The Effects of Biased Self-Perceptions in Teams*

Simon Gervais
Fuqua School of Business
Duke University

Itay Goldstein
Wharton School
University of Pennsylvania

2 June 2005

*This is an updated version of a paper previously circulated under the title "Overconfidence and Team Coordination." Financial support by the Rodney L. White Center for Financial Research is gratefully acknowledged. We would like to thank Roland Bénabou, Alon Brav, Thomas Chemmanur, Alexander Gümbel, Lars Hansen, Yael Hochberg, Ken Kavajecz, Leonid Kogan, Tracy Lewis, Robert Marquez, Sendhil Mullainathan, John Payne, Canice Prendergast, Manju Puri, Steve Ross, Steve Slezak, Anjan Thakor, 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, the University of Wisconsin at Madison, the University of Chicago, New York University, Duke University, the University of Pennsylvania, the 2003 Workshop on New Ideas and Open Issues in Corporate Finance held in Amsterdam, the 2004 FIRS Conference on Banking, Insurance and Intermediation held in Capri, and the 2005 Conference of the Caesarea Center in Israel. All remaining errors are the authors' responsibility.

Correspondence address: Simon Gervais, Fuqua School of Business, Duke University, One Towerview Drive, Durham, NC 27708-0120, (919) 660-7683. Itay Goldstein, Finance Department, Wharton School, University of Pennsylvania, Steinberg Hall – Dietrich Hall, Suite 2300, Philadelphia, PA 19104-6367, (215) 746-0499.

Email address: sgervais@duke.edu and itayg@wharton.upenn.edu.

The Effects of Biased Self-Perceptions in Teams

Abstract

Several finance and economics problems involve a team of agents in which the marginal productivity of any one agent increases with the effort of others on the team. Because the effort of each agent is not observable to any other agent, the performance of the team is negatively affected by a free-rider problem and by a lack of effort coordination across agents. In this context, we show that an agent who mistakenly 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 biased agent overworks, she benefits from the positive externality that other agents working harder generates. The presence of a team leader improves coordination and team value, but self-perception biases can never be Pareto-improving when they affect the leader. Because self-perception biases naturally make agents work harder, monitoring, even when it is costless, may hurt the team by causing an overinvestment in effort. Interestingly, the benefits of self-perception biases may be long-lived even if agents learn from team performance, as the biased agent attributes the team's success to her own ability, and not to the better coordination of the team.

1. Introduction

Cooperation, coordination and synergies are sources of value in many economic situations. For example, the success and viability of integrated firms (Grossman and Hart, 1986; Scharfstein and Stein, 2000), partnerships (Farrell and Scotchmer, 1988; Levin and Tadelis, 2005), strategic alliances (Jensen and Meckling, 1995; Holmström and Roberts, 1998), and joint ventures (Alchian and Woodward, 1987; Kamien, Muller and Zang, 1992; Aghion and Tirole, 1994) are affected by the efficiency with which various entities and agents interact with each other. In fact, the view that firms form endogenously as a way to gather and take advantage of complementary activities dates back to Alchian and Demsetz (1972). Similarly, some authors have argued that loan syndicates (Pichler and Wilhelm, 2001) and corporate boards (Hermalin and Weisbach, 1998) benefit from team cooperation. Even financial markets can benefit from coordination. For example, Dow (2004) shows that liquidity can be self-fulfilling: as long as liquidity demanders participate in markets, there will be liquidity in those markets. More generally, economic development and the efficient provision of public goods often require the concerted efforts of several people, businesses and organizations. Campbell (1986) even argues that cooperation is a healthy component of evolutionary behavior.

A natural approach to studying the incentives of an organization's members in the presence of cooperation and coordination issues has been Holmström's (1982) model of a team. The general idea behind this model is 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 problems 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), and Taylor and Brown (1988) document that individuals tend to perceive themselves as having more ability than is warranted. According to Kunda (1987), they also tend to believe in theories that imply that their own attributes cause desirable outcomes. Similarly, Fischhoff, Slovic and Lichtenstein (1977), and Alpert and Raiffa (1982) find that individuals tend to overestimate the precision of their information. In business settings, Larwood and Whittaker (1977) find that managers tend to believe that they are superior to the average manager, and Cooper, Woo and Dunkelberg (1988) find that entrepreneurs perceive their own chance for success as being higher than that of their peers. We incorporate such self-perception biases into the team problem by assuming that some players, which we refer to as *overconfident* players, overestimate the degree to which their effort contributes to team success (i.e., the marginal product of their effort). We show that the bias can not only overcome 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 agent who has an inflated opinion of herself 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 team members exert positive externalities on each other. 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 self-perception bias 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 imported behavioral considerations into team contexts. For example, Rotemberg (1994) analyzes the effect of altruism on coordination in teams. 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. 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.

Interestingly, it is not the concern for others or of others that solves coordination problems in our model. Instead, it is the extreme self-perception of some agents that does. Biased 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 matter little to overconfident agents but do 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. In that sense, our model is closer to that of Kelsey and Spanjers (2004) who show how the ambiguity aversion of some agents leads them to use personal effort as insurance for the effort of others, alleviating the free-rider problem in the process. Also closely related is the work of Gaynor and Kleindorfer (1991) who show that misperceptions about the production function can have positive effects.

Another approach to incentive problems in teams has concentrated on the organizational structure of the team. An example of this approach is found in the work of Hermalin (1998) who discusses the role of leadership in fostering team effort. We incorporate leadership into a two-agent framework by assuming that the effort choice of one agent — the leader — is made public before the other agent — the follower — makes her effort choice. This structure naturally mitigates the coordination problem to some extent as it creates an incentive for the leader to work harder knowing that her effort choice will affect the effort exerted by the follower. In essence, one agent leads the other by example. Using this framework, we study how self-perception biases and leadership interact in contributing to agents' welfare and to team value. We find that self-perception biases can generate Pareto improvements only when the rational agent is the leader. We also show that team production and value are maximized with a biased leader when her bias is small, but with a rational leader otherwise. As far as we know, this is the first set of results on the optimal organization design of a team in the presence of behavioral biases.

We also analyze the role played by monitoring in the presence of biased self-perceptions. 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 behavioral characteristics of these agents. Also, because monitoring and overconfidence can be substituted for each other, picking individuals with useful behavioral biases, like self-perception biases, becomes quite valuable for the firm when monitoring is costly.

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 their irrationality longer-lasting. Since self-perception biases lead to better performance in the presence of complementarities, an implication of this result is that complementarities across agents are responsible for both making the biases useful (for the team and its agents) and making them persist (through slower learning).

The 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.[1] More precisely, the two papers show that individuals who display an overconfidence bias will be better off in the long run, *assuming that they remain biased*. Our paper shows that the ability to learn about oneself can be mitigated by the very presence of the self-perception bias and that, as a result, *the bias tends to persist*. This result that biases are either slowly or never learned away further guarantees the survival of individuals with biased self-perceptions. Indeed, overconfident individuals will tend to survive in the early rounds and, in the process, will not learn their bias, 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),

[1]Our paper differs from these in two other respects. First, our paper analyzes the role of overconfidence in coordinating the decisions of team members, while these two papers focus on the survival of biased individuals. Second, we do not allow individuals to choose the biases that will make them better fit, as these authors do. Instead we treat overconfidence as innate, like preferences and skills.

but not for the reason the overconfident agent thinks (her high ability).

Before we embark on the details of the paper's main model, it is important to point out that our objective in this paper is not to argue that agent biases constitute a solution to team coordination problems. In particular, we certainly do not mean to suggest that agents' self-perceptions can be optimally modified at the turn of a dial, like compensation or monitoring. Instead we believe that behavioral biases are innate characteristics of individuals, just like risk aversion and ability. At the same time, we also believe that some of these biases can have useful roles, for their beholders and their peers, in some contexts. As such, our work should be viewed as an exploration of contexts in which biased individuals will flourish, as opposed to simply create exploitable opportunities for their rational counterparts.

The rest of the paper is organized as follows. In section 2, we start with a brief discussion of applications in which coordination issues arise as a result of a team context. One such application, the partnership, is the subject of our main model. In section 3, we set up the two-agent framework that is used throughout the paper, and highlight the coordination problems that arise in it. Section 4 introduces self-perception biases, and shows how they 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 biased agent's overinvestment in effort may not only benefit her team and teammates, but also herself. Section 5 looks at the effects of making one of the team members the team leader. More precisely, the rational-leader and biased-leader scenarios are compared in terms of individual welfare and team value. Section 6 looks at the joint roles of self-perception biases and monitoring in the team context, and shows them to be substitutes in the sense that the presence of one may render the other detrimental. The possibility that an agent's bias changes as she learns her skills through the team's output is considered in section 7. Finally, alternative interpretations and applications of our model are discussed in section 8, which also offers some final remarks and concludes. All proofs are contained in Appendix A.

2. Team Coordination: Applications

As mentioned in the paper's introduction, many applications of the team coordination problem originally studied by Holmström (1982) can be found in the finance and economics literature. In an effort to better situate our paper, we devote this section to a review of some of these applications. For now, we concentrate our review on corporate finance applications, and postpone our discussion of wider-ranging applications until the paper's conclusion. For each of the applications discussed, we explain the source of its inherent coordination and free-rider problem. We also reflect on the extent to which positive externalities can affect each problem, as such externalities are an important component of our theory. While we believe our model can apply to any of these settings, the rest of the paper will focus on the first of the applications reviewed below, namely the partnership setting.

