetz, 1972, "Production, Information Costs, and Economic Organization," *American Economic Review*, 62, 777-795.
- Alpert, M., and H. Raiffa, 1982, "A Progress Report on the Training of Probability Assessors," in *Judgment Under Uncertainty: Heuristics and Biases*, eds. D. Kahneman, P. Slovic, and A. Tversky, Cambridge and New York: Cambridge University Press, 294-305.
- Andolfatto, D., and E. Nosal, 1997, "Optimal Team Contracts," *Canadian Journal of Economics*, 30, 385-396.
- Becker, G. S., 1974, "A Theory of Social Interactions," *Journal of Political Economy*, 82, 1063-1093.
- Bénabou, R., and J. Tirole, 2002, "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 117, 871-915.
- Bénabou, R., and J. Tirole, 2003, "Intrinsic and Extrinsic Motivation," *Review of Economic Studies*, 70, 489-520.
- Bernardo, A., and I. Welch, 2001, "On the Evolution of Overconfidence and Entrepreneurs," *Journal of Economics and Management Strategy*, 10, 301-330.
- Butcher, L., 1988, "Accidental Millionaire: The Rise and Fall of Steve Jobs at Apple Computer," Paragon House Publishers, New York.

- Camerer, C., and D. Lovo, 1999, "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review*, 89, 306-318.
- Carrillo, J. D., and T. Mariotti, 2000, "Strategic Ignorance as a Self-Disciplining Device," *Review of Economic Studies*, 67, 529-544.
- Dawes, R. M., and R. H. Thaler, 1988, "Anomalies: Cooperation," *Journal of Economic Perspectives*, 2, 187-197.
- Eshel, I., L. Samuelson, and A. Shaked, 1998, "Altruists, Egoists, and Hooligans in a Local Interaction Model," *American Economic Review*, 88, 157-179.
- Faulí-Oller, R., and M. Giralt, 1995, "Competition and Cooperation Within a Multidivisional Firm," *Journal of Industrial Economics*, 43, 77-99.
- Ferreira, D., 2002, "Group Loyalty and Incentive Pay," working paper, Getulio Vargas Foundation.
- Fischhoff, B., P. Slovic, and S. Lichtenstein, 1977, "Knowing with Certainty: The Appropriateness of Extreme Confidence," *Journal of Experimental Psychology*, 3, 552-564.
- Gervais, S., J. B. Heaton, and T. Odean, 2003, "Overconfidence, Investment Policy, and Executive Stock Options," working paper, Duke University.
- Groves, T., 1973, "Incentives in Teams," *Econometrica*, 41, 617-631.
- Heaton, J. B., 2002, "Managerial Optimism and Corporate Finance," *Financial Management*, 31, 33-45.
- Heifetz, A., C. M. Shannon, and Y. Spiegel, 2002, "What to Maximize if You Must," working paper, Tel Aviv University.
- Heifetz, A., and Y. Spiegel, 2001, "The Evolution of Biased Perceptions," working paper, Tel Aviv University.
- Holmström, B., 1982, "Moral Hazard in Teams," *Bell Journal of Economics*, 13, 324-340.
- Itoh, H., 1991, "Incentives to Help in Multi-Agent Situations," *Econometrica*, 59, 611-636.

- Kandel, E., and E. P. Lazear, 1992, "Peer Pressure and Partnerships," *Journal of Political Economy*, 100, 801-817.
- Langer, E., and J. Roth, 1975, "Heads I Win, Tails It's Chance: The Illusion of Control as a Function of the Sequence of Outcomes in a Purely Chance Task," *Journal of Personality and Social Psychology*, 32, 951-955.
- McAfee, R. P., and J. McMillan, 1991, "Optimal Contracts for Teams," *International Economic Review*, 32, 561-577.
- Rasmusen, E., 1987, "Moral Hazard in Risk-Averse Teams," *RAND Journal of Economics*, 18, 428-435.
- Rotemberg, J. J., 1994, "Human Relations in the Workplace," *Journal of Political Economy*, 102, 684-717.
- Smith, A., 1976, *The Theory of Moral Sentiments*, Oxford: Clarendon Press, New York. (Originally published in 1759)
- Taylor, S., and J. D. Brown, 1988, "Illusion and Well-Being: A Social Psychological Perspective on Mental Health," *Psychological Bulletin*, 103, 193-210.
- Van Den Steen, E., 2002, "Skill or Luck? Biases of Rational Agents," working paper, Massachusetts Institute of Technology.
- Vander Veen, T. D., 1995, "Optimal Contracts for Teams: A Note on the Results of McAfee and McMillan," *International Economic Review*, 36, 1051-1056.
- Weinstein, N. D., 1980, "Unrealistic Optimism About Future Life Events," *Journal of Personality and Social Psychology*, 39, 806-820.

