



Working Paper 05-52
Business Economics Series 14
September 2005

Departamento de Economía de la Empresa
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34-91) 6249608

TEAM FORMATION AND BIASED SELF-ATTRIBUTION¹

Brice Corgnet²

Abstract

We analyze the impact of individuals' self-attribution biases on the formation of teams in the workplace. We consider a two periods model in which workers jointly decide whether to form a team or work alone. We assume workers' abilities are unknown. Agents update their beliefs about abilities after receiving a signal at the end of the first period. We show that allowing workers to learn about their abilities undermines cooperation when a fixed allocation of the group outcome is assumed. Consistent with the latter finding, we establish that making learning about workers' abilities less accessible increases workers' cooperation and welfare. When workers suffer from self-serving attribution, cooperation among agents is undermined whatever the allocation rule considered for the group outcome. We analyze possible solutions to insufficient teamwork. We find that team contracts based on a revelation game can improve cooperation as well as the presence of a manager in the team. Full efficiency is however never achieved. Our paper establishes a basic framework to analyze necessary psychological conditions for individuals to form teams. We apply our model to coauthorship and to organizational issues.

JEL codes: C72, D23, D82, K12, M12, M14, M54

Keywords: Biased self-attribution, team formation, behavioral contract theory

¹ I am especially grateful to my advisor Luis Ubeda. I am indebted as well to Pablo Ruiz-Verdu, Mikel Tapia, and to seminar participants at Universidad Carlos III (January 2005), IMEBE in Cordoba (December 2004) and at the workshop in Behavioral Economics in Budapest (July 2004).

² Ph.D. in Economics, Universidad Carlos III, email: bcorgnet@est-econ.uc3m.es.

Biased self-attribution and team formation¹

Brice Corgnet²

First version: December 2004

This version: August 2005

JEL codes: C72, D23, D82, K12, M12, M14, M54

Keywords: Biased self-attribution, team formation, behavioral contract theory

¹I am especially grateful to my advisor Luis Ubeda. I am indebted as well to Pablo Ruiz-Verdu, Mikel Tapia, and to seminar participants at Universidad Carlos III (January 2005), IMEBE in Cordoba (December 2004) and at the workshop in Behavioral Economics in Budapest (July 2004).

²Ph.D. in Economics, Universidad Carlos III, email: bcorgnet@est-econ.uc3m.es.

Abstract

We analyze the impact of individuals' self-attribution biases on the formation of teams in the workplace. We consider a two periods model in which workers jointly decide whether to form a team or work alone. We assume workers' abilities are unknown. Agents update their beliefs about abilities after receiving a signal at the end of the first period. We show that allowing workers to learn about their abilities undermines cooperation when a fixed allocation of the group outcome is assumed. Consistent with the latter finding, we establish that making learning about workers' abilities less accessible increases workers' cooperation and welfare. When workers suffer from self-serving attribution, cooperation among agents is undermined whatever the allocation rule considered for the group outcome. We analyze possible solutions to insufficient teamwork. We find that team contracts based on a revelation game can improve cooperation as well as the presence of a manager in the team. Full efficiency is however never achieved. Our paper establishes a basic framework to analyze necessary psychological conditions for individuals to form teams. We apply our model to coauthorship and to organizational issues.

1 Introduction

We analyze the impact of individuals' self-serving attribution on the formation of teams in the workplace. Workers suffering from self-serving biases tend to learn more positively about their ability than about others' abilities, for example by ignoring with some probability negative signals about their own ability.¹ The issue in our framework is that the information revealed about team members' abilities will be interpreted differently by coworkers suffering from self-attribution biases. As a result, conflicts in beliefs will arise and lead teams to split. Anticipating such an effect, individuals may not want to form teams. We consider a two periods model in which workers jointly decide whether to form a team or work alone. We assume workers' abilities are unknown. Agents update their beliefs about abilities after receiving a signal at the end of the first period. We find that as long as the group outcome is equally split among workers, the formation of teams in the workplace is less likely when agents have the possibility to learn about their abilities. This is the case since a rigid allocation rule like equal splitting does not provide the adequate incentives for continuing with the team project when coworkers receive asymmetric signals about their abilities. The worker receiving good news about his own ability will ask for a bigger share of the group outcome at time 1. When workers suffer from self-serving attribution, cooperation among agents is undermined whatever the allocation rule considered for the group outcome. As a result, no renegotiation-proof contracts can be designed to implement the efficient teams outcome (*ETO*). This is the case since in the presence of self-serving biases workers may not agree on the perceived ability of coworkers. Consequently, allocation rules permitting to reach a maximum level of cooperation in the case of rational workers may lead agents to work separately in the presence of self-serving biases even if there exists positive synergies for working as a team. Workers' welfare is negatively affected by team members' learning biases since optimal teams are not always formed as a result. The argument is that agents may not work as a group in the first period because they anticipate that biased coworkers (or they themselves) will more frequently leave the group after learning the outcome of the team project at the end of the first period. Ending the team in the second period implies a cost for the team workers since they will lose the positive effect on the marginal product of effort due to undertaking twice the same project. We analyze possible solutions to insufficient team formation and show how cooperation can be improved by team contracts based on a revelation mechanism or by the introduction of a manager in the team. We stress as well the importance of controlling the flows of information accessible to workers and the relevance of investments in social capital. Our paper establishes a basic framework to analyze necessary psychological conditions for individuals to form teams. We apply our model to coauthorship and organizations.

¹In the rest of the paper we will use equivalently the terms self-serving attribution and self-serving learning.

A number of researchers have studied the role of behavioral factors as a possible solution to free riding arising in teams when efforts of the group members are not observable.² Rotemberg (1994) shows how altruism can improve workers' cooperation and welfare when complementarities exist between team members. Kandel and Lazear (1992) show how peer pressure can increase cooperation among workers. This is the case because exerting more effort workers can reduce the negative effects of peer pressure. Although improving workers' cooperation, a high level of peer pressure in a team can reduce workers' expected utility. Gervais and Goldstein (2004) find that workers' biased self-perception facilitates cooperation among agents. The argument is that an overconfident agent overestimates his marginal product of effort leading himself and his coworker to exert more effort in the team. As long as the degree of self-confidence of the biased coworker is not too high, both the rational and the overconfident workers benefit from an increase in the degree of biased self-perception of the overconfident agent.

Our paper differs from the ones previously mentioned since by assuming observability of coworkers' actions we eliminate free riding issues. We consider the most favorable case for workers' cooperation considering teams with a sufficiently close level of collaboration such that agents' performances and actions are observable. Given favorable conditions for workers' cooperation: no moral hazard and asymmetry of information issues, we analyze in sections 3 to 5 (except 5.3) how learning about agents' abilities affects the formation of teams in the workplace. Issues related to private information are analyzed in sections 5.3 and 6. In opposition to Gervais and Goldstein (2004), we establish that workers' biased self-perception has a negative impact on cooperation in the workplace. In our work, we analyze cooperation defined as the decision of team formation whereas Gervais and Goldstein analyzes the level of effort undertaken by agents assuming a team is formed.

Other recent works have started to analyze team issues in the absence of moral hazard. Eaton and Hollis (2003) analyze teamwork when agents have private information about their own opportunities. They show that asymmetry of information among team members leads to insufficient teamwork and they justify the use of incentives schemes that "over-reward" joint work. Ishida (2005) focus on information asymmetry between the principal and team members. At each period, team members observe if an opportunity for collaboration is available. The principal is not able to observe such information and is then unable to interpret the non collaboration among workers as the absence of an opportunity for collaboration. In a repeated setting, the author justifies the use of low-powered incentives for individual work from the point of view of the principal. Under such incentives schemes the principal creates a motive for workers to build a reputation for being cooperative and facilitates peer monitoring. A very short summary of the issues considered in articles analyzing team issues is available in the next table.

²Free riding issues in teams have been studied in numerous articles such as Holmstrom 1982, Itoh 1991 or Che and Yoo 2001

Assumptions / Authors	Hidden actions	Hidden information	Behavioral elements
Holmstrom (1982) Itoh (1991) Che and Yoo (2001)	×		
Eaton and Hollis (2003) Ishida (2005)		×	
Kandel and Lazear (1992) Rotemberg (1994) Gervais and Goldstein (2004)	×		×
Corgnet (2005), sections 3, 4, 5.2 and 5.4			×
Corgnet (2005), sections 5.3 and 6		×	×

Focusing on agents self-serving attribution is motivated by a widespread evidence in Psychology literature showing that people tend to take credit for successes but deny responsibility for failure (Bradley 1978, Miller and Ross 1975, Zuckerman 1979). Individuals tend to process information in a distorted and motivated way in order to enhance one's self-perception (Fiske and Taylor 1991, Nisbett and Ross 1980). As noted by Gilbert et al.(1998):

"Psychologists from Freud to Festinger have described the artful methods by which the human mind ignores, augments, transforms and rearranges information in its unending battle against the affective consequences of negative events."³

There is an extensive evidence as well that people recall their successes better than their failures (Korner 1950, Silverman 1964, Mischel, Ebbesen and Zeiss 1976). Such a selective memory leads agents to hold excessively positive beliefs about themselves (Greenwald 1980) and to see themselves as above average (Svenson 1981 and Cooper, Woo and Dunkelberg 1988). Biases in inference and attribution have been mostly interpreted in the psychology literature as motivational biases. Agents are considered to feel better-off learning in a way that improves their self-image. As an example of motivated learning, we consider in section 3 workers ignoring negative signals about their abilities with some positive probability whereas positive signals are correctly processed. We assume that agents update as a Bayesian inferer the information actually processed, this corresponds to "quasi-Bayesian" learning as defined in Rabin (2002).⁴ Such learning rules in which agents misread or misremember the signals they observe but apply Bayes' rule have been used for example to explain financial anomalies and in particular over- and underreaction in asset prices (Barberis, Shleifer and Vishny 1998, Daniel and Hirshleifer

³I found this quotation thanks to Roland Bénabou.

⁴"A person is modeled as having a specific form of misreading of the world meant to correspond to a heuristic error, but then is assumed to operate as a Bayesian given this misreading." Rabin (2002).

1998). Rabin and Schrag (1999) have used quasi-Bayesian learning to analyze the effect of confirmatory bias, they show how individuals suffering from such a bias can become overconfident.

In section 2 we present and solve our model in the case of rational coworkers. We analyze the model with biased self-attribution in the third section. In section 4, we study the impact of information on team formation and coworkers' expected welfare. In section 5, we consider the issue of bargaining. We analyze in section 6, how team contracts can be used to foster team formation. Section 7 focuses on the role of team managers. We discuss applications of our model in section 8. Section 9 concludes. All proofs are available in the appendix.

2 The benchmark model

In this section, we analyze the benchmark framework in which workers are assumed to be Bayesian inferers.

2.1 Assumptions

We consider the case of two workers deciding jointly whether to complete an individual or a team project. The team project is undertaken only if both workers agree to do so.

We do not take into consideration moral hazard and signalling issues assuming that workers have access to common knowledge prior information and are able to observe others' actions and performances whether they form a team or not. The latter assumption discussed in the next subsection and in appendix *B* permits to focus our attention on the impact of Bayesian and self-serving learning on workers' decision to form teams.

We assume a two periods decision problem. At time 0, the two coworkers decide simultaneously whether to undertake the individual or the group project. At the end of the first period the outcome of the project chosen at time 0 is observed by both workers. At time 1, agents decide whether to continue with the project undertaken in the first period. The outcome associated to the project performed in the second period is observed at time 2. The sequence of decisions as well as the payoffs of the individual and team projects are represented in figure 1, where q^* stands for the prior expected ability of the workers.