A. Partnerships

Many firms choose a partnership as their ownership structure. For example, partnerships are the dominant organizational form in certain industries, such as law, accounting, and investment banking. In fact, overall, firms that provide human-capital-intensive services are more likely to organize as partnerships.

Effort coordination problems of the type studied in this paper are inherently prevalent in partnerships. This is in fact why we adopt a partnership as the basis of our model. In a partnership, partners share the firm's profits according to a pre-specified sharing rule. Clearly, although the effort of any one partner benefits all partners, the cost of this effort in entirely borne by the one partner. Thus, the free-rider problem arises naturally. Moreover, positive externalities among partners are likely to exist. They may exist simply because of synergies in production. Alternatively, they may be the product of the central role played by firm reputation, especially in firms that provide human-capital-intensive services. In such firms, the effort of a partner contributes to the firm's reputation, and this increases the productivity of effort of other partners. For example, a lawyer, who expects that his peers are going to shirk, realizes that the reputation of the firm is likely to deteriorate, and thus that his effort will have very little effect on the overall value. The importance of reputation in the context of partnerships is in fact emphasized in two recent papers by Morrison and Wilhelm (2004) and by Levin and Tadelis (2005).

B. Multi-Division Firms

The division managers of a multi-division firm will also typically face effort coordination problems. As in the case of partnerships, division managers bear the full costs of their efforts, but share the gains with other division managers. This point has been discussed and demonstrated in several articles, including Boot and Schmeits (2000), and Scharfstein and Stein (2000). In this context, positive externalities across divisions arise as a simple result of production synergies. They may also be a product of financing spillovers that efficient internal capital markets render possible. Indeed, as shown by Stein (1997), the success of one division will provide more resources to the firm, and thus will enable other divisions to get more financing for their investments. This may increase the productivity of the other divisions (or, more precisely, their incentive to be productive).

C. Syndicates

Financial institutions often form syndicates for the purpose of providing underwriting or lending services. Pichler and Wilhelm (2001) write that most securities sold in U.S. public markets are underwritten by syndicates of investment banks. Sufi (2005) documents that non-financial U.S. businesses obtain almost $1 trillion in new syndicated loans each year. This represents approximately 15 percent of their aggregate debt outstanding.

A syndicate of financial institutions is likely to encounter team coordination issues. To protect the financial interests of its members, the syndicate has to collect information and monitor the firm to which financial services are being provided. However, as argued by Pichler and Wilhelm (2001), a free-rider problem naturally arises inside the syndicate. Each member of the syndicate has to bear the full cost of his monitoring effort, while enjoying only a share of the resulting gain. Positive externalities may also exist between the information-gathering efforts of the different members of the syndicate. A piece of information collected by one member may not be enough to alter the behavior of the syndicate towards the firm. Thus the value of a piece of information may increase when other members are also making an effort to obtain information.

Interestingly, the syndicate also represents a particularly good setting for studying leadership issues, which we analyze in section 5. Indeed, as discussed by Pichler and Wilhelm (2001) and Sufi (2005), such syndicates are always explicitly led by one of their members (the 'lead bank' or the 'lead arranger'), and so leadership issues are, as in the later version of our model, intertwined with coordination issues.

D. Boards of Directors

Boards of directors are designed to protect the interests of shareholders by monitoring the managers and making sure that they act to maximize shareholder wealth. This role is particularly important in corporations with diffuse ownership structures, as direct monitoring by shareholders is difficult and costly.

The team effort problem studied in this paper applies directly to the monitoring effort by the board of directors, as proper monitoring of firm managers often requires the concerted effort of several directors. Directors, however, might free-ride on the effort of others as they share the reputation benefit of their own effort while bearing its full cost personally. The idea that these coordination problems may arise within boards of directors has been recognized, for example, by Hermalin and Weisbach (1998). Positive externalities across directors are also likely to affect the monitoring function of corporate boards. Indeed, as in syndicates, the informed input of one member may be worthless if it isn't coupled with the informed input of other members. Thus, as in our model, a director's effort has more value when other directors also exert some effort. In fact, given the distinct role played by the chairman of the board, it is also the case that leadership issues are likely to impact the effectiveness of corporate boards.

E. Venture Capital

The idea that the venture capital function is plagued by a double-sided moral hazard problem between the venture capitalist (VC) and the entrepreneur can be found in Sahlman (1990), Lerner (1995), Hellmann and Puri (2002), and Kaplan and Strömberg (2004). These authors argue that, in addition to the contribution that the entrepreneur's effort is bound to have on the potential success of the company, the VC's effort towards monitoring, advising and organizing the company can also impact its eventual fate. As such, it is reasonable to think of the relationship between the entrepreneur and the VC as a team problem in which the effort of one benefits both, as in the model of venture capital by Casamatta (2003).

In this context, complementarities between the entrepreneur and the VC are also likely to exist, as the dedication of one to the company can potentially make the other more dedicated as well. For example, a VC with limited human capital may choose to allocate more of it to a company in which the entrepreneur appears to be fully engaged. Likewise, the entrepreneur is less likely to turn his attention to alternative outside opportunities if she feels the committed support of the VC.

3. 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*. As we discuss later, the addition of a principal who hires these agents to operate and manage his firm can easily be accommodated in our model but, to keep the intuition simple, we start with this simpler framework. The value of the firm comes from a single one-period project, which can either succeed or fail with probabilities π and $1 - \pi$ respectively. The project generates two dollars at the end of the period if it succeeds, and it generates zero if it fails. Thus the firm's end-of-period cash flow is given by

$$\tilde{v} \equiv \begin{cases} 2 & \text{prob. } \pi \\ 0 & \text{prob. } 1 - \pi. \end{cases} \tag{1}$$

The probability of success π 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

$$\pi = ae_1 + ae_2 + be_1e_2, \tag{2}$$

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) view that teams (or firms) form to take advantage of positive externalities or complementarities. Of course, since π 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:

$$0 \leq 2a + b \leq 1. \tag{3}$$

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

$$\tilde{U}_i = \frac{1}{2}\tilde{v} - \tilde{c}_i e_i. \tag{4}$$

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.[2] Each agent privately observes her own cost,

[2]These distributional assumptions about $\tilde{c}_i$ are made purely for analytical convenience. The only required assumption is that effort costs are not perfectly correlated.

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, 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:

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

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

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

Solving for k_1 and k_2 in these equations, we find

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

Clearly, k_{BM} is increasing in both a and b. That is, agents work harder (or, more precisely in this model, work more often) when they are more skilled and when their partnership is more synergistic.

The result that skilled agents work harder is a direct product of the assumption that the marginal productivity of effort is increasing in a. The same result would obtain if we were to assume that the marginal disutility of effort is lower at all effort levels for higher ability agents.[3] Such an assumption is in fact made in several papers in which there is skill heterogeneity across agents, whether the models concentrate on signaling (e.g., Spence, 1973), rank-order tournaments (e.g., Lazear and Rosen, 1981), screening (e.g., Garen, 1985), or multi-period contracting (Lewis and Sappington, 1997).

Admittedly however, there is no economic theory justifying any assumption that implies a positive relationship between skill and effort.[4] Indeed, one can easily imagine contexts in which a highly skilled agent simply scales back on effort, as her lower but more productive effort achieves the same result as the more sustained effort of lower skilled agents and allows her to enjoy more leisure utility.[5] This is in fact the assumption made by Lazear (2000) in his model of performance pay. Alternatively, many authors prefer to stay away from any assumption that creates a relationship between skill and effort by making the production function additive in the two. This is, for example, the assumption carried through Prendergast's (1999) review of the literature on incentives in firms. For our purposes, the relationship between skill and effort becomes critical only when we introduce self-perception biases into the model. As we discuss then, alternative interpretations can be offered for the overinvestment in effort that will accompany these biases.

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

$$\bar{U}_1 \equiv \mathrm{E}\left[\tilde{U}_1\right] = ak_1 + ak_2 + bk_1k_2 - \int_0^{k_1} c_1\, dc_1 = ak_1 + ak_2 + bk_1k_2 - \frac{k_1^2}{2} \tag{7}$$

[3]We favored making the marginal productivity of effort increasing in a because, in later sections, the cost of effort would reveal an agent's own ability, which will be assumed misperceived. Making the utility cost of effort random, like the project's value, would reconcile the two approaches, but the additional structure of doing so appears unnecessary.

[4]Instead the assumption is likely to be driven by analytical considerations. Indeed, when effort, skill and output are monotonically related, the output allows the principal to (imperfectly) deduce the agent's skill, rendering screening possible, for example.

[5]Interestingly, Schor (1993) documents that workers do exactly the opposite: they allocate the hours that suddenly become available for leisure to extra work. So it appears that individuals benefit from leisure up to a certain point, but derive more utility from work once they have achieved a certain minimum leisure utility.

and

$$\bar{U}_2 \equiv \mathrm{E}\left[\tilde{U}_2\right] = ak_1 + ak_2 + bk_1k_2 - \frac{k_2^2}{2}. \tag{8}$$

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, $k_1 = k_2 = k_{\mathrm{BM}}$ and so both agents' expected utility is given by

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

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 effort 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 $k_1 = k_2 = k_{\mathrm{FB}}$ and will maximize (7) and (8). We find that the interior solution to this problem is given by

$$k_{\mathrm{FB}} = \frac{a}{\frac{1}{2} - b} \tag{10}$$

as long as $a + b < \frac{1}{2}$, which we assume from now on for convenience.[6]

Clearly, $k_{\mathrm{FB}} > k_{\mathrm{BM}}$. That is, agents do not exert enough effort in the equilibrium of the benchmark model.[7] 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_{FB} and k_{BM} is increasing in b.