## 9. Appendix

### Proof of Lemma 1

First, note that agent 2 thinks agent 1 is playing the benchmark game, and thus  $k_{21} = k_{\text{BM}}$ . Solving the maximization problem in (5), taking into account the fact that agent 2 thinks that her ability is  $a + d$ , we get that the threshold employed by agent 2 is

$$k_{22} = k_{\text{BM}} + d. \quad (13)$$

Now, agent 1 knows that agent 2 is overconfident and thus knows her threshold. As a result,  $k_{12} = k_{\text{BM}} + d$ . Finally, using this in the solution to the maximization problem in (5), we get that the threshold employed by agent 1 is

$$k_{11} = k_{\text{BM}} + bd. \quad (14)$$

This completes the proof. ■

### Proof of Proposition 1

(i) Using (7), the expected utility of agent 1 can be written as

$$\bar{U}_1 = a(k_{11} + k_{22}) + bk_{11}k_{22} - \frac{k_{11}^2}{2}.$$

Using (13) and (14) in this expression yields

$$\bar{U}_1 = a[2k_{\text{BM}} + d(b+1)] + b(k_{\text{BM}} + d)(k_{\text{BM}} + db) - \frac{(k_{\text{BM}} + db)^2}{2}.$$

Differentiation of this last expression with respect to  $d$  yields

$$\begin{aligned} \frac{\partial \bar{U}_1}{\partial d} &= ab + a + bk_{\text{BM}} + b^2k_{\text{BM}} + 2db^2 - (k_{\text{BM}} + db)b \\ &= ab + a + \frac{ab^2}{1-b} + db^2 = \frac{a}{1-b} + db^2 > 0. \end{aligned} \quad (15)$$

(ii) Using (8), the expected utility of agent 2 can be written as

$$\bar{U}_2 = a(k_{11} + k_{22}) + bk_{11}k_{22} - \frac{k_{22}^2}{2}.$$

Using (13) and (14) in this expression yields

$$\bar{U}_2 = a[2k_{\text{BM}} + d(b+1)] + b(k_{\text{BM}} + d)(k_{\text{BM}} + db) - \frac{(k_{\text{BM}} + d)^2}{2}.$$

Differentiation of this last expression with respect to  $d$  yields

$$\begin{aligned}
\frac{\partial \bar{U}_2}{\partial d} &= ab + a + bk_{\text{BM}} + b^2 k_{\text{BM}} + 2db^2 - k_{\text{BM}} - d \\
&= ab + a + \frac{ab}{1-b} + \frac{ab^2}{1-b} + 2db^2 - \frac{a}{1-b} - d \\
&= \frac{ab}{1-b} + d(2b^2 - 1).
\end{aligned} \tag{16}$$

This expression is positive if and only if (11) holds. ■

### Proof of Proposition 2

Using (15) and (16), we have

$$\frac{\partial (\bar{U}_1 + \bar{U}_2)}{\partial d} = \frac{\partial \bar{U}_1}{\partial d} + \frac{\partial \bar{U}_2}{\partial d} = \frac{a(1+b)}{1-b} + d(3b^2 - 1).$$

This expression is positive if and only if (12) holds. ■

### Proof of Proposition 3

Suppose that the project succeeds. Using Bayes' rule, agent 2 updates her (biased) beliefs about the mean of her own ability to

$$\mathbb{E}_{\text{B}}[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 2] = \frac{\int_0^{2(a+d)} [f_{\text{B}}(s) \cdot s \cdot (s + ak_{\text{BM}} + bk_{\text{BM}})] ds}{\int_0^{2(a+d)} [f_{\text{B}}(s) \cdot (s + ak_{\text{BM}} + bk_{\text{BM}})] ds},$$

where  $f_{\text{B}}(s) = \frac{1}{2(a+d)}$  is agent 2's biased density function for  $\tilde{s}_2$ . After some straightforward manipulations, this simplifies to

$$\mathbb{E}_{\text{B}}[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 2] = a + d + \frac{(a+d)^2}{3[a+d+(a+b)k_{\text{BM}}]}. \tag{17}$$