The team members do not know his and other's abilities necessary to undertake the projects. Workers update their beliefs about abilities at the end of the first period after observing the outcome of the project chosen in the first period. We assume the participation constraint is satisfied so that expected utility for both the individual and team projects are always higher than the payoffs of not working. The workers are always better-off exerting effort so that their decision is not about working or not working but about working alone or as a team. We assume agents are risk neutral, they select their projects by maximizing expected payoffs. The risk neutrality assumption

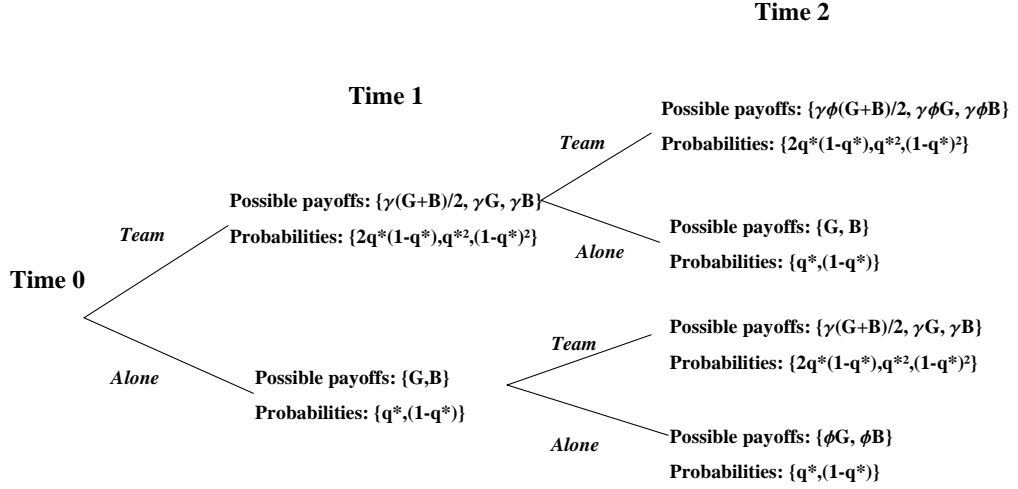


Figure 1: Payoffs tree

simplifies calculations. An agent $i \in \{1, 2\}$ when working alone undertakes a project that is a success (failure) with probability q_i ($1 - q_i$) and delivers a payoff $X_{I,i,t}^R = [R(\phi - 1) + 1]G$ ($[R(\phi - 1) + 1]B$).⁵ We assume in the rest of the paper without loss of generality that $G = 1$ and $B = 0$. We denote R the number of times the project has been previously undertaken, $R \in \{0, 1\}$. That is, we assume that when a project is repeated its expected value is multiplied by $\phi \geq 1$. As you work more time on a project you develop abilities specific to that project in addition to your initial level of ability on the task (q_i). The subscript t corresponds to time where $t \in \{0, 1, 2\}$, we drop the time subscript when not necessary. The outcomes of the two individuals' projects are assumed to be independent. If workers choose to form a team, they are involved in a project that delivers the following payoff to each worker: $X_{G,t}^R = [R(\phi - 1) + 1] \frac{\gamma(X_{I,1,t} + X_{I,2,t})}{2}$. The total outcome of the group project ($[R(\phi - 1) + 1] \gamma(X_{I,1,t} + X_{I,2,t})$) is shared equally among coworkers. The parameter γ represents synergies obtained for working in a team, the higher γ the higher the synergies, γ is known by the workers at time 0 and assumed to be positive. The absence of synergies corresponds to $\gamma = 1$, that is the team project total outcome is the sum of the individual projects outcomes. Strictly speaking synergies are defined by $\Gamma \equiv \gamma - 1$. We assume a beta prior distribution for individual abilities: $q_i \sim \text{Beta}(\alpha, \beta)$.⁶ We consider no discount factors, the effect of discounting would be to reduce the role of learning about workers' abilities at time 1.⁷

⁵We denote $X_{I,i,t} = \frac{X_{I,i,t}^R}{R(\phi-1)+1}$.

⁶The beta prior assumption is convenient since the beta distribution is a conjugate prior for the binomial problem considered here (Box and Tiao 1973). In addition, beta distributions can approximate any reasonably smooth unimodal distribution on $[0, 1]$ (Lee 1997).

⁷A low discount factor would not be consistent with our aim since we want to consider

2.2 Comments on the assumptions

Instead of assuming perfect observability of coworkers' performances, it may appear more natural to assume that workers learn more about their partner when they form a team with him. We may consider that workers are able to observe others' performances only when they form a team. This leads to a theoretically more complex game in which workers may decide to hide bad news about their abilities in order to signal themselves as high ability coworkers. We analyze such a game in appendix *B* in which both workers decide simultaneously whether to form a team or work alone. We find that propositions (1 & 2) are not modified for any of the multiple equilibria of the game. This is the case since the conditions for team formation at time 0 crucially depend on the conditions for team formation at time 1 after a team has been formed at time 0. Since the latter conditions do not change given that coworkers' performances are assumed to be mutually observable in the team, the conditions for team formation at time 0 are unchanged.

Concerning the risk neutrality assumption, we have to mention that taking into account risk aversion is likely to strengthen our results. The idea is that as self-serving biases increase, uncertainty about team continuation at time 1 increases so that the negative impact on workers' cooperation of biased learning will be higher for risk averse agents.

Rather than assuming the presence of a learning by doing effect (ϕ), we can rewrite our standard model considering a fixed cost $C > 0$ for shifting from the individual (team) project to the team (individual) project at time 1.⁸ The main results of our paper are not modified by introducing costs of shifting from one type of project to another instead of considering learning by doing effects.⁹

The equal splitting allocation rule is justified by the fact that at time 0 workers have the same prior beliefs. However, at time 1 updated workers' abilities may differ leading to bargaining opportunities. We assume in this section and in sections 3 and 4 that agents commit at time 0 to the equal splitting rule, alternative allocation rules are considered in sections 5 and 6.

We consider in our model a situation in which workers have the possibility to leave the team at time 1, there exists however cases in which agents may attempt to commit at time 0 to continue with the project started at time 0. We have to stress that commitment at time 0 may be broken at time 1 if one of the two workers has an interest to do so. In our framework commitment is not credible, this happens in many real life situations in which an ex ante agreement can be ended without further costs. We consider such examples in the case of the academic profession in section 8. Two coauthors may not be able to credibly commit to continue working together since they know that

projects for which learning and then self-serving biases matter.

⁸You may assume that there is a fixed cost to pay for undertaking a project. You do not have to pay such a cost if you repeat the same type of project in the second period.

⁹In Appendix *C*, the conditions for team formation are established for this alternative specification of our model.

one of the researchers can break the agreement at a low cost. The author who decides to leave the joint enterprise believes that cooperation was not profitable so that reputation costs that would threaten future joint work is reduced. Evidently, costs of reputation may be much higher if the coauthors are jointly involved in other projects or if the reputation of one coauthor can be communicated cheaply to other possible coauthors.¹⁰

2.3 Analysis of team formation

We solve the benchmark model by backward induction, establishing first the optimal conditions for team formation at time 1 given the history of signals and the project undertaken at time 0. The conditions for team formation at time 0 are derived in proposition 1 given the optimal decisions taken by workers at time 1. We denote $q_B = E[q_i | X_{I,i,1} = B] = \frac{\alpha}{\alpha+\beta+1}$ and $q_G = E[q_i | X_{I,i,1} = G] = \frac{\alpha+1}{\alpha+\beta+1}$.

Proposition 1 *At time 0, the condition for team formation is:*

$$\left\{ \begin{array}{l} \gamma \geq \gamma^* > 1, \text{ where } \gamma^* = \left\{ \begin{array}{l} \frac{(1+\phi)(\alpha^2+\alpha+\beta^2)+\phi(2\alpha+1)\beta}{(1+\phi)(\alpha^2+\alpha+\beta^2)+(2\alpha+1)\beta} \text{ for } \phi < \phi_* \\ \frac{2}{\phi(1+\frac{\alpha}{\alpha+1})} \text{ for } \phi_* \leq \phi < \phi^* \end{array} \right\} \\ \gamma \geq 1, \text{ for } \phi \geq \phi^* \end{array} \right\}$$

Where $\phi_* = \frac{\alpha^2+\alpha+\beta^2+\sqrt{(\alpha^2+\alpha+\beta^2)^2+4(\alpha^2+\alpha+\beta^2+2\alpha\beta+\beta)^2((2\alpha+1)^2+2\alpha+1)}}{2(\alpha^2+\alpha+\beta^2+2\alpha\beta+\beta)(2\alpha+1)}$ and $\phi^* = \frac{2}{1+\frac{\alpha}{\alpha+1}}$.

Given the optimal decisions of coworkers at time 1, agents will form a team at time 0 as long as synergies are sufficiently high.¹¹ A necessary condition for team formation at time 0 is $\gamma > 1$ so that positive synergies are required for agents to work together. The *ETO* in our model is obtained if teams are formed at time 0 and continued at time 1 whenever there exists positive synergies.¹² We call efficient teams equilibrium (*ETE*) an equilibrium that implements the efficient teams outcome. As it clearly appears in figure 2, teams characterized by strictly positive synergies ($\gamma > 1$) may not always be formed at time 0 according to proposition 1. This is a consequence of two important assumptions. First, the probability of losing the learning effect of undertaking twice the same projects is higher when choosing the team project than when working alone. The latter comes from the assumption that the

¹⁰We consider the issue of reputation in another paper Corgnet (2005c).

¹¹Using backward induction the equilibrium stated in proposition 1 is subgame-perfect since we consider a finite game with perfect information. Notice that the strategy consisting in never forming a team constitutes a Nash equilibrium of the subgame at time 1. As a result, there exists other subgame-perfect Nash equilibria than the ones captured in proposition 1. However, the latter equilibria are not considered in the rest of the paper since we want to analyze the decisions of coworkers when at least some degree of cooperation is possible (See appendix B).

¹²The efficient outcome is the expected payoffs a benevolent social planner would obtain by maximizing aggregate welfare.

agreement of the two coworkers is necessary for team formation. Assuming a team has been formed at time 0, it can split at time 1 either because of worker 1 or because of worker 2 whereas the individual projects are not continued only if both coworkers at the same time agree to do so. To compensate for this effect, coworkers at time 0 will require strictly positive synergies to engage in the formation of a team. The second important assumption underlying proposition 2 is the equal splitting rule.¹³ Such a rigid allocation rule prevents workers receiving better news about abilities from obtaining a bigger share of the group outcome. A rigid allocation rule is a rule decided at time 0 and maintained at time 1 independently of the signals received. The first proposition holds for any rigid allocation rules. This is the case since the equal splitting rule is at time 0 the most favorable allocation rule for workers' cooperation given that agents prior abilities are identical. As a result, if teams characterized by positive synergies are not formed under the equal splitting rule they will not be formed under any other rigid allocation rules. Our first proposition is closely related to the result obtained by Farrell and Scotchmer (1988). The authors analyze the formation of coalitions in a population of N agents. The members of a coalition are assumed to share the output of the group equally. They consider a static model in which no learning about abilities occur, the role of synergies (γ) is played by the economies of scale associated to the size of a coalition. Our model can be seen as a dynamic version of the framework of Farrell and Scotchmer (1988) when $N = 2$. The authors conclude that in equilibrium the sizes of coalitions are too small. This result is driven as in our model by the combination of agents differing in abilities (this occurs in our model at time 1) and equal splitting of the group outcome. Most able agents are reluctant to form coalitions with less able workers because of the equal splitting of the group outcome.

We show in section 5 that under the relative ability sharing rule, positive synergies is a sufficient condition for team formation at time 0 in the benchmark model. As shown in figure 1, the threshold for team formation is non-monotonic in the learning by doing effect ϕ , it is increasing in ϕ for $1 \leq \phi < \phi_*$ and decreasing for $\phi_* \leq \phi < \phi^*$, where ϕ_* stands for the level of learning by doing such that $\gamma_* = \frac{(1+\phi)(\alpha^2+\alpha+\beta^2)+\phi(2\alpha+1)\beta}{(1+\phi)(\alpha^2+\alpha+\beta^2)+(2\alpha+1)\beta}$ and ϕ^* is taken such that $\gamma_* = 1$. At $\phi = 1$ and $\phi = \phi^*$ the threshold for team formation takes its minimum value of 1. For low values of ϕ ($\phi < \phi_*$), an increase in ϕ leads workers to require more synergies to accept team formation at time 0. This occurs because the probability that learning by doing effects are lost at time 1 is higher for the team project than for the individual project. This is the case because the agreement of the two workers is required to form a team. For sufficiently large values of ϕ ($\phi \geq \phi^*$) the latter effect is more than compensated by the fact that an increase in ϕ has a more positive impact on the expected output of the team project than on the expected output of the

¹³As will appear clearly in section 5, the equal splitting rule is a necessary assumption to obtain insufficient teamwork.

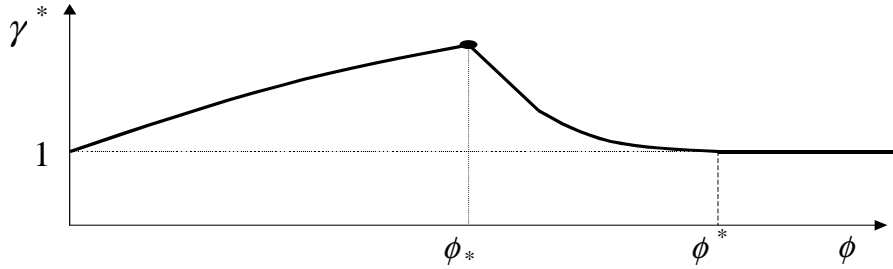


Figure 2: Threshold for team formation at time 0 (γ^*) as a function of learning by doing (ϕ).

individual project as long as synergies are positive ($\gamma \geq 1$). As a result, for $\phi \geq \phi^*$, an increase in ϕ will lead workers to be less demanding in terms of synergies to form a team at time 0.