[6] Otherwise, the corner solution given by $k_{\mathrm{FB}} = 1$ unnecessarily complicates the analysis.

[7] 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.

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.

4. Self-Perception Biases

In this section, we show how the coordination problem of the team is sometimes mitigated by the behavioral characteristics of its agents. Our approach emphasizes the role played by self-perception biases, a behavioral characteristic of individuals that has been extensively documented in the psychology literature. In particular, Langer and Roth (1975), and Taylor and Brown (1988) document the fact that people tend to overestimate their own skills, and Larwood and Whittaker (1977) show that business managers suffer from the same bias. Similarly, Greenwald (1980) documents that people's self-evaluations tend to be unrealistically positive. Moreover, Dunning, Meyerowitz and Holzberg (1989) find that such biases are more pronounced when the definition of competence is ambiguous, which is likely to be the case in many economic contexts. In what follows, we incorporate these findings into our model of the team, and show that self-perception biases can have useful coordination properties.

A. Introducing Self-Perception Biases

Suppose that agent 2 is biased about her own ability. Specifically, she thinks her ability is $A > a$, although it is truly only a. For now, we use the previously cited psychology literature as a justification for this assumption, and we do not discuss how and why this agent became biased about her ability. Later in the paper, we argue that the addition of a learning component to the model makes such biases persist naturally at both the firm and the individual levels. All the other details of the model remain the same as in the previous section.[8] Note that it is the departure from true ability to perceived ability that represents agent 2's bias. We denote this quantity by $d \equiv A - a$, and refer to it as her *self-perception bias* or *level of overconfidence.*

We assume that agent 1 knows that agent 2 is biased. 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.

[8]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 a biased agent, is between zero and one.

This is because our welfare results depend on whether other agents change their behavior when teamed with a biased agent. Still, as long as agent 1 assigns a positive probability to the possibility that agent 2 is biased, our welfare results will go through. 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 this. This assumption only affects the version of the model that incorporates learning, studied in section 7. Until that section, all of our results hold under the alternative assumption that agent 2 thinks that agent 1 recognizes her superior ability.[9] 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 they agree on the fact that she contributes more to it than the other agent.[10] 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 3. 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.[11] Using the same reasoning as in the benchmark model of section 3, we can derive the equilibrium strategies for the two agents,

[9] Appendix B includes a proof that this is indeed the case. Readers should note however that this appendix should be read only after reading this section, as it uses notation that we are about to introduce.

[10] Farrell and Scotchmer (1988) study a closely related issue in their model of partnerships with heterogenously skilled agents. Because each agent gets an equal share of the firm's value, it is the size of the partnership (i.e., the number of partners) that is solved for in equilibrium. A similar approach could be used here to determine the optimal size of the partnership before proceeding to (or as part of) our analysis of the effects of overconfidence. However, we feel that the added complexity would muddle the intuition without adding any insight.

[11] In section 3, we effectively had $k_1 = k_{11} = k_{21}$ and $k_2 = k_{22} = k_{12}$.

when agent 2 is biased about her own ability.

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

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

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

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

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

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 in fact 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 firms, organizations and industries that benefit more from synergies among workers.

As discussed in section 3, the positive relationship between productivity and effort cannot be unquestionably justified on economic grounds. A possible justification for it can be found in the work of Athey and Roberts (2001) on organizational design. In their model, agents must split their finite effort capital across all available endeavors (which, in theory, could include leisure), and in doing so tend to allocate more effort to activities and projects that they feel are productive. With this in mind, the bias that we introduce above is simply that of an agent who perceives her productivity to be higher in the one partnership modelled here.

Another perspective on Lemma 1 comes from the psychology literature on human motivation.[12] Atkinson (1978) argues that individuals differ in their tendency to approach achievement-related goals. In his theory, three factors influence this tendency: the need for achievement, the probability of success and the incentive value of success. The first of these factors amounts to a natural attraction that some individuals have for positively affecting the outcome of certain tasks. So, although Lemma 1's result that agent 2's biased self-perception leads her to exert more effort, her increased effort could have been equivalently assumed as a natural tendency for some individuals to work harder in order to fulfill their need for achievement. In fact, individuals with such high needs for achievement have been found to favor business occupations (Meyer, Walker and Litwin, 1961) and to exhibit positive biases when assessing success probabilities (Feather, 1965).

The natural tendency that some individuals have for exerting effort is also often referred to as intrinsic motivation (Deci, 1975). Intrinsically motivated individuals find some satisfaction in certain activities even when minimal external rewards exist for them to engage in those activities. In other words, the activity itself is the reward.[13] In economics terms, it is possible that a worker's pride in her work makes the disutility of effort negligible or even negative (i.e., the exertion of effort yields positive utility) for some levels of effort, as suggested by Baron (1988), Kreps (1997), and Baron and Kreps (1999). More than that, Deci (1975) also suggests that the perception of one's own skills will affect one's desire to work when he writes that "if a person's feelings of competence and self-determination are enhanced, his intrinsic motivation will increase" (p. 141). Similarly, Locke and Latham (1990), and Heath, Larrick and Wu (1999) show that optimism and dedication tend to go hand in hand. Indeed, these authors find that individuals with optimistic goals tend to work harder than individuals with more realistic goals. That our agent 2's biased perception of her own ability leads to a more sustained effort on her part is certainly consistent with this evidence.

Finally, 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 7.

[12] An excellent overview of this literature is contained in Weiner (1985).

[13] In fact, as argued by Deci (1975), Kruglanski (1975) and Kohn (1999), and as modelled by Bénabou and Tirole (2003), extrinsic motivation in the form of external rewards can crowd out intrinsic motivation.

C. Self-Perception Biases and Individual Welfare

Because, as discussed in section 3, 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.

Proposition 1 *Suppose that agent 2 is biased, 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

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

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 bias, as long as this bias is not too extreme, 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 severe. 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,[14] 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 bias 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 finance and economics literature, which describes how biased agents fail to realize the full value of their opportunities while others benefit from their mistakes.[15] 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), changes in their decision-making (choice of effort, in this case) may have a positive effect on everyone.

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 worse off when her overconfidence is too extreme.

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

[15] An exception is the aforementioned paper by Gaynor and Kleindorfer (1991) in which agents misperceive the production function.

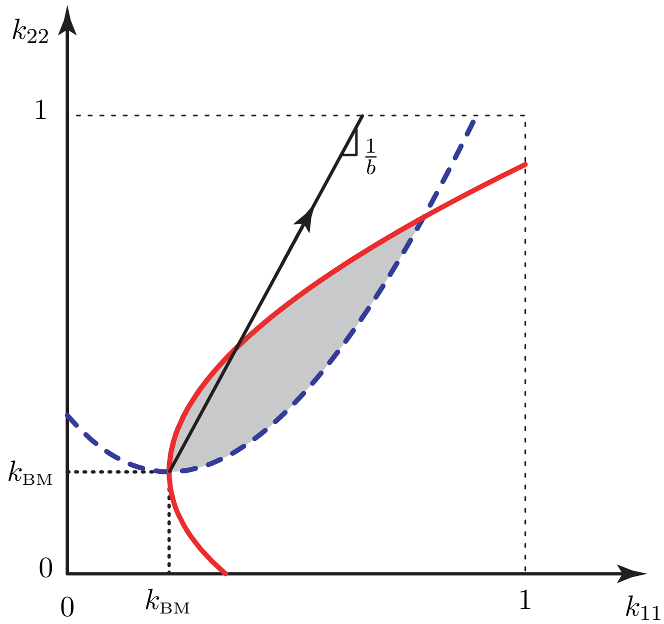

Figure 1: Iso-utility curves and agent welfare.

In the benchmark scenario, both agents use an effort cost threshold of k_{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.

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 complementarities 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 $k_{11} = k_{22} = k_{\mathrm{BM}}$, 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.

An alternative interpretation of our results about the welfare of agent 2 is that her self-perception bias motivates her to work harder, which in turn motivates her teammate to also work harder. The latter effect makes her better off. Bénabou and Tirole (2002) also 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.

Finally, it is important to point out that it is the existence of agent 2's overinvestment in effort that is key to this section's results, not the origin of her overinvestment in effort. Indeed, although we show in Lemma 1 that agent 2 works harder as a result of her biased self-perceptions, assuming that she does so because of her intrinsic motivation would be equally justifiable, based on the psychology literature. With such an assumption as a primitive to our model, Proposition 1 then says that intrinsic motivation can contribute to welfare, even if welfare is calculated without the intrinsic utility.[16] Our model then formalizes Galbraith's (1977) view that "other things equal, intrinsic motivation will increase the likelihood that spontaneous and cooperative behaviors are chosen by the individual" (p. 340).

D. Self-Perception Biases and Team Welfare

In our model, the sharing rule between the two teammates is prescribed: they each get half the team's output. This is why the model applies particularly well to a partnership. Another 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 employ-

[16] Clearly, adding intrinsic utility would make the result stronger.

ees (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 also refer to as *team welfare*, is studied in the following proposition.

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

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

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 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.