Similarly, if the project fails, agent 2 updates her beliefs about the mean of her ability to

$$\mathbb{E}_{\text{B}}[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 0] = \frac{\int_0^{2(a+d)} [f_{\text{B}}(s) \cdot s \cdot (1-s-ak_{\text{BM}}-bk_{\text{BM}})] ds}{\int_0^{2(a+d)} [f_{\text{B}}(s) \cdot (1-s-ak_{\text{BM}}-bk_{\text{BM}})] ds},$$

which simplifies to

$$\mathbb{E}_{\text{B}}[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 0] = a + d - \frac{(a+d)^2}{3[1-a-d-(a+b)k_{\text{BM}}]}. \tag{18}$$

Of course, the true probability of success (given that agent 2 works) is

$$\Pr\{\tilde{v} = 2 \mid e_2 = 1\} = a + ak_{11} + bk_{11} = a + (a+b)(k_{\text{BM}} + bd). \tag{19}$$



Thus, on average, agent 2's beliefs about the mean of her ability will be

$$\begin{aligned}\alpha &\equiv E_B[\tilde{s}_2 | e_2 = 1] \\ &= \Pr\{\tilde{v} = 2 | e_2 = 1\} E_B[\tilde{s}_2 | e_2 = 1, \tilde{v} = 2] + \left(1 - \Pr\{\tilde{v} = 2 | e_2 = 1\}\right) E_B[\tilde{s}_2 | e_2 = 1, \tilde{v} = 0],\end{aligned}$$

which, using (17)-(19), can be manipulated to yield

$$\alpha = a + d - \frac{d(a+d)^2[1 - (a+b)b]}{3[a + d + (a+b)k_{BM}][1 - a - d - (a+b)k_{BM}]}.\quad (20)$$

Because the last term in this expression can be shown to be larger than zero and smaller than  $d$ , we have the desired result. ■

### Proof of Proposition 4

We need to show that  $\frac{\partial \alpha}{\partial b} > 0$ . Differentiation of (20) with respect to  $b$  yields, after some manipulations,

$$\frac{\partial \alpha}{\partial b} = \frac{d}{P^2(1-P)^2} \left\{ (1+2b)P(1-P) + [1 - (a+b)b](1-2P)\frac{a(1+a)}{(1-b)^2} \right\},$$

where  $P \equiv a + d + (a+b)k_{BM}$ . Clearly, this last expression is greater than zero if  $P < \frac{1}{2}$ , a condition that can be rewritten as

$$b < \frac{1 - 2(a+d) - 2a^2}{1 - 2d}.$$

Because the right-hand side of this last expression is larger than  $1 - 4a - 2d$ , this condition is implied by our assumption that  $4a + 2d + b < 1$ . ■

### Proof of Proposition 5

From the perspective of the overconfident agent, there are two groups of projects: those in which she did not exert any effort, and those in which she did make an effort. The success rate of projects in which agent 2 chooses not to work is

$$\tilde{s}_1 k_{11} = \tilde{s}_1 (k_{BM} + bd),$$

where  $\tilde{s}_1$  is the realization of agent 1's skill. However, agent 2 expects agent 1 to exert an effort on only  $k_{BM}$  of the projects, that is, she expects these projects to be successful at a rate of  $\tilde{s}'_1 k_{BM}$ , where  $\tilde{s}'_1$  denotes the realized skill of agent 1 as perceived by agent 2. As a result, the overconfident agent (wrongfully) infers that the skill of her teammate is  $\tilde{s}'_1 = \tilde{s}_1 \left(1 + \frac{bd}{k_{BM}}\right)$ .