3 The biased self-attribution model

3.1 The self-serving attribution case

3.1.1 Assumptions

In this section we consider that workers suffer from biases in their learning process whether they work alone or in a team.¹⁴ Motivated learning as mentioned in the introduction can be seen as Bayesian learning with imperfect processing of negative signals. As presented in the first section, psychologists have found extensive evidence that agents tend to recall and process their successes more easily than their failures (Korner 1950, Silverman 1964, Mischel, Ebbesen and Zeiss 1976). Researchers have found that positive personality information is efficiently processed whereas negative personality information is poorly processed (Kuiper and Derry 1982, Kuiper and MacDonald 1982, Kuiper et al. 1985). We introduce inferences biases by assuming that workers ignore bad signals about their ability with probability p . Our assumption implies a different treatment of bad and good signals, such an asymmetry in the learning process is what we refer to as motivated or self-serving learning. Workers are more easily tempted to ignore bad signals than good signals about their abilities in order to build a positive self-image. Through time, above average effects will arise in which workers will see themselves as more talented than others. The latter effects will generate a dispersion in coworkers' beliefs about one's and other's abilities. Differences in perceptions about abilities will lead agents to break teams. The learning process considered in this section is described in assumption 1. Workers are assumed to suffer from self-serving attribution mistakenly interpreting bad signals about their abilities as the result of bad luck rather than being the result of insufficient

¹⁴Assuming that self-serving biases only arise when workers are in a team would strengthen our results by increasing the negative impact of learning biases on team formation.

talent. A worker believes with probability p that bad news are uninformative about his ability and decide to ignore such information.¹⁵

Assumption 1 (self-serving learning)

A worker suffering from self-serving attribution biases will update his ability as a Bayesian inferrer suffering from imperfect information processing. The updating rule is described as follows, $\forall i \in \{1, 2\}$:

$$\left[\begin{array}{l} E_{i,S} [q_i | X_{I,i} = G] = E [q_i | X_{I,i} = G] \\ E_{i,S} [q_i | X_{I,i} = B] = E [q_i | X_{I,i} = B] \text{ with probability } (1 - p) \\ E_{i,S} [q_i | X_{I,i} = B] = E [q_i] \text{ with probability } p \end{array} \right]$$

A worker updates his coworker’s ability as a Bayesian inferrer.

We denote $E_{i,S}$ the expected workers’ abilities as computed by worker i taking into account that with probability p the agent suffers from biased information processing.¹⁶ We introduce a subscript i for the expectation of worker i since when learning biases are present coworkers’ expectations may not coincide. We assume that the two coworkers suffer from learning biases. We denote Π the probability that at least one of the two workers exhibit self-serving biases at time 1, by construction: $\Pi = 2p - p^2$ as long as the biases suffered by the two workers are taken to be independent. The inference process described above can be referred to as self-enhancing since it leads agents to believe they are more talented than they really are. According to our learning process, workers are assumed to update their beliefs about one’s and other’s abilities differently. Agents are Bayesian inferers when updating others’ abilities but assumed to suffer from self-serving attribution biases when updating their own ability. There is evidence in the psychology literature that individuals see themselves more positively than they are seen by others. For example, Lewinsohn et al. (1980) compared the ratings made by observers and by college students themselves about personality characteristics like friendliness, warmth and assertiveness of students involved in a group interaction task. They found that self-ratings were significantly more positive than observers’ ratings. In our framework, workers being closer to each other by forming a team, self-serving biases may arise as well in learning about one’s coworker’s ability. We consider for simplicity the case in which team workers do not suffer from learning biases in assessing others’ ability. The results of this section would still hold as long as self-serving biases in assessing one’s ability are more pronounced than in estimating one’s coworker’s ability. A discussion of this assumption is provided in sections 8.1.3 and 8.2.1. This is the case since our inference process leads workers to overestimate their contribution to the team outcome.

¹⁵The information received by the worker at time 1 is however always informative about his ability. We do not model explicitly the fact that some signals may be uninformative.

¹⁶Alternatively, we can consider the case of two agents with different degrees of self-serving attribution: $p_1 \neq p_2$. The results derived below continue to hold taking $p = \text{Max}\{p_1; p_2\}$ and $\Pi = p_1 + p_2(1 - p_1)$.

Another important element of our quasi-Bayesian learning process model concerns the degree with which workers are aware of their biases. A span of possible assumptions about the level of awareness about coworkers biases is considered below.

Assumption 2 (awareness of biases)

Assumption 2a: (the naive case). Workers are unaware of neither their learning biases nor their coworker’s biases.

Assumption 2b: (asymmetric awareness of biases). Workers are unaware of their learning biases but aware of their coworker’s biases.

Assumption 2c [2c’]: (partial [full] awareness of biases). Workers are aware of their coworker’s biases as well as their own biases and know that a worker’s biases are recognized at time 1 with probability $\rho < 1$ [$\rho = 1$].

A first possibility is to consider that workers are naive inferers totally unaware of their biases. In this case the effect of learning biases will not translate to the formation of teams at time 0. As many teams will be formed at time 0 in the biased learning and in the Bayesian learning cases but more teams will be split at time 1 in the case learning biases are introduced. An alternative assumption is to consider an asymmetry in awareness of biases. Workers may recognize that others tend to learn positively about themselves without seeing themselves as suffering from such a bias. As argued in Gilbert et al.(1998) people are able to deceive themselves as long as they are not fully aware of the incentives they have to do so. As a result, people willing to build a positive self-image are unlikely to be fully aware of their biases whereas they may be fully aware of others’ biases as long as they are not interested in building a positive image of others. In our model, under assumption 2c workers are able to recognize their biases with probability ρ (we assume for simplicity that the degree of awareness ρ is identical for both workers), as long as $\rho < 1$ the effects of learning biases at time 1 are not fully eliminated. Finally, we can refer to the case studied in Bénabou and Tirole (2002) in which workers are fully aware of their learning biases.¹⁷ Full awareness of biases (assumption 2c’) is of no interest in the present model since for $\rho = 1$ the effects of learning biases entirely disappear. We analyze in appendix *D* a variant of our model in which there exists a probability of not receiving any signals at time 1. In that case the effects of learning biases are not fully eliminated since workers may not be able to recover the true information about their ability. We find that self-serving biases undermine workers’ cooperation similarly to proposition 2 stated below. Certainly, the most realistic assumption lies in between cases 2a and 2c’ implying an intermediate degree of sophistication of agents. Psychologists have stressed the limited awareness agents have of their mental processes (Epstein 1983, Gilbert et al.1998). As stated in Fiske and Taylor (1991):

¹⁷Bénabou and Tirole refer to this assumption as metacognition. Their case corresponds to assumption 2c’.

"...the process by which self-structures influence the processing of self-relevant information and subsequent affect and behavior are largely preconscious without conscious directive activity on the part of the self."

The former arguments are consistent with assumptions 2a to 2c. According to assumptions 2a-2b, the conditions for team formation at time 1 are modified compared to the benchmark model if at least one coworker suffers from self-serving biases, this occurs with probability Π . Assuming 2c, the conditions for team formation at time 1 are modified if at least one coworker suffers from self-serving biases and does not realize he is biased, this occurs with probability $\Pi_\rho = 2p(1 - \rho) - p^2(1 - \rho)^2$.

3.1.2 Analysis

The model is solved as for the benchmark case by establishing first the conditions for team formation at time 1 and by determining the optimal decisions of coworkers at time 0 once taken into account their optimal behavior at time 1. Conditions for team formation at time 0 are derived in the appendix and represented in figure 2. We denote π the variable such that $\pi \equiv p$ under assumption 2b and $\pi \equiv \Pi_\rho$ under assumption 2c. In the rest of the paper, when not mentioned explicitly, our propositions are derived under both assumptions 2b and 2c. In the next proposition we compare the conditions for team formation at time 0 when coworkers are rational and when they suffer from self-serving attribution.

Proposition 2 *If rational coworkers decide not to form a team at time 0, coworkers suffering from self-serving biases will not form a team at time 0 neither.*

There exists a range of values for γ and ϕ such that coworkers with self-serving biases do not form a team at time 0 whereas rational coworkers do.

The proposition states a fundamental result of our paper, that is the cooperation between coworkers is undermined by self-serving attribution. The introduction of self-serving learning leads coworkers to interpret information differently. Agents tend to ignore part of the negative signals they receive about their abilities whereas updating other's ability as a Bayesian inferrer. The direct consequence of this learning process is that workers are likely to see themselves as contributing more to the success of the team project than their coworker. As a result, coworkers that formed a team at time 0 will more frequently split the group at time 1 than in the rational model. Anticipating such an effect, coworkers at time 0 will decide to build teams under more restrictive conditions than in the rational model. Workers will require higher levels of synergies to form teams at time 0. There exists levels of synergies that are not sufficient for team formation in the model with self-serving biases whereas sufficient for team formation in the rational model. We have to emphasize that there exists no range of values of synergies such that

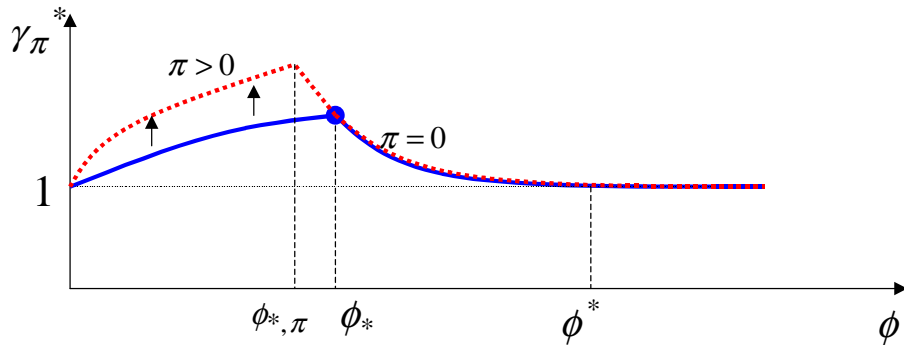


Figure 3: Threshold for team formation at time 0 (γ_π^*) as a function of learning by doing (ϕ) when self-serving biases are present. For assumption 2b: $\pi = p$ and for assumption 2c, $\pi = \Pi_\rho$.

learning biases can have a positive impact on team formation. It may appear surprising since for very large levels of synergies an increase in the estimate of one's ability has a more positive impact on the team project than on the individual project expected outcome. However, in our model when synergies are sufficiently high the formation of a team is ensured whether agents suffer or not from self-serving biases. Biased learning has an impact on team formation for intermediate values of synergies, that is for $\gamma_B \leq \gamma < \gamma_*$ when a team has been formed at time 0 and $\phi^2 \gamma_B \leq \gamma < \phi^2 \gamma_*$ after working alone at time 0. But, for intermediate values of synergies, an increase in one's estimate of ability leads to a higher increase in expected outcome for the individual project than for the team project.

Our proposition holds since the team project is preferred to the individual project when an increase in workers' estimate of their ability has a higher effect on the expected output of the group project. Our proposition is robust to any synergy functions as discussed in section 5.2 and in appendix A.

Assuming workers possess private information about their own opportunities, Eaton and Hollis (2003) obtain a result of insufficient teamwork. Their result depends however on the fact that the team project output of a worker is independent on his coworker's ability. If this is not the case, private information may lead to excessive as well as insufficient teamwork. Consider our benchmark model in which agent's ability is private information and both workers have low abilities (q_B). At time 1, teamwork will be chosen for $\gamma\phi \geq \frac{2q_B}{q_B+q^*}$ ($\frac{\gamma}{\phi} \geq \frac{2q_B}{q_B+q^*}$) after a team has (not) been formed at time 0.¹⁸ Given that $\frac{2q_B}{q_B+q^*} < 1$, teams associated with negative synergies will be formed. This shows that private information and biased self-attribution

¹⁸In the rest of the paper, when writing $\gamma\phi \geq g$ ($\frac{\gamma}{\phi} \geq g$) we will refer implicitly to the case in which a team has (not) been formed at time 0.

have different consequences on team formation. Biased self-attribution is an effect that adds to standard informational issues like moral hazard and ex ante information asymmetry. In the absence of such informational issues, biased self-attribution still operates to deter teamwork. In addition, the Biased self-attribution explanation to insufficient team formation leads to specific recommendations. We analyze in the next sections possible solutions to insufficient cooperation: the management of information and the selection of projects (section 4), team contracts for coworkers with self-serving learning (section 6), the role of team managers (section 7), the psychological and cultural selection of coworkers (section 8.1.3), building social capital and corporate culture (section 8.2.1).