As mentioned above, Locke and Latham (1990), and Heath, Larrick and Wu (1999) argue that individuals who have optimistic views of certain projects tend to work harder at these projects. Although their work is silent about the ultimate welfare of these individuals, our model shows that certain team contexts are well suited for them. Indeed, in synergistic teams, they are indirectly rewarded for their excessive effort. In that sense, our results agree with Armor and Taylor's (1998) findings that some excess motivation to work can be beneficial in certain situations. Also, note that, as before, it is not necessary for the agent who overworks to be biased about her own ability for the above results to hold. The tendency to overwork may come from other behavioral characteristics of the agent. As long as one agent mistakenly overinvests in costly effort and that her teammate is aware of that bias, the complementarities between the two agents will ensure that they both benefit from the bias. This would be the case, for example, if the rational agent only attributed a positive

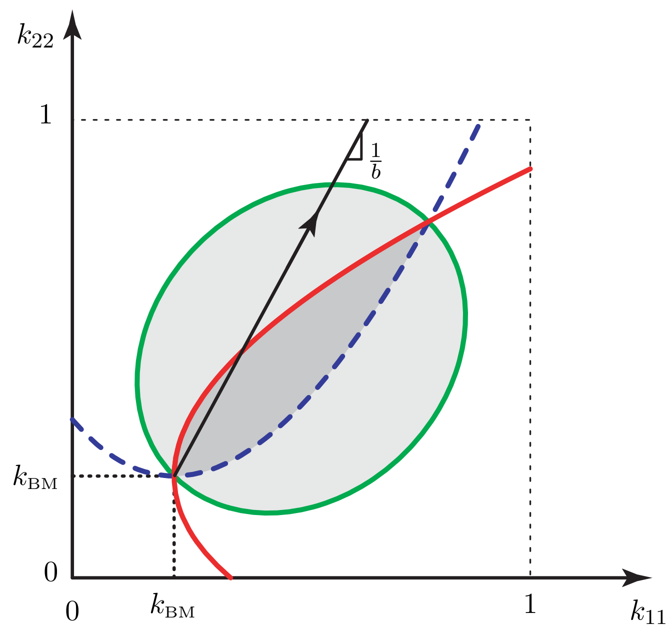

Figure 2: Iso-utility curves and team welfare.

In the benchmark scenario, both agents use an effort cost threshold of k_{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.

probability that her teammate is biased.

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

5. Leadership

As shown in section 4, the presence of an overconfident agent can increase a firm's value by increasing the equilibrium levels of effort. That is, some behavioral traits of agents can naturally make them valuable teammates. 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 (firm) can choose to make some of its members (employees) overconfident; instead, the teams that include overconfident members will simply tend to do better than competing teams.[17] Other solutions to team problems have been offered in the literature. These solutions revolve around more explicit mechanisms designed to restore some of the surplus that the lack of coordination fails to generate. In this and the next sections, we explore two such mechanisms: the presence of a team leader and monitoring.

A. Introducing a Leader

Because the unobservability of effort is partially responsible for the coordination problems of the team, it is natural to expect public effort choices to reduce the extent of these problems. Indeed, if an agent knows more about the presence or absence of synergies when she makes her effort decision, she is more likely to work at the same time as her teammate. The notion of leadership that we explore in this section captures this sequential aspect in effort choices. In particular, we assume that the effort choice of one agent, the *leader*, is made public before the other agent, the *follower*, decides whether or not to exert an effort.

In such a setting, the leader uses her public choice of effort to influence the effort decision of the follower. In particular, the leader can internalize the externalities that her effort choice has on her teammate, as her actions affect her teammate's actions. In that sense, she leads by example. This notion of leadership is similar to that developed by Hermalin (1998). In his work, the leader is endowed with some information about the profitability of a project, and uses her public effort choice to boost the credibility of her attempt to signal it to the other agents. Our model differs from Hermalin's in that our leader does not have any informational advantage about the project. In particular, her leadership role is limited to the fact that she moves first and her effort is publicly observable. We also allow a biased agent to lead a rational follower or to follow a rational leader.

The model with a leader is solved in essentially the same way as the no-leader model of section 4, with the exception that the follower can make her effort choice based on the observable effort of the leader. This, of course, means that the follower will have a different effort threshold depending on whether the leader exerted an effort or not. In terms of notation, we denote the effort threshold of the leader by k_{L} and those of the follower by k_{F1} (after the leader exerts an effort) and k_{F0} (after the leader chooses not to work). These thresholds are derived in the following lemma.

Lemma 2 *With a biased leader, the equilibrium thresholds are given by* $k_{\mathrm{L}} = a + d + b(2a + b)$,

[17] We return to this argument in section 7.

$k_{\text{F1}} = a + b$, and $k_{\text{F0}} = a$. With a rational leader, the equilibrium thresholds are given by $k_{\text{L}} = a + b(2a + b + d)$, $k_{\text{F1}} = a + d + b$, and $k_{\text{F0}} = a + d$.

The effort decision of the follower is straightforward to derive: she makes an effort if and only if her effort cost is smaller than the probability that (she thinks) she contributes to the project being successful (since her payoff is one in that case). Because of the effort complementarities of the agents, the leader's effort increases this probability by b. Also, the biased follower thinks that her effort contributes an additional probability of d.

Clearly, the rational follower's thresholds are not affected by the leader's bias; only the frequency with which the higher threshold (k_{F1} vs. k_{F0}) is chosen is affected by d. This frequency, of course, is given by k_{L}, and so the average effort level of the follower is given by $k_{\text{L}}k_{\text{F1}} + (1 - k_{\text{L}})k_{\text{F0}}$. Given this, it is easy to show that both agents work harder with a leader than without one, whether the leader is the rational agent or the biased agent. The fact that her action is observed by the follower commits the leader to exerting effort more frequently. Because of complementarities, the higher effort exerted by the leader creates an incentive for the follower to work harder (or more often), on average, as well. Incidentally, this is another difference between our model and that of Hermalin (1998): because our model's leader does not have any information about the project's fundamentals, she can only commit the follower to working harder when synergies exist between them (i.e., when $b > 0$).

B. The Effects of Leadership

Without any self-perception bias, the presence of a leader trivially improves the welfare of both agents by increasing the ex ante likelihood that coordination will take place (i.e., that they will exert effort at the same time). Of course, team value is then also trivially improved by the leadership structure. For the balance of this section, we explore how the team's organizational structure and the self-perception biases of its agents combine to affect individual welfare and team value. In doing so, we can determine when and why a rational agent makes a better leader than the biased agent and vice versa. Our first result establishes that self-perception biases can only be Pareto-improving when the team leader is rational.

Proposition 3 *With a biased leader, the expected utility of the (biased) leader is decreasing in d, whereas the expected utility of the (rational) follower is increasing in d. With a rational leader, the expected utility of the (rational) leader is increasing in d, while the expected utility of the (biased)* *follower is increasing in d if and only if $d < \frac{b^3}{2} + ab(1+b)$.*

Interestingly, when she follows the rational agent's lead, the overconfident agent's bias can make her better off, provided that her bias is not too extreme. The intuition for this result is the same as before: the rational agent anticipates that her biased teammate will react incorrectly to her lead but, knowing that an overinvestment in effort will follow, she works harder in order to create more synergy between the two of them; this additional synergy feeds back into the biased agent's utility, making her better off. This cannot occur when the biased agent leads because, in that structure, the rational agent's strategy is not a function of her biased teammate's expected effort, but her actual effort. In other words, the rational agent does not work harder to anticipate more synergy; only the frequency with which she uses her $k_{\text{F}1}$ threshold (as opposed to her $k_{\text{F}0}$ threshold) increases. As a result, the biased leader does not benefit from any synergy feedback; that is, her overinvestment in effort is a pure cost to herself.

The decision of the firm to be led by the rational agent or the biased agent is ultimately driven by the firm's value under each scenario. The following proposition sheds more light on the issue. As discussed earlier, in this model, firm value corresponds to team welfare as proper contracting would ensure that agents get their reservation utility.

Proposition 4 *(i) Expected firm output is larger with a rational (biased) leader when d is greater (smaller) than $1-4a-2b$. (ii) Firm value is larger with a rational (biased) leader when d is greater (smaller) than $1-6a-3b-\frac{2a}{b}$.*

As long as the overconfident agent's bias is not too extreme, making her the team leader is beneficial to team output and to team value. To see how the first part of this result works, suppose that we increase the bias of the overconfident agent from 0 to d, while keeping the strategy (i.e., thresholds) of the rational agent unchanged. With the biased leader, k_{L} increases by d, increasing the direct contribution of the biased agent to the probability of success by da, that of the rational agent by $d(k_{\text{F}1}-k_{\text{F}0})a = dba$, and the synergistic contribution of the two agents by $dk_{\text{F}1}b = d(a+b)b$, for a total increase of $d\big[a+b(2a+b)\big]$. With the rational leader, both $k_{\text{F}1}$ and $k_{\text{F}0}$ increase by d. Although this does not affect the rational agent's direct contribution to the probability of success (as k_{L} is assumed to stay the same), it does increase the direct contribution of the biased agent by da, and increases the synergistic contribution of the two agents by $dk_{\text{L}}b = d\big[a+b(2a+b)\big]b$, for a total increase of $d\Big\{a+b\big[a+b(2a+b)\big]\Big\}$. Because $2a+b<1$, the biased-leader scenario is always preferable, *keeping the rational agent's strategy constant.* The value of the rational leader comes from the fact that the rational agent changes her strategy with a change in d only when she leads (as a follower, the rational agent only reacts to the leader's observable effort, not to her expected effort). As shown in the proof to the proposition, this additional contribution to the probability of success is large when the follower's bias is large, yielding the first result.