This information is used by agent 2 to learn about her own skill from the projects she worked on. These projects are successful at a rate of

$$\tilde{s}_1 k_{11} + \tilde{s}_2 + b k_{11} = \tilde{s}_1 (k_{\text{BM}} + b d) + \tilde{s}_2 + b (k_{\text{BM}} + b d) = \tilde{s}'_1 k_{\text{BM}} + \tilde{s}_2 + b (k_{\text{BM}} + b d),$$

but agent 2 thinks they are successful at a rate of  $\tilde{s}'_1 k_{\text{BM}} + \tilde{s}'_2 + b k_{\text{BM}}$ . Thus she infers that her own skill is  $\tilde{s}'_2 = \tilde{s}_2 + b^2 d$ . ■

## Proof of Lemma 2

This equilibrium is derived in the exact same manner as the equilibrium in Lemma 1 (taking into account the new sharing rule). As such, the proof is omitted. ■

## Proof of Proposition 6

The team welfare is given by

$$\bar{U}_1 + \bar{U}_2 = 2a(2k_{11} + \kappa) - \frac{1}{2}k_{11}^2 - \frac{1}{2}(k_{11} + \kappa)^2.$$

The derivative of this expression with respect to  $q$  is

$$\frac{\partial (\bar{U}_1 + \bar{U}_2)}{\partial q} = 2a \left( 2 \frac{\partial k_{11}}{\partial q} + \frac{\partial \kappa}{\partial q} \right) - k_{11} \frac{\partial k_{11}}{\partial q} - (k_{11} + \kappa) \left( \frac{\partial k_{11}}{\partial q} + \frac{\partial \kappa}{\partial q} \right).$$

We know that  $\frac{\partial k_{11}}{\partial q} = a$  and that  $\frac{\partial \kappa}{\partial q} = d(1 - k_{11} - aq)$ . Thus,

$$\begin{aligned} \frac{\partial (\bar{U}_1 + \bar{U}_2)}{\partial q} &= 2a[2a + d(1 - k_{11} - aq)] - k_{11}a - (k_{11} + \kappa)[a + d(1 - k_{11} - aq)] \\ &= 2a^2(1 - q) + d \left[ (2a - k_{11} - \kappa)(1 - k_{11} - aq) - \frac{\kappa}{d}a \right] \equiv V(q, d). \end{aligned}$$

We can find  $q^*$  by setting  $V(q, d)$  equal to 0 and solving for  $q$ . We can see immediately that  $q^* = 1$  when  $d = 0$ . In order to find the effect of  $d$  on  $q^*$ , we calculate the derivatives of  $V(q, d)$  with respect to  $q$  and with respect to  $d$ :

$$\begin{aligned} \frac{\partial V(q, d)}{\partial d} &= (2a - k_{11} - \kappa)(1 - k_{11} - aq) - \frac{\kappa}{d}a - \kappa(1 - k_{11} - aq) \\ &= (2a - k_{11}) \cdot (1 - k_{11} - aq) - a(1 + q(1 - k_{11})) - 2\kappa(1 - k_{11} - aq) \\ &= a - 2ak_{11} - 2a^2q - k_{11} + k_{11}^2 + 2k_{11}aq - aq - 2\kappa(1 - k_{11} - aq) \\ &= a - 2a^2 - 4a^2q - a - aq + a^2 + 2a^2q + a^2q^2 + 2a^2q + 2a^2q^2 \\ &\quad - aq - 2\kappa(1 - k_{11} - aq) \\ &= -a^2 - 2aq(1 - \kappa) + 3a^2q^2 - 2\kappa(1 - k_{11}) \\ &= -a^2(1 - q^2) - 2aq(1 - \kappa - aq) - 2\kappa(1 - k_{11}) \end{aligned} \tag{21}$$

and

$$\begin{aligned}
\frac{\partial V(q, d)}{\partial q} &= -2a^2 + d \begin{bmatrix} -[a + d(1 - k_{11} - aq)](1 - k_{11} - aq) \\ -2a(2a - k_{11} - \kappa) - a(1 - k_{11} - aq) \end{bmatrix} \\
&= -2a^2 + d \begin{bmatrix} -[2a + d(1 - k_{11} - aq)](1 - k_{11} - aq) \\ -2a(2a - k_{11} - \kappa) \end{bmatrix} \\
&= -2a^2 + d \left[ -d(1 - k_{11} - aq)^2 - 2a(1 - \kappa - 3aq) \right] \\
&= -2a^2(1 - dq) + d \left[ -d(1 - k_{11} - aq)^2 - 2a(1 - \kappa - 2aq) \right]. \tag{22}
\end{aligned}$$

Because  $k_{11} + \kappa < 1$ , we have  $1 - \kappa - aq > 0$  and  $1 - \kappa - 2aq > 0$ , which imply that (21) and (22) are both negative. By the implicit function theorem, this further implies that the optimal monitoring intensity  $q^*$  is decreasing in  $d$ . ■