In our second proposition, we establish that for some range of values of synergies, the coworkers suffering from self-serving biases make different decisions than in the case of rational coworkers. In particular, workers may not form teams that would be formed by rational coworkers. Agents are behaving non optimally for some range of synergy values by refusing to form profitable teams. The presence of learning biases leads workers to be excessively demanding in terms of synergies to accept the formation of teams, this has a negative effect on coworkers' expected payoffs. Corollary 1 captures this intuition.

Corollary 1 *The expected welfare of rational coworkers is at least as high and for a range of values of γ and ϕ strictly higher than the expected welfare of coworkers suffering from self-serving biases.*

Consider that risk neutral managers have to select workers for the completion of a project. Managers' payoffs are assumed to be a fraction of workers projects outcomes. The optimal decision for managers consists in selecting workers with the lowest degrees of self-serving biases. Taking into account that low self-confidence agents have lower degrees of self-serving biases (Kuiper and Derry 1982), managers will maximize their expected payoffs by hiring workers with the lowest self-confidence levels. However, since there is evidence of positive correlation between self-confidence and ability, the selection of team workers will lead to a trade-off consisting in choosing agents with high abilities but not excessively high self-confidence.^{19,20} We obtain results in apparent opposition to Gervais and Goldstein (2004). In their model they find that biased self-perception leads workers to increase their level of cooperation. Gervais and Goldstein do not analyze the decision of team formation, they assume a team has been formed and analyze the optimal level of effort exerted by team members when workers suffer from biased self-perception. In their model, biased self-perception facilitates cooperation among workers by increasing the perceived marginal product of effort. In our model, coworkers'

¹⁹There is evidence that experts tend to be more overconfident than relatively inexperienced people (Griffin and Tversky 1992).

²⁰In section 7, we evoke team managers' incentives to hire self-serving workers.

biased self-perception obtained through self-serving learning has a negative impact on the formation of teams since it creates conflicts in beliefs among agents. Workers do not form teams sufficiently often leading to a lower expected utility than in the benchmark model.

4 The impact of information on team formation and coworkers' expected welfare

In this section, we analyze how the formation of teams and the expected payoffs of coworkers is modified by the frequency with which information about workers' abilities is released at time 1. We consider an alternative model to the one presented in section 2 in which performances on the projects completed in the first period are observable at time 1 with probability (ω) and only observable at the end of the second period with probability $(1-\omega)$.²¹ The absence of signals about workers' abilities at time 1 implies that the decision for working in a team is the same at times 0 and 1. In the no time 1 signals model ($\omega = 0$), a team will be formed at time 0 if synergies are positive ($\gamma \geq 1$). We establish in the next proposition that the expected utility of coworkers is always at least as high and sometimes strictly higher for sets of projects with lower signals availability at time 1, that is for sets of projects characterized by a lower ω .²²

Proposition 3 *Whatever $0 \leq p \leq 1$, coworkers' expected welfare in a model with low signals availability (ω_L) is at least as high as the expected welfare of coworkers in a model with higher signals availability ($\omega_H > \omega_L$), and strictly higher for a range of values of γ and ϕ .*

Our proposition states that receiving information at time 1 about agents' abilities can never increase coworkers' expected welfare. The signals received at time 1, when leading to asymmetric beliefs about agents' abilities, may lead one of the two coworkers to stop with the team project undertaken at time 0. In figure 4, we see that as signals availability increases (ω goes up) the conditions for team formation at time 0 are more demanding in terms of synergies. As ω increases, coworkers' expected utility will be reduced since teams with positive synergies will not be formed. The release of signals about workers' abilities tends to create heterogeneity in workers' beliefs, this effect is even stronger when self-serving biases are considered since then identical signals about one's and other's abilities can be interpreted differently. The release of signals about workers' abilities has a negative effect on agents' welfare since it undermines profitable cooperation. The two workers would

²¹We assume that the observability of individuals' performances are perfectly correlated. This is a simplifying assumption that is not in contradiction with the fact that individual projects performances are independent. This assumption is relaxed in proposition 4.

²²Notice that proposition 3 holds on assumptions 2a and 2c. For assumption 2b (asymmetric awareness of biases) there may exist a range of values for which coworkers' expected welfare is higher when ω goes up. For p or ω close to 0 proposition 3 holds as well for 2b (see appendix A).

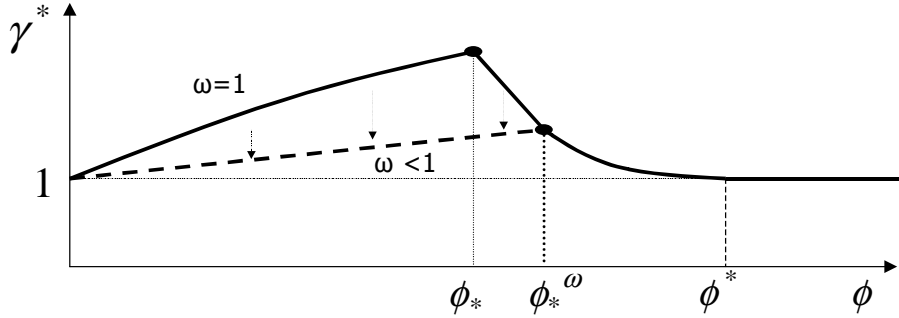


Figure 4: Threshold for team formation at time 0 in the case of total signals availability ($\omega = 1$) and in the case of partial signals availability ($\omega < 1$), for $p = 0$. We denote ϕ_*^ω the threshold such that $\gamma_\omega = \gamma_*$ where $\gamma_\omega = \frac{(1+\phi)(\alpha+\beta)(\alpha+\beta+1)-\omega(2\alpha+1)\beta}{(1+\phi)(\alpha+\beta)(\alpha+\beta+1)-\omega(2\alpha+1)\beta\phi}$.

be better-off ignoring the signals received at time 1. However, given that one of the worker ignores the signals at time 1 it is optimal for the other worker to take the signals into account. A possible solution for a manager to maximize coworkers' expected payoffs is to design sets of projects in which performances are observed with delay.

The level of workers' cooperation is undermined by the possibility of learning about workers' abilities, this leaves a role for projects designs fostering teamwork by limiting signals observability. In order to improve workers' cooperation, projects for which no learning about workers' abilities is possible should be preferred, especially when coworkers suffer from self-serving biases. We have assumed up to now that either the two signals were observable or none of them were. We consider now the case in which one of the two signals can be observed whereas the other is not. To analyze this case we assume that the observability of σ_1 is independent of the observability and σ_2 , where $\sigma_i, \forall i \in \{1, 2\}$ is the vector of payoffs associated to the time 0 project as perceived by worker i .²³ We denote $0 \leq \omega_i \leq 1$ the probability that signal $i \in \{1, 2\}$ is observed. We find that the unique scalars ω_1 and ω_2 such that the *ETO* is attained for any parameters values of γ and ϕ are $\omega_1 = \omega_2 = 0$. This result is stated in proposition 4, we assume for simplicity the absence of self-serving biases ($p = 0$).

Proposition 4 *Assuming that the observability of coworkers' signals is independent and denoting ω_i the probability that the signal about worker i 's ability is observable, the unique scalars ω_1 and ω_2 such that the efficient teams outcome is attained for any parameters γ, ϕ of the model are $\omega_1 = \omega_2 = 0$.*

Proposition 4 states that the *ETE* can only be achieved when no information about coworkers' abilities is provided to agents, this is true even when

²³ $\sigma_i \in S \equiv \{(B, B), (G, B), (B, G), (G, G)\}$.

we allow for independence in the observability of coworkers' signals.

Propositions 3 and 4 establish that as learning becomes less accessible, the level of synergies required for team formation becomes lower and coworkers' expected utility is increased.²⁴ Learning has a negative effect on workers' cooperation since it leads to changes in agents' beliefs that may induce teams to split at time 1 when workers stick to rigid allocation rules as for example equal splitting.

For some projects performances will be automatically observed by the coworkers so that managers are not able to hide signals about agents' abilities. Managers may avoid learning by giving workers independent instead of interdependent projects. In the case of interdependent projects, tasks undertaken in the first and second periods require the same type of abilities so that the performance observed at time 1 is informative about one's ability required to complete the second period project. In addition, performing twice the same project (individual or team project) will imply a learning by doing effect ϕ . For the independent set of projects, the first period project performance observed at time 1 will be uninformative about the ability required to complete the second period project. No learning by doing effect is present in this case since time 0 and time 1 projects do not require the same kind of abilities. The sets of independent projects, by preventing workers from learning about abilities, should be associated with a higher level of cooperation. However, as $\phi > 1$, the expected utility of coworkers need not be higher than in the case of interdependent projects. We find that for intermediate levels of synergies and learning by doing effects workers' expected utility in the independent projects case is strictly higher than workers' expected utility in the case of interdependent projects.²⁵

Another important element that can affect the learning process of coworkers and the level of cooperation among agents is the difficulty of the task. If we consider an extremely easy project, the learning by doing effect (ϕ) is likely to be very close to one. This implies that coworkers will be close to reach the *ETO* attained for $\phi = 1$. In the case of a very difficult project, the first period project performance is likely to be uninformative about workers' abilities so that no learning will occur at time 1, the payoff of the individual i task can be specified as follows: $a + bX_{I,i}$ for $b = 0$ so that no learning is possible. For projects with no learning issues teams will be formed at time 0 for any level of positive synergies ($\gamma \geq 1$). Such projects are characterized by higher levels of workers' cooperation and expected welfare than projects for which learning about workers' abilities is possible.

Finally, notice that the *ETO* is attained once coworkers use the outcome of the group project instead of the individual outcomes to update coworkers'

²⁴Further results on the impact of information are obtained in Corgnet (2005b). In particular, teams are formed more often as prior information becomes sufficiently precise.

²⁵If γ is very close to 1, the benefits for increased team formation are reduced and interdependent projects may be preferred, and if γ is very large the effect of learning when teams are formed will be especially high making interdependent projects more attractive.

abilities. This learning behavior can be interpreted as the result of a "team spirit" characterized by the fact that coworkers do not dissociate individual and group performances in order to assess agents' abilities. Interestingly this "team spirit learning" leads workers to ignore part of the information they receive about coworkers' abilities to focus only on the group achievements. The latter is consistent with proposition 3 in which we find that releasing signals with less frequency leads to more cooperation among workers. In order to improve coworkers' cooperation it may not be necessary to hide information to agents if their learning process is characterized by a "team spirit" that leads them to regard individual achievements as irrelevant information.

5 The inefficiency result, bargaining and contingent allocation rules

We have assumed that the equal splitting of the group outcome decided at time 0 was maintained at time 1. However, in case asymmetric performances are observed workers' posterior abilities are different. If worker 1 (2) receives $\sigma_1 = (G, B)$ ($\sigma_2 = (B, G)$) he will see the equal splitting rule as unfair and will have an incentive to bargain his share of the group outcome. We analyze in this section the case of allocation rules that depend on the signals received by coworkers at time 1, we refer to such rules as contingent allocation rules. An important assumption used in section 5.2 and 5.4 is that workers do not attempt to learn about their biases in the bargaining process.²⁶ The case of sophisticated workers willing to learn about their biases is considered in section 5.3.

5.1 The benchmark model under the relative ability allocation rule

We consider an allocation rule under which the share of the group outcome obtained by an individual is equal to his relative ability. The relative ability of worker i is defined as: $\frac{\hat{q}_{i,t}}{\hat{q}_{i,t} + \hat{q}_{j,t}} \forall i, j \in \{1, 2\}, \forall t \in \{0, 1\}$. We denote $\hat{q}_{i,t}$ the level of ability of worker i as updated by a Bayesian inferrer given information up to time t . Under this allocation rule, worker i 's expected payoffs for a team project undertaken for the first time is: $\gamma \hat{q}_{i,t}$. The next proposition shows that, in this case, workers will form teams at time 0 whenever there exists positive synergies ($\gamma \geq 1$). This result still holds if coworkers' prior abilities are different as long as both workers agree on the priors.

Proposition 5 *Under the relative ability allocation rule and in the absence of self-serving biases, teams are formed at time 0 whenever synergies are positive.*

Our proposition shows that selecting a flexible splitting rule, cooperation among workers can be increased up to its efficient level. The allocation rule

²⁶Instead of a bargaining process, one can think that workers are proposed at time 1 a new allocation rule η selected randomly in the interval $[0, 1]$. Agents then decide whether to accept or reject such a rule. In this setting workers do not have the opportunity to learn about their biases. That way, the allocation rule is flexible but not contingent.

considered constitutes a Pareto improvement compared to the equal splitting rule since it rises both workers' expected welfare. Under the relative ability allocation rule, propositions 3 and 4 do not hold and the no signals and signals models lead to the same coworkers' expected payoffs. We have to stress however that the results established in section 4 apply to any situation in which rigid allocation rules are used. There exists many real life examples in which bargaining is limited. A fixed allocation rule may be optimal when the costs of implementing a flexible allocation rule are too high. A rule that can be modified whenever agents receive new pieces of information can be very costly to implement.