Part (ii) of Proposition 4 adds effort costs to the tradeoff. As in part (i), the rational agent is a better choice for a leader when d is large enough. However, this occurs for more values of d than before. The reason is simple: the effort gap between the leader and the follower is larger with an overconfident leader and increases with d; given that effort costs are quadratic in thresholds (see footnote 14), reducing this gap by making the rational agent team leader saves on effort costs. Notice also that when b is small relative to a, $1 - 6a - 3b - \frac{2a}{b}$ is negative, and it is then always preferable for the team to be led by the rational agent. Thus, when synergies are small relative to the individual contribution of each agent, rational agents who can anticipate and thus better harness the behavioral motives (overconfidence, intrinsic motivation, the need for achievement, etc.) of other agents make better leaders.

Together, Propositions 3 and 4 represent, as far as we know, the first results on the optimal structure of an organization based on the behavioral characteristics of its agents. Given that the main trust of this paper is not about firm organization per se, we do not pursue this line of thought further. However, we do believe that behavioral considerations are likely to impact the optimal organizational structure of firms. As such, we anticipate this topic to be a fruitful area for future research.

6. Monitoring

As shown in the previous section, the presence of a team leader facilitates coordination between the team's agents by making the effort of one public. In this section, we explore another mechanism that makes effort decisions public, namely monitoring. Such a mechanism has been explored in team contexts before (see, e.g., Aoki, 1994). We study its effectiveness in restoring team surplus in the presence of self-perception biases.

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 (e.g., the firm's owner) 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.[18] 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 biased and believes that her ability is $A = a+d$ although it is truly only a. For most of this section however, we assume that $b = 0$. Thus, our analysis focuses on the joint role of self-perception biases 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 4, 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 firm 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.

[18] As long as the agent who works gets more than the one who doesn't, any sharing rule will lead to the same qualitative results.

B. The Effect of Monitoring

The question we seek to address in the rest of this section is whether self-perception biases complement the monitoring mechanism or reduce 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. 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 4, 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 3 *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

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

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 bias does not affect the tradeoff of working and shirking for agent 1. The extra effort exerted by agent 2 as a result of her bias 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 firm value that is lost to the coordination problems described in section 3. With self-perception biases, 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 biases, 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 5 *The intensity q^* of monitoring that maximizes firm value is*

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

(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 self-perception biases. 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 firm value. 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 5 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. Incidentally, the same figure also shows that q^* is increasing in b: as complementarities across agents increase, monitoring becomes more valuable.

In sum, monitoring and overconfidence are substitutes rather than complements. Free monitoring can be detrimental for a firm when some of its agents are positively biased about their own ability. 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.

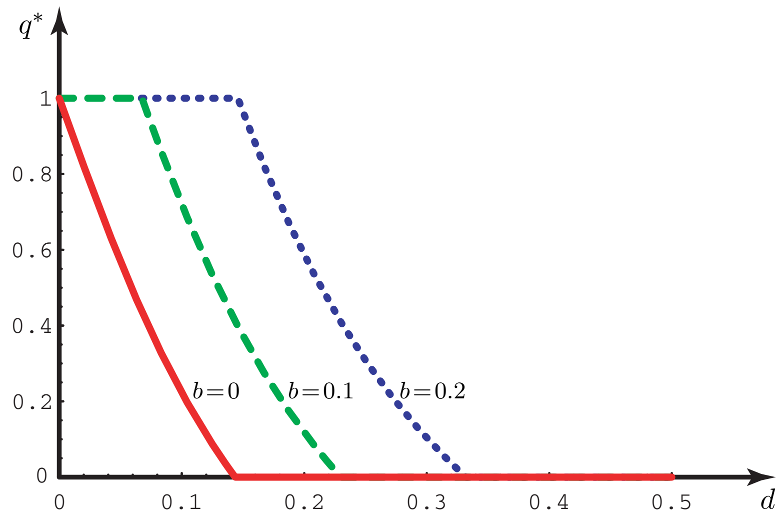

Figure 3: Optimal monitoring as a function of overconfidence.

This figure shows the level of monitoring q^* that maximizes firm value as a function of agent 2's self-perception bias 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.

7. Learning

As shown in section 4, the presence of biased 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 biased self-perceptions 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 the biases 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 biased members are more likely to prosper over time.[19]

A related question is whether, as the team and its members prosper, overconfidence can sustain itself. Overconfidence, like preferences, cannot be faked. In fact, it is crucial that biased agents be unaware of their biases when making decisions, as it is this unawareness that affects their behavior. Thus, if biased agents eventually learn their true skills, the benefits of self-perception

[19]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.

biases disappear, and the team starts suffering from the same coordination problems described in section 3. 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 $[0, 2(a+d)]$.[20] Thus, although each agent's average skill is really a, agent 2 thinks that her average skill is $a+d$ where, as in section 4, d denotes the extent of her self-perception bias. Because the rest of the model is unchanged and because agents are risk-neutral (and so care only about average skills when choosing effort), we can use Lemma 1 directly to obtain the equilibrium strategies of the two agents. The only difference from before is that the agents' beliefs about $\tilde{s}_1$ and $\tilde{s}_2$ will change after the project's outcome is realized. As a result, agent 2's overconfidence will also change at the end of the period. This change in overconfidence is what we focus on in this section. 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 self-perceived expected ability for the second period.[21]

Since her ability gets impounded into the probability of the project being successful only when

[20] 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 $4a + 2d + b \leq 1$.

[21] The results of Propositions 6 and 7 below are obtained under the assumption that agents are myopic when choosing their initial effort threshold. In particular, these results do not account for the fact that exerting more effort allows agents to learn more about their own ability (as they learn nothing new about themselves when they don't work). Incorporating this strategic concern into the model significantly complicates the analysis, because the agents then have to take into account how they will choose their later thresholds depending on earlier effort choices and project outcomes. This makes the number of endogenous thresholds grow exponentially with each period. In fact, even a two-period model proved to be too complicated for analytical solutions to be derived. To assess the robustness of our results, we solve the strategic version of the two-period model numerically (the details of this numerical solution are available from the authors upon request). Because the parameter space for the model's only three parameters, a, b and d, is closed and bounded, it is relatively straightforward to span the whole space numerically. As expected, both agents choose a first-period effort threshold that is higher than the myopic one. Furthermore, Propositions 6 and 7, this section's two main results, are shown to hold, even in this strategic setting.

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 $\alpha \equiv \mathrm{E}_{\mathrm{B}}\left[\tilde{s}_2 \mid e_2 = 1\right]$, 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 a biased agent.

Proposition 6 *On average, the bias 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 direct positive impact of d on this probability. The second, using k_{BM} instead of $k_{\mathrm{BM}} + bd$ for agent 1's effort cost threshold, has an indirect negative impact of $(a + b)bd$ (that is, the marginal productivity of agent 1 times bd). Because this indirect impact is felt to a lesser extent (that is, 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 7 *The end-of-period bias 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 self-perception biases to make team members better off by facilitating coordination and, *at the same time*, makes convergence to rationality more difficult and thus slower. This makes overconfidence a good candidate as an ingredient for long-term team success.

B. Convergence of Beliefs

Proposition 6 shows that the biased agent remains biased 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.[22] We assume that each project has an independent payoff, and that the effort cost of both agents are independent across projects.[23]

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 bias.

Proposition 8 *With an infinite number of projects, the biased agent concludes that her ability is $\tilde{s}_2 + b^2 d$, 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 $b^2 d$ 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 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 8 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

[22] 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. As mentioned in footnote 21, however, such a model would quickly prove to be analytically intractable.

[23] 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.

agent's high skill.[24]

Notice also that b and d combine to make the overconfident agent's learning biased. In other words, self-perception biases and complementarities again go hand in hand: they improve team performance and welfare and, at the same time, they prolong the positive effects of the biases. In fact, as in Proposition 7, it is the case that the ex post bias of agent 2 is increasing in b and that learning is impaired by larger values of b (as the bias is reduced by $(1-b^2)d$).

8. 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 the role of biased self-perceptions in team problems. 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 biased ones, better off. On the one hand, the overinvestment in effort by a biased 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 biased agent's share of that payoff.

[24] Incidentally, the fact that the overconfident agent does not realize that the rational agent works as hard as she really does is also ultimately responsible for her reaching incorrect ex post beliefs about her own skill. Indeed, if agent 2 knew k_1, she would learn agent 1's skill correctly from the projects on which she chooses not to work. Knowing k_1 and $\tilde{s}_1$ would in turn allow agent 2 to correctly infer her own skill from the projects on which she exerts an effort. This is the sense in which our results about learning, including the ones in Propositions 6 and 7 in fact, depend on the assumption that agent 2 does not realize that agent 1 knows about her overconfidence.

The presence of biased agents also creates opportunities for teams to set their organizational structure according to the behavioral characteristics of their agents. In particular, when one agent is responsible for setting the tone through a choice of effort that is made public, it is tempting to conclude that individuals who tend to overinvest in effort make good leaders. This, as we show, is only the case when the extent of their bias is relatively small. Otherwise, the team and its agents are better off with a rational leader who can anticipate the biased effort choices made by her followers.

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.

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.[25] 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 self-perception biases 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.

As we argue throughout the paper, our theory is more general than the role of self-perception biases in teams. First, any behavioral tendency that creates a natural incentive for an agent to invest in effort will have the same positive effects on her and her team as biased self-perceptions. In

[25]It does so for a different reason than it does in Kyle and Wang's (1997) model of overconfident trading. In that model, overconfident trading is suboptimal, but it is the optimal response to rational trading. This is why the welfare implications of our and their models are diametrically opposed.

particular, intrinsic motivation and the need for achievement will also facilitate team coordination when synergies exist across teammates. More than that, the lessons that can be learned from our theory are relevant not only to the various finance applications discussed in section 2 but to a wide range of economic situations.

For example, an alternative interpretation of the model is that of a public good that is more likely to come to fruitful completion or have value if agents contribute their effort to it. In this light, our model shows that overconfidence can mitigate the efficiency loss prompted by free-riding behavior in the allocation of public goods, as originally pointed out by Samuelson (1954). Our model can also be viewed as a description of technological innovations problems. Indeed, coordinated goals and efforts make technological innovations more likely to occur and to benefit its participants. Our model then points out that there may be 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.[26]

In fact, 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. 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. Finally, 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. 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.

[26] 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.

9. References

Aghion, P., and J. Tirole, 1994, "The Management of Innovation," *Quarterly Journal of Economics*, 109, 1185-1209.

Alchian, A. A., and H. Demsetz, 1972, "Production, Information Costs, and Economic Organization," *American Economic Review*, 62, 777-795.

Alchian, A., and S. E. Woodward, 1987, "Reflections on the Theory of the Firm," *Journal of Institutional and Theoretical Economics*, 143, 110-136.

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.

Aoki, M., 1994, "The Contingent Governance of Teams: Analysis of Institutional Complementarity," *International Economic Review*, 35, 657-676.

Armor, D. A., and S. E. and Taylor, 1998, "Situated Optimism: Specific Outcome Expectancies and Self-Regulation," in *Advances in Experimental Social Psychology*, 30, eds. M. P. Zanna, New York: Academic Press, 309-379.

Athey, S., and J. Roberts, 2001, "Organizational Design: Decision Rights and Incentive Contracts," *American Economic Review*, 91, 200-205.

Atkinson, J. W., 1978, *An Introduction to Motivation*, 2nd Edition, New York: Van Nostrand.

Baron, J. N., 1988, "The Employment Relation as a Social Relation," *Journal of the Japanese and International Economies*, 2, 492-525.

Baron, J. N., and D. M. Kreps, 1999, *Strategic Human Resources: Frameworks for General Managers*, New York: John Wiley.

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.

Boot, A. W., and A. Schmeits, 2000, "Market Discipline and Incentive Problems in Conglomerate Firms with Applications to Banking," *Journal of Financial Intermediation*, 9, 240-273.

Campbell, D. T., 1986, "Rationality and Utility from the Standpoint of Evolutionary Biology," *Journal of Business*, 59, S355-S364.

Casamatta, C., 2003, "Financing and Advising: Optimal Financial Contracts with Venture Capitalists," *Journal of Finance*, 58, 2059-2085.

Cooper, A. C., C. Y. Woo, and W. C. Dunkelberg, 1988, "Entrepreneurs' Perceived Chances for Success," *Journal of Business Venturing*, 3, 97-108.

Deci, E. L., 1975, *Intrinsic Motivation*, Plenum Press: New York.

Dow, J., 2004, "Is Liquidity Self-Fulfilling?" *Journal of Business*, 77, 895-908.

Dunning, D., J. A. Meyerowitz, and A. D. Holzberg, 1989, "Ambiguity and Self-Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability," *Journal of Personality and Social Psychology*, 57, 1082-1090.

Eshel, I., L. Samuelson, and A. Shaked, 1998, "Altruists, Egoists, and Hooligans in a Local Interaction Model," *American Economic Review*, 88, 157-179.

Farrell, J., and S. Scotchmer, 1988, "Partnerships," *Quarterly Journal of Economics*, 103, 279-297.

Faulí-Oller, R., and M. Giralt, 1995, "Competition and Cooperation Within a Multidivisional Firm," *Journal of Industrial Economics*, 43, 77-99.

Feather, N. T., 1965, "The Relationship of Expectation of Success to Need Achievement and Test Anxiety," *Journal of Personality and Social Psychology*, 1, 118-126.

Fischhoff, B., P. Slovic, and S. Lichtenstein, 1977, "Knowing with Certainty: The Appropriateness of Extreme Confidence," *Journal of Experimental Psychology*, 3, 552-564.

Galbraith, J. R., 1977, *Organization Design*, Reading, MA: Addison-Wesley.

Garen, J. E., 1985, "Worker Heterogeneity, Job Screening, and Firm Size," *Journal of Political Economy Volume*, 93, 715-739.

Gaynor, M., and P. Kleindorfer, 1991, "Equilibrium Misperceptions," *Economics Letters*, 35, 27-30.

Greenwald, A. G., 1980, "The Totalitarian Ego: Fabrication and Revision of Personal History," *American Psychologist*, 35, 603-618.

Grossman, S. J., and O. D. Hart, 1986, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691-719.

Groves, T., 1973, "Incentives in Teams," *Econometrica*, 41, 617-631.

Heath, C., R. P. Larrick, and G. Wu, 1999, "Goals as Reference Points," *Cognitive Psychology*, 38, 79-109.

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.

Hellmann, T., and M. Puri, 2002, "Venture Capital and the Professionalization of Start-Up Firms: Empirical Evidence," *Journal of Finance*, 57, 169-197.

Hermalin, B. E., 1998, "Toward an Economic Theory of Leadership: Leading by Example," *American Economic Review*, 88, 1188-1206.

Hermalin, B. E., and M. S. Weisbach, 1998, "Endogenously Chosen Boards of Directors and Their Monitoring of the CEO," *American Economic Review*, 88, 96-118.

Holmström, B., 1982, "Moral Hazard in Teams," *Bell Journal of Economics*, 13, 324-340.

Holmström, B., and J. Roberts, 1998, "The Boundaries of the Firm Revisited," *Journal of Economic Perspectives*, 12, 73-94.

Itoh, H., 1991, "Incentives to Help in Multi-Agent Situations," *Econometrica*, 59, 611-636.

Jensen, M. C., and W. H. Meckling, 1995, "Specific and General Knowledge and Organizational Structure," *Journal of Applied Corporate Finance*, 8, 4-18.

Kamien, M. I., E. Muller, and I. Zang, 1992, "Research Joint Ventures and R&D Cartels," *American Economic Review*, 82, 1293-1306.

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

Kaplan, S. N., and P. Strömberg, 2004, "Characteristics, Contracts, and Actions: Evidence from Venture Capitalist Analyses," *Journal of Finance*, 59, 2177-2210.

Kelsey, D., and W. Spanjers, 2004, "Ambiguity in Partnerships," *Economic Journal*, 114, 528-546.

Kohn, A., 1999, *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*, Boston: Houghton Mifflin Co.

Kreps, D. M., 1997, "Intrinsic Motivation and Extrinsic Incentives," *American Economic Review*, 87, 359-364.

Kruglanski, A. W., 1975, "The Endogenous-Exogenous Partition in Attribution Theory," *Psychological Review*, 82, 387-406.

Kunda, Z., 1987, "Motivated Inference: Self-Serving Generation and Evaluation of Causal Theories," *Journal of Personality and Social Psychology*, 53, 636-647.

Kyle, A. S., and F. A. Wang, 1997, "Speculation Duopoly With Agreement to Disagree: Can Overconfidence Survive the Market Test?" *Journal of Finance*, 52, 2073-2090.

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.

Larwood, L., and W. Whittaker, 1977, "Managerial Myopia: Self-Serving Biases in Organizational Planning," *Journal of Applied Psychology*, 62, 194-198.

Lazear, E. P., 2000, "Performance Pay and Productivity," *American Economic Review*, 90, 1346-1361.

Lazear, E. P., and S. Rosen, 1981, "Rank-Order Tournaments as Optimum Labor Contracts Journal," *Journal of Political Economy*, 89, 841-864.

Lerner, J., 1995, "Venture Capitalists and the Oversight of Private Firms," *Journal of Finance*, 50, 301-318.

Levin, J., and S. Tadelis, 2005, "Profit Sharing and the Role of Professional Partnerships," *Quarterly Journal of Economics*, 120, 131-171.

Lewis, T. R., and D. E. M. Sappington, 1997, "Penalizing Success in Dynamic Incentive Contracts: No Good Deed Goes Unpunished?" *RAND Journal of Economics*, 28, 346-358.

Locke, E. A., and G. P. Latham, 1990, *A Theory of Goal Setting and Task Performance*, Englewood Cliffs, N.J.: Prentice Hall.

McAfee, R. P., and J. McMillan, 1991, "Optimal Contracts for Teams," *International Economic Review*, 32, 561-577.

Meyer, H. H., W. B. Walker, and G. H. Litwin, 1961, "Motive Patterns and Risk Preferences Associated with Entrepreneurship," *Journal of Abnormal and Social Psychology*, 63, 570-574.

Morrison, A. D., and W. J. Wilhelm, 2004, "Partnership Firms, Reputation, and Human Capital," *American Economic Review*, 94, 1682-1692.

Pichler, P., and W. Wilhelm, 2001, "A Theory of the Syndicate: Form Follows Function," *Journal of Finance*, 56, 2237-2264.

Prendergast, C., 1999, "The Provision of Incentives in Firms," *Journal of Economic Literature*, 37, 7-63.

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.