5.2 Self-serving learning, unsophisticated agents and flexible allocation rules

The negative effect of self-serving biases on team formation does not disappear under the relative ability allocation rule. Actually, this negative effect may persist under any allocation rule. We present this intuition by showing that it may be impossible to establish an allocation rule under which workers perceive themselves better-off working as a team whereas positive synergies exist. We consider that both workers suffer from self-serving biases at time 1 and that the signal received consists of two failures, that is: $(X_{I,1,1}, X_{I,2,1}) = (B, B)$, $\sigma_1 = (G, B)$ and $\sigma_2 = (B, G)$. The latter occurs a priori with probability $\frac{\beta^2}{(\alpha+\beta)^2} p^2$ and $\left[\frac{\beta^2}{(\alpha+\beta)^2} (1-\rho)^2 p^2 \right]$ under assumptions $2a-2b$ and $[2c]$. Denoting $\hat{q}_{i,t}^j$ the ability of agent i as estimated by agent j at time t , we have by assumptions: $\hat{q}_{1,1}^1 = \hat{q}_{2,1}^2 = \frac{\alpha}{\alpha+\beta}$ and $\hat{q}_{2,1}^1 = \hat{q}_{1,1}^2 = \frac{\alpha}{\alpha+\beta+1}$. Assuming a team has been formed at time 0 and denoting η_i the share of the group outcome obtained by worker $i \in \{1, 2\}$ ($\eta_1 + \eta_2 = 1$), a team will be formed at time 1 if both workers are better-off working as a team, that is if the following conditions are satisfied:

$$\left\{ \begin{array}{l} \phi \eta_1 (\hat{q}_{1,1}^1 + \hat{q}_{2,1}^1) \gamma \geq \hat{q}_{1,1}^1 \\ \phi (1 - \eta_1) (\hat{q}_{1,1}^2 + \hat{q}_{2,1}^2) \gamma \geq \hat{q}_{2,1}^2 \end{array} \right\}$$

$$\Leftrightarrow \gamma \phi \geq \frac{\frac{\alpha+\beta+1}{2(\alpha+\beta)+1}}{\text{Min}\{\eta_1; 1-\eta_1\}} > 1 \text{ since: } \text{Min}_{\eta_1} \frac{\frac{\alpha+\beta+1}{2(\alpha+\beta)+1}}{\text{Min}\{\eta_1; 1-\eta_1\}} = \frac{2(\alpha+\beta)+2}{2(\alpha+\beta)+1} > 1.$$

As a result, whatever the sharing rule η_1 , the condition for team formation is more demanding in terms of synergies than in the *ETE* since $\frac{\frac{\alpha+\beta+1}{2(\alpha+\beta)+1}}{\text{Min}\{\eta_1; 1-\eta_1\}} > 1, \forall 0 \leq \eta_1 \leq 1$. Recall that we refer to the *ETO* as the situation in which workers are willing to form teams at time 0 whenever there exists positive synergies. We establish in proposition 6 that as long as at least one coworker suffers from self-serving biases the level of cooperation reached in the *ETE* cannot be attained whatever the splitting rule considered.

Proposition 6 *Whatever the allocation rule considered, as long as at least one coworker suffers from self-serving biases there exists a range of values of parameters γ and ϕ such that the *ETO* cannot be attained.*

Workers' beliefs diverge once agents suffer from self-serving biases. A

direct consequence is that even when bargaining is possible and positive synergies are present, an allocation rule permitting team formation and making both agents feel better-off may not exist. Moreover, even if such an allocation exists, it is not unique and workers may not agree on the rule since they have different beliefs about abilities. As a result, allowing workers to bargain is not a solution to obtain the *ETO* once agents suffer from self-serving biases. In the appendix (proposition 6'), we analyze the case in which one of the two coworkers' abilities is common knowledge whereas the other coworker suffers from self-serving biases. Our team inefficiency result first stated in proposition 2 is robust to flexible allocation rules. It is robust as well to any synergy function as is established in appendix A.

Our result about the impossibility to find an allocation rule that ensures a sufficiently high level of cooperation is in line with the experimental results of Babcock et al.(1995) and Babcock and Loewenstein (1997) in which self-serving biases tend to prevent defendants and plaintiffs from reaching an agreement about a settlement. The next corollary states that the negative effect of self-serving biases on aggregate expected welfare is maintained under any allocation rules.²⁷

Corollary 2 *Whatever the allocation rule considered, the aggregate expected welfare in the case of self-serving learning is at most as high as in the case of Bayesian workers using a relative ability allocation rule.*

We have shown that when self-serving biases are present, it is impossible to obtain the maximum level of workers' cooperation that consists in forming teams whenever there exists positive synergies. This is so because biased information processing of at least one coworker implies a divergence in beliefs among agents. Conflicting beliefs prevent workers from agreeing on a sharing rule that would make both agents better-off by attaining a higher level of cooperation. The intuition is that cooperation is undermined either by a rigidity in allocation rules (e.g: equal splitting) or by a "rigidity in beliefs". Rigidity in beliefs arise when self-serving biases are introduced since then agents disagree on workers' abilities being convinced that they hold the correct beliefs when it may not be the case.

As a corollary of proposition 6, we show that there exists no long term commitment contracts implementing the *ETO*. To do so we define a contract as the share of the group outcome η_i distributed to worker i at time 1, $\forall i \in \{1, 2\}$. We consider budget balanced contracts ($\eta_1 + \eta_2 = 1$) as well as contracts involving a third party ($\eta_1 + \eta_2 + \eta_3 = 1$).²⁸ We use the following definitions:

Definition 1 *A long term commitment contract is such that it can be renegotiated if all parties agree to do so.*²⁹

²⁷The aggregate welfare is the sum of the outcomes obtained by the two workers.

²⁸The group outcome is distributed in its totality to workers. This definition is similar to the one used in Bartling and von Siemens (2004).

²⁹This definition is taken from Salanié (1997).

Definition 2 *A contract is renegotiation-proof if it is impossible at time 1 to design a new contract that increases the utility of one agent without reducing the utility of the other agent.*

By definition, a long term commitment contract is renegotiation-proof. Corollary 3 can then be derived from proposition 6.

Corollary 3 *In the model with two agents and in the case of a third party, there exists no long term commitment contracts that can implement the ETO.*

According to Corollary 3, it is impossible to design a contract enforcing the ETO. This is the case because self-serving biases prevent workers from agreeing on the efficient allocation rule that consists in rewarding agents according to their relative ability updated at time 1. In the second part of the corollary, we extend the result of the initial framework to the presence of a third agent. The third party is assumed to be able to observe without biases the outcome of the two coworkers. We show that there exist no long term commitment contracts implementing the ETO. This is the case since by penalizing team breaks the third party will be better-off when teams are broken for a range of positive synergies. As a result, the third agent is exposed to contracts offers at time 1 from the part of the other two coworkers preventing the implementation of the ETO for a range of positive synergies.

The results of section 4 still hold under any sharing rules when self-serving biases are taken into account. The intuition is that signals received at time 1 may lead agents suffering from self-serving biases to hold distinct beliefs about workers' abilities. A discrepancy in beliefs among workers will prevent them from designing an allocation rule leading to the efficient outcome, even in the set of flexible allocation rules. As a result, making learning about abilities less accessible will limit beliefs dispersion among workers and improve cooperation and aggregate welfare.

Instead of considering the case in which coworkers become overconfident about their abilities as a result of self-serving biases we could assume that agents suffer from initial overconfidence. Such a model would lead to similar results concerning the impossibility to find allocation rules implementing the ETO. Two workers convinced that their ability is above average will have a different perception of what is a fair allocation for the team outcome. Consequently, they may not find an allocation rule that ensures a sufficiently high degree of cooperation.

5.3 The inefficiency result with sophisticated agents

5.3.1 Assumptions: the revelation game

In this subsection agents are assumed to be aware of their self-serving biases. We modify slightly the model presented in section 3 considering that workers suffering from biased self-attribution misinterpret with probability p bad news about their ability as being good news. As in the previous sections, workers update others' abilities as a Bayesian inferrer.

Assumption 1” (self-serving learning)

The updating rule is described as follows:

$$\left[\begin{array}{l} E_{i,S} [q_i | X_{I,i} = G] = E [q_i | X_{I,i} = G] \\ E_{i,S} [q_i | X_{I,i} = B] = E [q_i | X_{I,i} = B] \text{ with probability } (1 - p) \\ E_{i,S} [q_i | X_{I,i} = B] = E [q_i | X_{I,i} = G] \text{ with probability } p \end{array} \right]$$

We consider that workers give a positive to the possibility that they are biased. Workers will try to learn over their biases and recover the correct signals about their abilities. Our impossibility result captured in proposition 6 is based on the assumption that workers are unable to recover any information about their ability in the bargaining process. This behavior is consistent with assumptions 2a, 2b and 2c. In this section, we assess the robustness of our impossibility result by considering sophisticated agents of the type considered in Bénabou and Tirole (2002). The assumption that workers are aware of their biases is of particular relevance when we consider the issue of contracting. The process of writing the contract and its implementation will deliver some information about workers’ biases that are likely to be used by agents in order to reduce the negative effects of their biases. Assuming that workers assign a probability zero to the possibility that they are biased would prevent us from analyzing informational issues associated to the contracting activity.

We define a contract as in section 5.2, that is as the share of the group outcome η_i distributed to worker i at time 1, $\forall i \in \{1, 2\}$.³⁰ The set of contracts analyzed are budget balanced, that is the group outcome is distributed in its totality to workers ($\eta_1 + \eta_2 = 1$). We considered contracts that can be contingent on the signals received about workers’ abilities at time 1. The issue is that workers’ suffering from biased self-attribution may disagree about the signals received at time 1. To tackle this issue we consider that contracts are contingent on signals revealed rather than observed by the agents (σ_i). We modify the initial model by introducing a revelation game after workers have received signals about their abilities and before they select their second period project (figure 5).

The revelation game played at time 1 is as follows: at time 1^- each coworker chooses an action $a_i = (a_{i1}, a_{i2})$ that is a message vector formed of two components that belongs to the set S of possible signals observed at time 0’. The set S is actually the set of possible types of coworkers. This is the case since the perception of signals by the agents is private information.³¹

At time $t = 1^+ > 1^-$, workers decide either to continue with the project selected at time 0 or to undertake the other project.

³⁰The share of the group outcome given to worker 1 at time 1 cannot be contingent. This is why we do not consider it further.

³¹The set of possible messages being the set of types, we can use the Revelation Principle and conclude that our results continue to hold for any message space. The Revelation Principle can be applied to our model since it can be represented as a normal form game of a static Bayesian game.

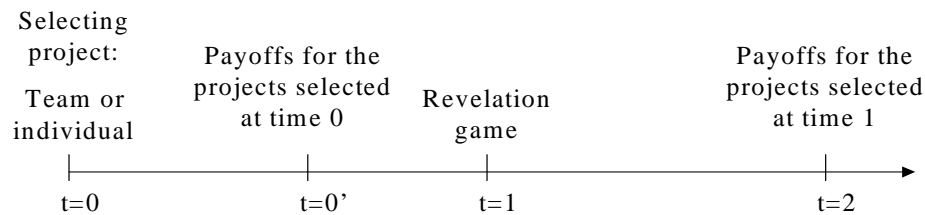


Figure 5: Time line for the model with sophisticated agents and a revelation game

The actions of the two agents will determine the share of the group outcome given to the first coworker (η) as a function of the revealed signals: $\eta \equiv \eta(a_{11}, a_{12}, a_{21}, a_{22})$. Given that workers assess others' abilities as Bayesian inferers, they possess on aggregate the correct information about their abilities. As a result, we may wonder if allowing workers to communicate will lead agents to eliminate their learning biases and lead them to cooperate efficiently.³²

5.3.2 The inefficiency result

The result captured in proposition 7 shows that such a conjecture is not verified, an *ETE* being impossible to achieve for some values of the parameters γ and ϕ . Proposition 7 is the counterpart of proposition 6 when agents are learning about their biases and have the possibility to communicate about their observed signals through a revelation game.

Proposition 7 *There exists no Bayesian Perfect Equilibria (BPE) that implement the ETO for all γ and ϕ .*

Proposition 7 shows that our impossibility result is robust to highly sophisticated workers' behaviors.

5.4 Bargaining power and biased self-attribution

We consider two coworkers, an agent suffering from self-serving learning with probability p (worker 1) and an unbiased agent (worker 2). The learning behavior of the two workers is known to both agents. The unbiased worker, endowed with the bargaining power, selects at time 0 a non contingent allocation rule $\bar{\eta}$ for the projects undertaken in the first and second periods. The rest of the assumptions are the same as in the benchmark model and in particular coworkers are assumed to have the same ability a priori. We show in the next proposition that the unbiased worker may decide to give a higher share of the group outcome to the biased worker. We consider the case in which a team has been formed at time 0. A similar reasoning holds for the case in which a team has not been formed at time 0. We denote $\bar{p} = \frac{\left(\frac{2q_G}{q_B+q_G} - \gamma\phi\right)(p_{BB}q_B + p_{GG}q_G) + (1-\gamma\phi)(p_{GB}q_B + p_{GB}q_G)}{(\gamma\phi-1)p_{BB}q_B}$ and \bar{P} the set of values

³²That is to form teams whenever there is positive synergies $\gamma \geq 1$.

satisfying $\bar{p} < 1$ for $\gamma\phi \in \left[\frac{q_B+3q_G}{2(q_B+q_G)}, \frac{2q_G}{q_B+q_G} \right] \equiv F$. For $\gamma\phi \in F$, there always exists γ and ϕ such that $\bar{P} \neq \{\emptyset\}$.

Proposition 8 *There exists a range of parameter values*

$\gamma\phi \in \left[\frac{q_B+3q_G}{2(q_B+q_G)}, \frac{2q_G}{q_B+q_G} \right]$ and $p > \bar{p}$ such that the unbiased worker, endowed with the bargaining power, prefers to select $\bar{\eta} = \frac{q_G}{\gamma\phi(q_B+q_G)} > \frac{1}{2}$ than equal splitting.

Proposition 8 puts forward that biased coworkers can benefit from their self-serving learning. In the case in which worker 2 is endowed with the bargaining power and worker 1 is known to be unbiased, the optimal choice for the second worker is to take $\bar{\eta} \leq \frac{1}{2}$. The expected payoffs of the first agent are then higher when he suffers from biased self-attribution ($p > 0$). For $\gamma\phi \in F$ the *ETO* is attained when $p = 0$ so that worker's 1 expected payoffs in the equal splitting case is $\gamma\phi q^*$. However, for $\gamma\phi \in F$, $p > \bar{p} > 0$ and assuming worker 2 has the bargaining power, the expected payoffs for worker 1 becomes $\gamma\phi\bar{\eta}q^* > \gamma\phi q^*$. Our model provides a rational motive for self-serving learning. The learning biases of a worker function as a credible threat that pushes his colleague to abandon his bargaining power to ensure sufficient teamwork.

6 Team contracts

To analyze team contracts in details we consider the hypothesis of sophisticated coworkers willing to learn over their biases. This is the adequate assumption to use since workers are faced with opportunities to learn about their biases in the contracting process.

In the next proposition we show, using the revelation game presented in section 5.3, that there exists a pooling equilibrium in which every type of worker plays the same strategy. This equilibrium leads to team formation for sufficiently high level of synergies. Under the pooling equilibrium, no information about workers' biases is revealed so that agents are unable to reduce their learning errors. The revelation game is then of no use to improve cooperation among workers. We denote p_G (p_B) the probability that $X_{I,i,1} = G$ ($X_{I,i,1} = B$), $\forall i \in \{1, 2\}$.

Proposition 9 *For $\gamma\phi \geq M \equiv \text{Max} \left\{ \frac{\hat{q}_G}{\eta(q_B+\hat{q}_G)}; \frac{\hat{q}_G}{(1-\eta)(q_B+\hat{q}_G)} \right\} \left(\frac{\gamma}{\phi} \geq M \right)$, there exists an uninformative BPE that implements the *ETO* when a team has (not) been formed at time 0, where $\hat{q}_G = wq_G + (1-w)q_B$ with $w = \frac{p_G}{p_G+pp_B}$ and $\eta(a_1, a_2) = \eta$.*

As a consequence of our last result, workers may not always be willing to play the revelation game at time 0" since they can be better-off by selecting an equal splitting rule contract ($\eta = \frac{1}{2}$). The intuition is that a truthful telling equilibrium (*TTE*) implementing the *ETO* is attainable only if the contract specifies rigid allocation rules independent of workers' types. Under truthful telling, team formation is obtained at best for $\gamma\phi \geq \frac{2q_G}{q_B+q_G} >$

$\frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} \geq \frac{2q_G}{q_B+q_G} \right)$ when $\eta = \frac{1}{2}$ whereas team formation is achieved in the absence of a revelation game for $\gamma\phi \geq \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} \geq \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right)$ when $\eta = \frac{1}{2}$. For $\gamma\phi \in \left[\frac{2\hat{q}_G}{q_B+\hat{q}_G}, \frac{2q_G}{q_B+q_G} \right] \equiv H \left(\frac{\gamma}{\phi} \in H \right)$ and $\eta = \frac{1}{2}$, information revelation will lead to an equilibrium in which teams are formed only when the signals received at time 0' are symmetric whereas in the absence of a revelation game teams would always be formed. This proposition implies that full revelation of information can lead to less team formation than no revelation of information.³³

Proposition 9 shows how the process of acquiring information about one's biases can be costly in terms of cooperation. This stresses how efficient it can be to ignore the possibility of one being biased.³⁴ This behavior would be justified in our case for sufficiently high synergies $\gamma\phi \geq \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} \geq \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right)$. However, for $\gamma\phi < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right)$ and $\eta = \frac{1}{2}$, teams are formed when signals are symmetric ($w.p : p_{BB} + p_{GG}$) in the *TTE* whereas they are formed only when signals are symmetric and no biases have occurred ($w.p : (1-p)^2 p_{BB} + p_{GG}$) in the absence of a revelation game.³⁵ The following corollary captures this result. Being unaware of one's biases is logically more detrimental as the intensity of one's biases increases, that is as p rises.

Corollary 4 *The revelation mechanism will be accepted by coworkers as long as $\gamma\phi < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right)$.*

It appears that promoting communication among workers is beneficial in terms of increased cooperation for sufficiently small levels of synergies. For high levels of synergies contracts based on rigid allocation rules and no communication mechanisms should be preferred. Corollary 4 stresses how simple contracts can be favored to more complex ones without assuming differences in costs of writing and implementation. We show that there exists a psychological cost to contracting.³⁶ In the next proposition we derive the contracts leading to the highest degree of team formation when $\gamma\phi < \frac{2\hat{q}_G}{q_B+\hat{q}_G}$, that is when contracts based on fixed allocation rules cannot ensure anymore team formation for positive synergies. We define such contracts before comparing them in proposition 10. The three contracts stated below (C_{TTE}^0), (C_{TTE}^1) and (C_{TTE}') are based on a *TTE*, we use the set S' defined as follows:

³³Further developments on *BPE* leading to partial revelation of information are available in Cornet (2005b).

³⁴A behaviour that has been observed by psychologists, see Epstein 1983 and Gilbert et al.(1998).

³⁵We use the following notations: $p_{GG} = P[X_{I,1,1} = G; X_{I,2,1} = G]$ and $p_{BB} = P[X_{I,1,1} = B; X_{I,2,1} = B]$.

³⁶This cost arises endogenously. In the literature on Contract Theory and inequity aversion, a psychological cost is directly introduced in the utility function (Englmaier and Wambach 2002, Bartling and von Siemens 2004, Rey Biel 2004).

$$S' \equiv \left\{ \begin{array}{l} (q, r, q, r) \forall (q, r) \in S, (G, B, k, l) \forall (k, l) \in S / (G, B) \\ (m, m, B, G) \forall m \in \{B, G\} \end{array} \right\}$$

Definition 3 The contract (C_{TTE}^0) is defined by the following system of equations:

$$\{ \eta_{(ijkl)} = \eta^*, \forall (i, j, k, l) \in S^2 \}$$

This contract is the *TTE* derived in proposition 8. It defines allocation rules that are independent of the coworkers' revealed signals and leads to the *ETO* for $\gamma\phi \geq \frac{2q_G}{q_B+q_G}$ $\left(\frac{\gamma}{\phi} \geq \frac{2q_G}{q_B+q_G} \right)$. The following contracts are defined in the case a team has been formed at time 0, similar contracts can be defined if a team has not been formed at time 0 by substituting ϕ by $\frac{1}{\phi}$ in systems (C_{TTE}^1) and $(C_{TTE}'^1)$.

Definition 4 The contract (C_{TTE}^1) is defined by the following conditions:

$$\left\{ \begin{array}{l} \eta_{GBGB} \in \left[\frac{1}{\gamma\phi} - \frac{q_B}{\gamma\phi(q_B+q_G)}, 1 - \frac{q_B}{\gamma\phi(q_B+q_G)} \right], \eta_{GBGG} = \eta_{GBGB} \\ (\eta_{GBBB}, \eta_{GBBG}, \eta_{BBBB}) \in A^3, A \equiv \left[0, \frac{1}{2\gamma\phi} \left[\right. \right. \\ \eta_{BGGB} \in \left[\frac{q_B}{(q_B+q_G)\gamma\phi}, 1 + \frac{q_B}{\gamma\phi(q_B+q_G)} - \frac{1}{\gamma\phi} \right], \eta_{GGGG} = \eta_{BGGB} \\ \left. \left. (\eta_{GGGG}, \eta_{BBBB}) \in B^2, B \equiv \left[\frac{1}{2\gamma\phi}, 1 - \frac{1}{2\gamma\phi} \right] \right. \right. \\ \left. \left. \forall (i, j, k, l) \notin S', \eta_{ijkl} = 0 \right. \right. \end{array} \right\}$$

The contract (C_{TTE}^1) is such that allocations depend on signals revealed by coworkers at time 1. In particular, considering $\gamma\phi < \frac{2q_G}{q_B+q_G}$, for $(X_{I,1,1}, X_{I,2,1}) = (G, B) [(B, G)]$ the share of the group outcome given to the first worker is higher [lower] than equal splitting since $\frac{1}{\gamma\phi} - \frac{q_B}{\gamma\phi(q_B+q_G)} > \frac{1}{2} \left[\frac{\gamma\phi-1}{\gamma\phi} + \frac{q_B}{\gamma\phi(q_B+q_G)} < \frac{1}{2} \right]$. The contingent contract associated to full revelation of information in equilibrium allows workers to be rewarded based on their true relative ability. However, this contingent contract does not permit teams to be formed when both workers receive a bad signal and at least one of them suffers from self-serving biases. This is the case since truthful revelation is not a possible equilibrium when $\gamma\phi < \frac{q_B+3q_G}{2(q_B+q_G)}$ $\left(\frac{\gamma}{\phi} < \frac{q_B+3q_G}{2(q_B+q_G)} \right)$ if teams are formed both for $(\sigma_1, \sigma_2) \in \Sigma$ and $(X_{I,1,1}, X_{I,2,1}) \in V$, where:

$\Sigma_1 \equiv \{ \{(G, B), (B, G)\}; \{(G, B), (B, B)\}; \{(B, B), (B, G)\} \}$ and $V \equiv \{(G, B); (B, G)\}$. To ensure that teams are not formed when $(\sigma_1, \sigma_2) \in \Sigma$ in order to make team formation possible for $(X_{I,1,1}, X_{I,2,1}) \in V$, the allocations η_{GBBG} , η_{GBBB} and η_{BBBB} are taken to be sufficiently low, that is inferior to $\frac{1}{2\gamma\phi}$.

Definition 5 The contract $(C_{TTE}'^1)$ is defined by the following system:

$$\left\{ \begin{array}{l} \eta_{GBGB} \in \left[\frac{1}{\gamma\phi} - \frac{q_B}{\gamma\phi(q_B+q_G)}, 1 - \frac{1}{2\gamma\phi} \right] \Rightarrow \gamma\phi \geq \frac{q_B+3q_G}{2(q_B+q_G)} \\ \eta_{BGBG} \in \left[\frac{1}{2\gamma\phi}, 1 - \frac{q_G}{\gamma\phi(q_B+q_G)} \right] \Rightarrow \gamma\phi \geq \frac{q_B+3q_G}{2(q_B+q_G)} \\ \eta_{GGGG} = \eta_{BBBB} = \eta_{BGBG}, \eta_{GBGG} = \eta_{G BBB} = \eta_{GBGB} \\ (\eta_{GGGG}, \eta_{BBBB}) \in B^2, B \equiv \left[\frac{1}{2\gamma\phi}, 1 - \frac{1}{2\gamma\phi} \right] \\ \eta_{G BBB} \in \left[0, \frac{1}{2\gamma\phi} \right], \forall (i, j, k, l) \notin S', \eta_{ijkl} = 0 \end{array} \right\}$$

The contract (C'_{TTE}) depends as well on signals revealed by coworkers at time 1. The contract (C^1_{TTE}) is defined only for $\gamma\phi \geq \frac{q_B+3q_G}{2(q_B+q_G)}$ whereas (C^1_{TTE}) is defined for any positive level of synergies. The reason is that for $\gamma\phi \geq \frac{q_B+3q_G}{2(q_B+q_G)}$, the contract (C^1_{TTE}) can be improved by taking $\eta_{BBBB} = \eta_{BGBG}$ and $\eta_{G BBB} = \eta_{GBGB}$ since then teams can be formed at the same time for $(\sigma_1, \sigma_2) \in \{(G, B), (B, B)\}; \{(B, B), (B, G)\}$ and $(X_{I,1,1}, X_{I,2,1}) \in \{(G, B); (B, G)\}$. The contract (C'_{TTE}) does not permit however the implementation of the efficient team outcome since teams are not formed when both workers receive a bad signal and both workers exhibit self-serving learning.