Sahlman, W. A., 1990, "The Structure and Governance of Venture-Capital Organizations," *Journal of Financial Economics*, 27, 473-521.

Samuelson, P. A., 1954, "The Pure Theory of Public Expenditure," *Review of Economics and Statistics*, 36, 387-389.

Scharfstein, D. S., and J. C. Stein, 2000, "The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment," *Journal of Finance*, 55, 2537-2564.

Schor, J. B., 1993, *The Overworked American: The Unexpected Decline of Leisure Edition*, New York: Basic Books.

Spence, M., 1973, "Job Market Signaling," *Quarterly Journal of Economics*, 87, 355-374.

Stein, J. C., 1997, "Internal Capital Markets and the Competition for Corporate Resources," *Journal of Finance*, 52, 111-133.

Sufi, A., 2005, "Information Asymmetry and Financing Arrangements: Evidence from Syndicated Loans," Working Paper, MIT.

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.

Weiner, B., 1985, *Human Motivation*, New York: Springer-Verlag.

10. Appendix A

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

$$k_{22} = k_{\text{BM}} + d. \tag{13}$$

Now, agent 1 knows that agent 2 is biased 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

$$k_{11} = k_{\text{BM}} + bd. \tag{14}$$

This completes the proof. ∎

Proof of Proposition 1

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

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

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

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

Differentiation of this last expression with respect to d yields

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

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

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

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

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

Differentiation of this last expression with respect to d yields

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

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

Proof of Proposition 2

Using (15) and (16), we have

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

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

Proof of Lemma 2

Biased leader. If the biased agent works, the rational agent's expected utility is $2a + b - \tilde{c}_1$ if she works, and a if she does not. Her effort cost threshold is therefore $k_{\mathrm{F1}} = a + b$. If the biased agent does not work, the rational agent's expected utility is $a - \tilde{c}_1$ if she works, and 0 if she does not. Thus her effort cost threshold is $k_{\mathrm{F0}} = a$. Taking these subsequent thresholds into account, the biased agent's expected utility is $a + d + (a+b)^2 - \tilde{c}_2$ if she works, and a^2 if she does not work. Her threshold cost of effort is therefore $k_{\mathrm{L}} = a + d + b(2a + b)$.

Rational leader. If the rational agent works, the biased agent's expected utility is $2a + d + b - \tilde{c}_2$ if she works, and a if she does not. Her effort cost threshold is therefore $k_{\mathrm{F1}} = a + d + b$. If the rational agent does not work, the biased agent's expected utility is $a + d - \tilde{c}_2$ if she works, and 0 if she does not. Thus her effort cost threshold is $k_{\mathrm{F0}} = a + d$. Taking these subsequent thresholds into account, the rational agent's expected utility is $a + (a+b)^2 + d(a+b) - \tilde{c}_1$ if she works, and $a^2 + ad$ if she does not work. Her threshold cost of effort is therefore $k_{\mathrm{L}} = a + b(2a + b + d)$. ∎

Proof of Proposition 3

With a leader, the expected payoff of the team's project is given by

$$\mathrm{E}[\tilde{v}] = 2\Big(ak_{\mathrm{L}} + a\big[k_{\mathrm{L}}k_{\mathrm{F1}} + (1 - k_{\mathrm{L}})k_{\mathrm{F0}}\big] + bk_{\mathrm{L}}k_{\mathrm{F1}}\Big). \qquad (17)$$

The expected effort cost of the leader is

$$\mathrm{E}[\tilde{c}_{\mathrm{L}}] = \frac{k_{\mathrm{L}}^2}{2}, \tag{18}$$

whereas that of the follower is

$$\mathrm{E}[\tilde{c}_{\mathrm{F}}] = k_{\mathrm{L}}\frac{k_{\mathrm{F1}}^2}{2} + (1-k_{\mathrm{L}})\frac{k_{\mathrm{F0}}^2}{2}. \tag{19}$$

Biased leader. We can use the effort thresholds of Lemma 2 in (17), (18) and (19) to calculate the expected utility of the two agents. After some manipulations, the biased leader's expected utility is given by

$$\mathrm{E}\big[\tilde{U}_{\mathrm{L}}\big] = \frac{1}{2}\mathrm{E}[\tilde{v}] - \mathrm{E}[\tilde{c}_{\mathrm{L}}] = \frac{1}{2}\Big[a^2(3+4b+4b^2) + 2ab^2(1+2b) + b^4 - d^2\Big], \tag{20}$$

and that of the rational follower is given by

$$\mathrm{E}\big[\tilde{U}_{\mathrm{F}}\big] = \frac{1}{2}\mathrm{E}[\tilde{v}] - \mathrm{E}[\tilde{c}_{\mathrm{F}}] = \frac{1}{2}\Big[a^2(3+6b+4b^2) + a(4b^3+3b^2+2bd+2d) + b^2(b^2+d)\Big]. \tag{21}$$

Clearly (20) is decreasing in d, while (21) is increasing in d.

Rational leader. Similar manipulations yield the following expected utilities for the rational leader and biased follower:

$$\mathrm{E}\big[\tilde{U}_{\mathrm{L}}\big] = \frac{1}{2}\mathrm{E}[\tilde{v}] - \mathrm{E}[\tilde{c}_{\mathrm{L}}] = \frac{1}{2}\Big[a^2(3+4b+4b^2) + 2a(d+bd+b^2+2b^2d+2b^3) + b^2(b+d)^2\Big], \tag{22}$$

and

$$\mathrm{E}\big[\tilde{U}_{\mathrm{F}}\big] = \frac{1}{2}\mathrm{E}[\tilde{v}] - \mathrm{E}[\tilde{c}_{\mathrm{F}}] = \frac{1}{2}\Big[a^2(3+6b+4b^2) + ab(3b+4b^2+2bd+2d) + b^4 + b^3d - d^2\Big]. \tag{23}$$

Clearly (22) is increasing in d, whereas

$$\frac{\partial}{\partial d}\mathrm{E}\big[\tilde{U}_{\mathrm{F}}\big] = \frac{1}{2}\Big[2ab(1+b) + b^3 - 2d\Big],$$

which is positive if and only if $d < \frac{b^3}{2} + ab(1+b)$. ∎

Proof of Proposition 4

(i) To calculate the expected team output, we can use the effort thresholds of Lemma 2 in (17). After some manipulations, we find that expected team output is given by

$$\mathrm{E}[\tilde{v}] = 2\Big[2a^2(1+2b+2b^2) + a(1+2b)(2b^2+d) + b^2(b^2+d)\Big]$$

with a biased leader, and by

$$\mathrm{E}[\tilde{v}] = 2\Big[2a^2(1+2b+2b^2) + a(4b^3+2b^2+4b^2d+2bd+d) + b^2(b+d)^2\Big]$$

with a rational leader. After subtracting the first of these expressions from the second one, and performing a few simple algebraic steps, we find that the expected team output with a rational leader exceeds that with a biased leader by

$$2b^2d\,(4a+2b-1+d)\,,$$

which is positive if and only if $d > 1-4a-2b$.

(ii) Team value under the biased-leader structure can be obtained by summing up (20) and (21). Team value under the rational-leader structure can be obtained by summing up (22) and (23). After taking the difference and simplifying, we find that team value with a rational leader exceeds team value with a biased leader by

$$\frac{1}{2}b^2d\left(6a+3b+\frac{2a}{b}-1+d\right),$$

which is positive if and only if $d > 1-6a-3b-\frac{2a}{b}$. ∎

Proof of Lemma 3

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 5

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

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

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

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

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:

$$\begin{aligned}
\frac{\partial V(q,d)}{\partial d} &= (2a-k_{11}-\kappa)(1-k_{11}-aq)-\frac{\kappa}{d}a-\kappa(1-k_{11}-aq) \\
&= (2a-k_{11})\cdot(1-k_{11}-aq)-a(1+q(1-k_{11}))-2\kappa(1-k_{11}-aq) \\
&= a-2ak_{11}-2a^2q-k_{11}+k_{11}^2+2k_{11}aq-aq-2\kappa(1-k_{11}-aq) \\
&= a-2a^2-4a^2q-a-aq+a^2+2a^2q+a^2q^2+2a^2q+2a^2q^2 \\
&\quad -aq-2\kappa(1-k_{11}-aq) \\
&= -a^2-2aq(1-\kappa)+3a^2q^2-2\kappa(1-k_{11}) \\
&= -a^2\left(1-q^2\right)-2aq(1-\kappa-aq)-2\kappa(1-k_{11}) \qquad (24)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial V(q,d)}{\partial q} &= -2a^2+d\begin{bmatrix} -\big[a+d(1-k_{11}-aq)\big](1-k_{11}-aq) \\ -2a(2a-k_{11}-\kappa)-a(1-k_{11}-aq) \end{bmatrix} \\
&= -2a^2+d\begin{bmatrix} -\big[2a+d(1-k_{11}-aq)\big](1-k_{11}-aq) \\ -2a(2a-k_{11}-\kappa) \end{bmatrix} \\
&= -2a^2+d\Big[-d(1-k_{11}-aq)^2-2a(1-\kappa-3aq)\Big] \\
&= -2a^2(1-dq)+d\Big[-d(1-k_{11}-aq)^2-2a(1-\kappa-2aq)\Big]. \qquad (25)
\end{aligned}$$

Because $k_{11}+\kappa<1$, we have $1-\kappa-aq>0$ and $1-\kappa-2aq>0$, which imply that (24) and (25) are both negative. By the implicit function theorem, this further implies that the optimal monitoring intensity q^* is decreasing in d. ∎