The three contracts proposed do not strictly dominate each other, choosing among them will depend on the level of synergies as well as on learning biases. This is captured in the next proposition.

Proposition 10 *The contracts leading to the highest coworkers' expected payoffs for $\gamma\phi < \frac{2q_G}{q_B+q_G}$ satisfy the following conditions:*

For $\gamma\phi < \frac{q_B+3q_G}{2(q_B+q_G)}$ ($\gamma\phi \geq \frac{q_B+3q_G}{2(q_B+q_G)}$) and $p_G < \frac{(2p-p^2)p_B}{2}$ ($p_G > \frac{p^2 p_B}{2}$), (C^0_{TTE}) is the best contract.

For $\gamma\phi < \frac{q_B+3q_G}{2(q_B+q_G)}$ and $p_G > \frac{p_B(2p-p^2)}{2}$, (C^1_{TTE}) is the best contract.

For $\gamma\phi \geq \frac{q_B+3q_G}{2(q_B+q_G)}$ and $p_G > \frac{p_B(p^2)}{2}$, (C'^1_{TTE}) is the best contract.

We know from proposition 8 that the contract (C^0_{TTE}) leads to individual work when the signals received are asymmetric whereas from the definition of the other two contracts team formation is obtained in that case. The three contracts are not equivalent, contracts (C^1_{TTE}) and (C'^1_{TTE}) being preferred when self-serving biases are not too high respectively when $p(2-p) < \frac{2p_G}{p_B}$ and $p < \sqrt{\frac{2p_G}{p_B}}$. An increase in coworkers' learning biases (p) does not affect the probability $(p_{GG} + p_{BB})$ with which a team is formed under contract (C^0_{TTE}) whereas teams are formed less frequently when p rises for contracts (C^1_{TTE}) and (C'^1_{TTE}) .

In the next corollary, we derive from propositions 9 and 10 the conditions under which contracts stating fixed allocation rules are dominated by contracts based on allocation rules contingent on coworkers' revealed signals.

Corollary 5 For $\gamma\phi < \frac{2\hat{q}_G}{q_B + \hat{q}_G}$, contingent allocation rules are strictly preferred to fixed allocation rules when: $\gamma\phi < \frac{q_B + 3q_G}{2(q_B + q_G)}$ and $p(2 - p) < \frac{2p_G}{p_B}$ and when: $\gamma\phi \geq \frac{q_B + 3q_G}{2(q_B + q_G)}$ and $p < \sqrt{\frac{2p_G}{p_B}}$.

For $\gamma\phi \geq \frac{2\hat{q}_G}{q_B + \hat{q}_G}$ contingent allocation rules are strictly dominated by fixed allocation rules.

Corollary 5 motivates the use of contingent allocation rules as long as synergies and biases are not too high. It helps as well to understand why fixed allocation rules are commonly observed. We show that it may be due to the impossibility for agents to identify their learning errors from their coworker. Our corollary implies that partnerships based on fixed allocation rules can be justified when synergies are sufficiently high or when self-serving biases are prevalent.³⁷

7 Team managers

7.1 Introduction: the case of unsophisticated workers

To introduce the issue of team managers we consider the case of unsophisticated agents. In section 5.2, we establish in corollary 3 that the presence of a third agent (e.g: a manager) in the team is not a definitive solution to achieve efficient cooperation. This result stresses the impossibility for a manager to force teamwork when renegotiation is possible. The possibility to impose teamwork has been evoked in the context of workers forming part of an organization.³⁸ However, if renegotiation is possible, agents in the organization willing to work alone, will be able to do so by offering side payments to their managers. In addition, an agent working in a firm has still the opportunity to leave the organization and work alone so that forcing teamwork should be ineffective.

7.2 A model for team managers

We consider a situation in which a third agent called a manager has the possibility to observe the outcomes of the team projects. The manager is assumed to update workers' abilities without biases. This assumption is in agreement with the motivational explanation underlying biased self-attribution. The manager is paid a proportion ($\xi > 0$) of the total payoffs of coworkers' projects. The timing of the game is described as follows. At time 0, workers decide simultaneously whether to be involved in an individual or team project rewarded according to equal splitting. They decide as well if a manager should be hired, if it is the case and the manager accepts the offer the game continues as described below. If it is not the case, the game becomes the one presented in sections 3.1 or 6.1. At time 1, workers receive the payoffs of the first period

³⁷Such conditions may be satisfied for consulting partnerships if we take into account that experts tend to be more self-confident than non experts (Griffin and Tversky 1992).

³⁸Eaton and Hollis (2003) consider that teamwork can be imposed in an organization if individual work is not rewarded at all.

project and the manager decides the allocation rule for the team project in the second period. This is equivalent to say that the manager has the full bargaining power. At time 1', workers decide whether to continue with the first period project or undertake another project. It is straightforward to see that if no bribery is possible, a *ETE* can be obtained in our framework. In the absence of bribery, the *BPE* is truthful telling: the manager pays the coworkers based on their true relative ability (that is the relative ability as perceived by the manager). We have to ensure however that a manager has an interest to participate, that is $\xi > \bar{\xi}$, and that the presence of the manager will be accepted by both workers.³⁹ We provide in the next proposition a rationale for the existence of managers. Managers prevent team conflicts by designing contracts based on informed and objective beliefs about workers' abilities. Managers are then paid an informational rent that is the result of their unbiased assessments.

As long as ξ can be taken sufficiently close to 0 the presence of a manager will be optimal for any of the situations in which the *ETE* is not achievable. That is, as stated in proposition 11, if $\bar{\xi} = 0$ a manager is hired whenever $\gamma\phi < \frac{2\hat{q}_G}{q_B + \hat{q}_G}$ ($\frac{\gamma}{\phi} < \frac{2\hat{q}_G}{q_B + \hat{q}_G}$). We state as well in the next proposition that, when bribery is possible, the rent of the manager is lower and the manager is hired less often by team members.

Proposition 11 *For $\gamma\phi < \frac{2\hat{q}_G}{q_B + \hat{q}_G}$ ($\frac{\gamma}{\phi} < \frac{2\hat{q}_G}{q_B + \hat{q}_G}$), in the absence of opportunities for bribery, a manager is hired with a strictly positive maximum rent increasing in the level of synergies (γ) and in the level of coworkers' biases (p).*

If a manager is hired when bribery is possible it will be hired as well if no bribery is possible.

For $\gamma\phi < \frac{2\hat{q}_G}{q_B + \hat{q}_G}$ ($\frac{\gamma}{\phi} < \frac{2\hat{q}_G}{q_B + \hat{q}_G}$), the best contracts are respectively: the rigid allocation rule in the absence of a revelation game and the contracts (C_{TTE}^1) and ($C_{TTE}^{1'}$) when the revelation game is available. In the absence of a communication game, the manager will be hired as long as: $\xi \leq 1 - \chi(\gamma, \phi, p)$, where $\chi(\gamma, \phi, p) = p_{BB} \frac{q_B}{q^*} \left[(1-p)^2 + \frac{1-(1-p)^2}{\gamma\phi} \right] + (p_{GG} + p_{GB}\gamma\phi) \frac{q_G}{q^*}$ and $\frac{\partial\chi(\gamma, \phi, p)}{\partial\gamma} < 0$, $\frac{\partial\chi(\gamma, \phi, p)}{\partial p} < 0$. We call $[1 - \chi(\gamma, \phi, p)]\gamma\phi$ the maximum rent of the manager. If contracts (C_{TTE}^1) and ($C_{TTE}^{1'}$) are available, the maximum rent for the manager are respectively:

$$[1 - \chi^1(\gamma, \phi, p)]\gamma\phi 2q^* \text{ and } [1 - \chi^{1'}(\gamma, \phi, p)]\gamma\phi 2q^* \text{ where:}$$

$$\chi^1(\gamma, \phi, p) = \frac{p_{BB}q_B}{q^*} \left[(1-p)^2 + \frac{1-(1-p)^2}{\gamma\phi} \right] + (p_{GG} + p_{GB}\gamma\phi) \frac{q_G}{q^*}$$

$$\chi^{1'}(\gamma, \phi, p) = \frac{p_{BB}q_B}{q^*} \left[(1-p)^2 + p_{BB} \frac{p^2}{\gamma\phi} + p_{BG} \right] + (p_{GG} + p_{GB}) \frac{q_G}{q^*}.$$

It is straightforward to see that these maximum rents are increasing in both γ and p .

³⁹We denote $\bar{\xi}$ the revenue associated to the manager's outside option.

For $\gamma\phi \geq \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} \geq \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right)$, as long as there exists a contract that implements the *ETO* (proposition 9), a manager is not hired since his maximum rent is zero.

The proposition stresses how the objectivity of managers can be rewarded in equilibrium. For the manager to be hired, the level of team synergies cannot be too high $\left[\gamma\phi < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right) \right]$, however his salary is increasing in γ for $\gamma\phi \in \left[1, \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left[\left(\frac{\gamma}{\phi} \in \left[1, \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right] \right) \right] \right]$. The rent of managers is increasing in the synergy parameter since their presence allows more teams to be formed. The more team formation is valued, the more team managers should earn in equilibrium. Another reasonable result is that managers' pay increases as coworkers' cognitive biases increase. This is the case because managers' earnings depend on their informational advantage compared to team members. If workers misperceive their ability more often, the team manager will more often have superior information compared to coworkers. As a result, the manager's informational rent and then his pay are increasing in p . In our model, there exists an incentive for team managers to maintain workers' biased self-attribution at a high level in order to maximize their informational rent. This behavior of managers is a limitation to the process of debiasing coworkers that adds to the individual's psychological cost.

In a context in which bribery is possible, the conditions for the implementation of the *ETO* are more difficult to meet. We state in the proposition that managers are hired less frequently in that case. The argument is that, in the presence of bribery, managers are paid more in equilibrium than in the case of no bribery. If it was not the case, managers could reject the bribery offer. Bribery increases the cost of hiring a manager whereas the benefits of his presence are at most the same.⁴⁰ The gains of engaging a manager are maximum when the *ETO* can be implemented in his presence, this happens when no bribery is possible but may not happen with opportunities for bribery. As a result, if γ or p are not sufficiently high, managers may not be hired in the bribery model for $\gamma\phi < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right)$. Bribery is detrimental to everybody ex ante and should be restricted as much as possible.

8 Applications

In this section, we consider first an analysis of cooperation in the academic profession. In a second part, we study the more general case of organizations. We provide an economic justification for corporate culture and investments in social capital.

⁴⁰ Further results are available upon request. We consider the modified framework in which at time 1, the two workers have the possibility to play a bribery game (B). The game (B) is such that the two coworkers simultaneously offer an amount of money $a_i \in \mathfrak{R}_+, \forall i \in \{1, 2\}$ to their manager. The manager can accept ($a_3 = Y$) this offer or not ($a_3 = N$). If the manager accepts the offer of worker i , the bargaining power is entirely transferred to worker i so that he is able to decide the allocation rule for the team project in the second period.

8.1 Coauthorship

8.1.1 The issue: insufficient cooperation among researchers

There exists evidence that research institutions and academic departments are concerned with insufficient coauthorship. To solve this issue institutions tend to "over-reward" joint works. McDowell and Smith (1992) reject the hypothesis according to which the weighting of coauthored articles in the promotion of academics is equal to the inverse of the number of coauthors. The authors use answers to a questionnaire given to 378 Economics researchers in 20 of the most important *US* institutions from 1968 to 1975. They find that each author of a n -authors article is rewarded more than $\frac{1}{n}$ of the credits given to a single authored article of the same quality. Schinski, Kugler and Wick (1998) using a survey of 140 Finance academics in 1995 confirm the result of McDowell and Smith (1992). Similar results have been obtained by Liebowitz and Palmer (1983) and in other fields than Economics by Long and McGinnis (1982). The rules used by research institutions can be interpreted as a mechanism to provide incentives for researchers to write joint papers.

8.1.2 An explanation based on biased self-attribution

As a direct application of our model, we consider the decision of two researchers to write a joint paper. As we have shown in proposition 1, when self-serving biases are present, researchers may not want to write a joint paper even when it would be optimal to do so under a rational perspective. By anticipating the possibility that you yourself or your coauthor will attribute success to his own talent rather than other's talent, you may decide to write a paper alone even in the presence of strictly positive synergies. As long as researchers suffer from learning biases and take these biases into account, less joint papers will be written than in the absence of such biases. If researchers exhibit self-serving learning but do not realize theirs and others' biases, more joint projects will be started but a larger proportion than in the rational case will end after the first article has been written.