Proof of Proposition 6

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

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

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

$$\mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2=1,\tilde{v}=2\big]=a+d+\frac{(a+d)^2}{3\big[a+d+(a+b)k_{\mathrm{BM}}\big]}. \qquad (26)$$

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

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

which simplifies to

$$\mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 0\big] = a + d - \frac{(a+d)^2}{3\big[1 - a - d - (a+b)k_{\mathrm{BM}}\big]}. \tag{27}$$

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

$$\Pr\big\{\tilde{v} = 2 \mid e_2 = 1\big\} = a + ak_{11} + bk_{11} = a + (a+b)(k_{\mathrm{BM}} + bd). \tag{28}$$

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

$$\begin{aligned} \alpha &\equiv \mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1\big] \\ &= \Pr\big\{\tilde{v} = 2 \mid e_2 = 1\big\}\mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 2\big] \\ &\quad + \Big(1 - \Pr\big\{\tilde{v} = 2 \mid e_2 = 1\big\}\Big)\mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 0\big], \end{aligned} \tag{29}$$

which, using (26)-(28), can be manipulated to yield

$$\alpha = a + d - \frac{d(a+d)^2\big[1 - (a+b)b\big]}{3\big[a + d + (a+b)k_{\mathrm{BM}}\big]\big[1 - a - d - (a+b)k_{\mathrm{BM}}\big]}. \tag{30}$$

Because the last term in this expression is clearly larger than zero, we have $\alpha < a + d$. To establish that $\alpha > a$, first notice from (30) that $\alpha = a$ when $d = 0$. Thus we only need to show that $\frac{\partial \alpha}{\partial d} > 0$. Using (29), we can write

$$\begin{aligned} \frac{\partial \alpha}{\partial d} &= \Pr\big\{\tilde{v} = 2 \mid e_2 = 1\big\}\frac{\partial \mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 2\big]}{\partial d} \\ &\quad + \Big(1 - \Pr\big\{\tilde{v} = 2 \mid e_2 = 1\big\}\Big)\frac{\partial \mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 0\big]}{\partial d} \\ &\quad + \Big(\mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 2\big] - \mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 0\big]\Big)\frac{\partial \Pr\big\{\tilde{v} = 2 \mid e_2 = 1\big\}}{\partial d}. \end{aligned} \tag{31}$$

From (26) and (27), it is clear that $\mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 2\big] > \mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 0\big]$ and, from (28), we have

$$\frac{\partial \Pr\big\{\tilde{v} = 2 \mid e_2 = 1\big\}}{\partial d} = (a+b)b > 0.$$

Thus the last line of (31) is positive, and so the result will be established if we can show that $\frac{\partial \mathrm{E}_{\mathrm{B}}\left[\tilde{s}_2 \mid e_2=1, \tilde{v}=2\right]}{\partial d}$ and $\frac{\partial \mathrm{E}_{\mathrm{B}}\left[\tilde{s}_2 \mid e_2=1, \tilde{v}=0\right]}{\partial d}$ are positive. Differentiation of (26) with respect to d, followed by some manipulations, yields

$$\frac{\partial \mathrm{E}_{\mathrm{B}}\big[\tilde{s}_2 \mid e_2 = 1, \tilde{v} = 2\big]}{\partial d} = 1 + \left(\frac{a+d}{3}\right)\frac{a + d + 2(a+b)k_{\mathrm{BM}}}{\big[a + d + (a+b)k_{\mathrm{BM}}\big]^2},$$

which is clearly positive. Differentiation of (27) with respect to d, followed by some manipulations, yields

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

This last expression is positive if

$$0 < \frac{a+d}{1-(a+d)-(a+b)k_{\mathrm{BM}}} < 1.$$

The first inequality is obvious. The second inequality is equivalent to $2(a+d)+(a+b)k_{\mathrm{BM}} < 1$, which is implied by our assumption that $4a+2d+b \leq 1$ (see footnote 20), as

$$2(a+d)+(a+b)k_{\mathrm{BM}} < 4a+2d+b.$$

This completes the proof. ∎

Proof of Proposition 7

We need to show that $\frac{\partial \alpha}{\partial b} > 0$. Differentiation of (30) with respect to b yields, after some manipulations,

$$\frac{\partial \alpha}{\partial b} = \frac{d(a+d)^2}{3P^2(1-P)^2}\left\{(a+2b)P(1-P) + \left[1-(a+b)b\right](1-2P)\frac{a(1+a)}{(1-b)^2}\right\},$$

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

$$b < \frac{1-2(a+d)-2a^2}{1-2d}. \tag{32}$$

Notice that the right-hand side of this last expression satisfies

$$\begin{aligned}\frac{1-2(a+d)-2a^2}{1-2d} &> 1-2(a+d)-2a^2\\ &= 1-4a-2d+2a-2a^2\\ &= 1-4a-2d+2a(1-a)\\ &> 1-4a-2d.\end{aligned}$$

Thus condition (32) is implied by our assumption that $4a+2d+b \leq 1$ (see footnote 20), which is equivalent to $b < 1-4a-2d$. ∎

Proof of Proposition 8

From the perspective of the biased 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

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

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 $\tilde{s}_1' k_{\mathrm{BM}}$, where $\tilde{s}_1'$ denotes the realized skill of agent 1 as perceived by agent 2. As a result, the biased agent (wrongfully) infers that the skill of her teammate is[27]

$$\tilde{s}_1' = \tilde{s}_1 \left(1 + \frac{bd}{k_{\mathrm{BM}}}\right).$$

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

$$\tilde{s}_1 k_{11} + \tilde{s}_2 + bk_{11} = \tilde{s}_1 \left(k_{\mathrm{BM}} + bd\right) + \tilde{s}_2 + b\left(k_{\mathrm{BM}} + bd\right) = \tilde{s}_1' k_{\mathrm{BM}} + \tilde{s}_2 + b\left(k_{\mathrm{BM}} + bd\right),$$

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

[27] Note that it is possible that $\tilde{s}_1'$ falls outside the $[0, 2a]$ support for $\tilde{s}_1$. This minor inconsistency could be easily fixed by assuming that the overconfident agent assigns a small probability that agent 1's skill is above $2a$. Because the added complexity of doing so contributes nothing to the economics of the paper, we ignore this technical detail in our analysis.

11. Appendix B

This appendix re-derives the results of section 4 under the alternative assumption that the overconfident agent (agent 2) knows that her teammate (agent 1) considers her to be overconfident. Because many of the steps of the derivation are similar to those contained in the proofs to the results of section 4, we keep the analysis short by skipping some of these steps. Before we turn to the statements and proofs of these results, it is worth noting that, because the agents now know each other's beliefs, it is no longer necessary to use k_{21} and k_{12}. Indeed, it is always the case that $k_{21} = k_{11}$ and $k_{12} = k_{22}$. We begin with the analogue to Lemma 1 under this alternative assumption.

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

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

$$k_{11} = k_{\text{BM}} + \frac{bd}{1-b^2};$$

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

$$k_{22} = k_{\text{BM}} + \frac{d}{1-b^2}.$$

Proof. Using the same arguments as in section 3 and the fact that agent 2 thinks that he ability is $A = a + d$, we have

$$\begin{aligned} k_{11} &= a + bk_{22}, \quad \text{and} \\ k_{22} &= a + d + bk_{11}. \end{aligned}$$

Solving for k_{11} and k_{22} in these equations, we find

$$k_{11} = \frac{a + (a+d)b}{1-b^2}, \quad \text{and} \tag{33}$$

$$k_{22} = \frac{a + d + ab}{1-b^2}, \tag{34}$$

which can be rewritten as in the lemma's statement above. ∎

Notice that, as before, both k_{11} and k_{22} are increasing in d, that is, both agents work harder as a result of agent 2's bias. The following two propositions, the analogues to Propositions 1 and 2, show that the same welfare results obtain under the assumption that agent 2 knows the beliefs of agent 1.

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

(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

$$d < \frac{ab(1+b)}{1-2b^2}. \tag{35}$$

Proof.

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

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

Using (33) and (34) in this expression yields, after some manipulations,

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

Differentiation of this last expression with respect to d yields

$$\frac{\partial \bar{U}_1}{\partial d} = \frac{a(1+b) + b^2 d}{(1-b^2)^2} > 0. \tag{36}$$

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

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

Using (33) and (34) in this expression yields, after some manipulations,

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

Differentiation of this last expression with respect to d yields

$$\frac{\partial \bar{U}_2}{\partial d} = \frac{ab(1+b) - (1-2b^2)d}{(1-b^2)^2}. \tag{37}$$

This expression is positive if and only if (35) holds. ∎

As in section 4 therefore, agent 1 is always made better off by agent 2's bias, and agent 2 is better off only if her bias is not too extreme. The reason is the same: agent 2 works harder as a result of her bias, and this makes agent 1 react with a more sustained effort, which in turn benefits agent 2. As long as the extent of her overinvestment in effort is not too large, agent 2's net benefit is positive.

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

$$d < \frac{a(1+b)^2}{1-3b^2}. \tag{38}$$

Proof. Using (36) and (37), we have

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

This expression is positive if and only if (38) holds. ∎

Finally, as before, team welfare increases with agent 2's bias up to a certain point. When agent 2's overinvestment in effort becomes too large, agent 2's welfare loss becomes larger than agent 1's welfare gain.