8.1.3 A solution: find the optimal coauthor

In order to limit the negative effects of motivated learning, researchers may consider to select their coauthors according to their degree of self-serving biases. A possibility is to use gender, nationality or self-confidence as selection criteria to improve the choice of one's coauthor. There is empirical evidence of significant cultural and gender differences in the magnitude of self-serving biases. Japanese have been found to be particularly responsive to negative signals about their ability whereas North Americans tended to discount such evidence (Kitayama et al. 1997, Heine, Kitayama, Lehman, 2001). Beyer (1990) found that self-serving evaluation biases are more representative of men than women. Another possibility for researchers is to select coauthors with low self-esteem. Psychologists have found that agents with low levels of self-confidence are equally likely to recall positive and negative self-relevant information (Kuiper and Derry 1982). Finally, establishing friendship among

coauthors can reduce the negative effects of self-serving biases on cooperation. As long as workers learn as positively about others as they learn about themselves self-serving biases will not undermine team formation. Brown (1986) found that friends and relatives are evaluated less negatively than an average person since they receive more credit for success and less blame for failure. However, the negative impact of self-serving biases on group formation will not disappear as long as learning biases are stronger in evaluating one's ability than one's coworker's abilities.

8.2 Applications to the organization

8.2.1 Corporate culture and social capital

Following our model, we can regard corporate culture as a system of values facilitating cooperation in the workplace. This is the case since corporate culture by inducing workers to share common and exclusive (only the members of the firm) values functions as a mechanism to develop fellowship in the workplace. As found in Brown (1986) people tend to learn more positively about friends than about other agents so that fellowship among coworkers can reduce the negative effects of self-serving attribution on team formation. Assuming workers learn identically about themselves and about others, the *ETE* is obtained, that is teams are formed whenever there exists positive synergies. This occurs since workers discounting negative evidence about their ability will discount as well negative evidence about their coworker's ability. In that case, the conditions for team formation at time 1 are identical to the maximum cooperation case consisting in forming a team for $\gamma \geq \frac{1}{\phi}$ ($\gamma \geq \phi$) if a team has (not) been formed at time 0. According to our model, a corporate culture will be developed as long as the costs for establishing common values among workers are compensated by an increase in firm's profits due to an increase in workers' cooperation.

Formally, we modify the model described in section 2 by assuming the existence of a manager that is rewarded a proportion (δ) of the outcomes obtained by the two workers in the individual and team projects.⁴¹ In the first period of our model, we assume there exists an opportunity for the manager of our team to establish a corporate culture at a cost $c \geq 0$.⁴² We consider that the strength of cultural rules (proportional to c) reduces the probability (π in section 3) with which at least one worker learns more positively about himself than about his coworker, this probability is denoted $\pi(c)$ where $\pi'(c) < 0$. We assume $\pi(c) = \pi_0 - \epsilon c$ where $\epsilon > 0$ measures the impact of cultural rules on workers' assessment of others' abilities.

We consider the following situation: $\phi < \phi_{*,\pi(c)}$ and $1 < \gamma_{**,\pi(c)} < \gamma <$

⁴¹The existence of a manager is justified as long as his investment in corporate culture makes both workers better-off.

⁴²A corporate culture is unlikely to be initiated by managers but they have a role in promoting cultural rules. As stated by Camerer and Vepsalainen (1988): "Cultural rules spring up randomly like weeds in a garden; the manager's job is to pull the weeds and cultivate the plants".

γ_{**,π_0} such that a profitable team ($\gamma \geq 1$) will not be formed if no corporate culture is created whereas it will be formed if cultural rules are established.⁴³ Under the latter ranges of values for parameters, investing in corporate culture may rise manager's expected utility since the cost of establishing cultural rules may be outweighed by the increase in workers' cooperation. Under the latter specification, a manager will invest in corporate culture if: $c < 2\delta(1+\phi)(\gamma-1)q^* + (1-\gamma\phi)(p_{BG}(q_B+q_G) + 2p_{BB}q_B) \equiv \bar{c}$ for $\gamma < \gamma_B$ and if: $c < \bar{c} - 2(1-\gamma\phi)p_{BB}q_B$ for $\gamma_B \leq \gamma < \gamma_*$.^{44,45} As long as the cost of building social capital is not too high, the manager will be better-off investing in corporate values. More positive synergies will favour investment in corporate culture since then gains for increased cooperation are larger. For a given cost c , the range of synergy values γ for which social capital is valuable to the manager increases in coworkers' responsiveness to corporate values (ϵ).

Our model provides a motive for investing in corporate culture since it improves workers' cooperation. Corporate culture is seen as an opportunity to develop social capital in the organization. Social capital is defined by Cohen and Prusak (2001) as:

"the stock of active connections among people: the trust, mutual understanding, and shared values and behaviors that bind the members of human networks and communities and make cooperative action possible."

Cohen and Prusak provide a large number of examples of investment in social capital that improved cooperation in the organization.⁴⁶ Our framework provides a theoretical support for such investments by establishing how a high stock of social capital can lead to economic gains through an increased cooperation among workers.

Our framework can be used to justify the existence of consulting firms specialized in promoting teamwork in the organization as for example *Laxa*, *the Plexus corporation*, *the Teambuildinginc* or *the Thinkshop*.⁴⁷ Firms appear to be concerned about facilitating teamwork and are ready to invest in expensive programs that propose to develop team spirit by organizing group activities such as fly-fishing, sailing or rockclimbing. All these activities have in common that they participate in developing fellowship among coworkers so as to promote teamwork in the organization. Consistent with our model, Japanese organizations are characterized by strong personal links among employees and the extensive use of teamwork (Haitani 1990, Koike 1988).

⁴³ $\gamma_{**,\pi(c)} = \frac{(1+\phi)(\alpha^2+\alpha+\beta^2)+\phi(2\alpha+1)\beta-\pi(c)\beta^2}{(1+\phi)(\alpha^2+\alpha+\beta^2)+(2\alpha+1)\beta-\phi\pi(c)\beta^2}$ and $\frac{\partial\gamma_{**,\pi(c)}}{\partial\pi(c)} > 0$.

⁴⁴ Recall that: $\gamma_B = \frac{2q^*}{\phi(q_B+q^*)}$ and $\gamma_* = \frac{2q_G}{\phi(q_B+q_G)}$.

⁴⁵ It is easy to see that the two thresholds for (c) are strictly positive.

⁴⁶ As an example, the authors evoke the investments of the firm Alcoa that consisted in building headquarters providing more space for employees to meet and talk (p82).

⁴⁷ This is only a small subsample of consulting firms dedicated to boost teamwork in the organization.

8.2.2 Organizational structure and cultural differences

Cultural differences in the way people learn about themselves have been documented. Japanese tend to be more self-critical than individuals from Western countries (Kitayama et al. 1997, Heine et al. 1999, Heine, Kitayama, and Lehman 2001). In agreement with our model is the observation that the Japanese society in which self-criticism is seen more positively than self-confidence is characterized by a corporate culture based on teamwork and cooperation (Abegglen 1958, Haitani 1990, Koike 1988).⁴⁸ Florida and Kenney (1991) analyzing the transfer of Japanese organizations to the US report that a firm like Toyota has been working with local school systems in order to develop group-oriented behavior of his future employees. The authors document that Japanese firms select their US employees taking into account their ability to work in teams. However, an excessive degree of self-criticism may imply an inefficient use of teamwork in Japanese organizations, that is teams are formed for $\gamma < 1$.

9 Conclusion

The objective of our paper is to analyze the conditions for team formation when learning about workers' ability is introduced. In particular, we focus on learning when agents suffer from self-serving biases. We find that when the group outcome is shared equally among workers, learning with or without biases undermines workers' cooperation. This is the case because as workers receive asymmetric news about their abilities they have an incentive to split the team and work alone. A lower level of cooperation among workers leads to a reduction in agents' expected welfare. As a consequence, making learning less accessible has a positive impact on workers' expected utility.

Interestingly, when a more flexible sharing rule depending on workers' relative abilities is introduced, learning has a negative impact on team formation only in the presence of self-serving biases. The idea is that in the presence of learning biases agents may hold different beliefs about abilities preventing any flexible sharing rule to ensure a sufficiently high level of workers' cooperation. For some positive levels of synergies it may be impossible to design an allocation rule such that both workers see themselves better-off working as a team.

We are able to provide recommendations to foster workers' cooperation when agents suffer from self-serving biases as for example designing adequate team contracts, introducing a team manager or favoring projects for which learning is not accessible.

We apply our model to the academic profession. We predict that too little cooperation among researchers may occur as a result of self-serving biases. A possible solution is to select one's coauthor according to his age, gender or nationality in order to minimize one's coworker's learning biases. Our model

⁴⁸A formal analysis of team formation with self-critical workers is developed in Corgnet (2005b).

is finally used to analyze the relevance of corporate culture as a mechanism to improve workers' cooperation.

Our work by establishing the negative impact of learning biases on cooperation in the workplace challenges the view of researchers emphasizing the positive role of biased self-perception.⁴⁹ In the words of Taylor and Brown (1988):

"(...) the capacity to develop and maintain positive illusions may be thought of as valuable human resource to be nurtured and promoted, rather than an error-prone processing system to be corrected. In any case, these illusions help make each individual's world a warmer and more active and beneficent place in which to live".

Our model can be extended to more complex teams. A possible direction of research is to analyze the impact of learning biases on the formation of networks and on the optimal organizational structure of firms.

⁴⁹Many authors have focused on how biased self-perception can increase motivation and lead agents to work harder. At a theoretical level we can refer to the works of Bénabou and Tirole (2002) and Gervais and Goldstein (2004). At the empirical level, Felson (1984) found that a positive view of oneself was associated to working harder and longer on tasks.

10 Appendix

Appendix A

Proof of Proposition 1. We denote $U_T (U_{I,i})$ the total payoffs of worker i associated to the decision of forming a team at time 0. We drop subscript i since workers' utility functions are identical.

Conditions for team formation at time 1:

Using simple algebra, one can derive the following conditions for team formation at time 1.

Assuming a team has been formed at time 0, the conditions for team formation at time 1 are as follows:

For $\gamma \leq \frac{1}{\phi} \equiv \frac{1}{\phi}$, workers will decide to work alone whatever the history of signals. For $\frac{1}{\phi} \leq \gamma < \gamma_* \equiv \frac{2q_G}{\phi(q_B+q_G)} = \frac{2}{\phi(1+\frac{\alpha}{\alpha+1})}$, workers will form a team at time 1 if the signals received are either both failures or both successes. For $\gamma \geq \gamma_*$, agents will work as a team at time 1 whatever the history of signals. Assuming workers did not form a team at time 0, the conditions for team formation at time 1 are the same multiplying the thresholds $\frac{1}{\phi}$ and γ_* by ϕ^2 .

Conditions for team formation at time 0:

Assume first: $\phi \leq \gamma_*$:

For $\gamma < \frac{1}{\phi}$ whatever the history of signals at time 1 no teams are formed at time 0. For $\frac{1}{\phi} \leq \gamma < \phi$ teams are formed at 0 if the expected utility of forming a group at 0 is higher for worker $i \in \{1, 2\}$ than working alone, that is for $i = 1$ (the case $i = 2$ is symmetric):

$$\begin{aligned} E \left[U_T \mid I_0, \frac{1}{\phi} \leq \gamma < \phi < \gamma_* \right] &\geq E \left[U_I \mid I_0, \frac{1}{\phi} \leq \gamma < \phi < \gamma_* \right] \Leftrightarrow \\ P \left[X_{I,1,1} = G; X_{I,2,1} = G \right] &\left(\gamma G + \frac{\gamma \phi}{2} E \left[X_{I,1,2} + X_{I,2,2} \mid X_{I,1,1} = G; X_{I,2,1} = G \right] \right) \\ + P \left[X_{I,1,1} = B; X_{I,2,1} = B \right] &\left(\gamma B + \frac{\gamma \phi}{2} E \left[X_{I,1,2} + X_{I,2,2} \mid X_{I,1,1} = B; X_{I,2,1} = B \right] \right) \\ + P \left[X_{I,1,1} = G; X_{I,2,1} = B \right] &\left(\frac{\gamma}{2} G + E \left[X_{I,1,2} \mid X_{I,1,1} = G \right] \right) \\ + P \left[X_{I,1,1} = B; X_{I,2,1} = G \right] &\left(\frac{\gamma}{2} B + E \left[X_{I,1,2} \mid X_{I,1,1} = B \right] \right) \\ &\geq (1 + \phi) \frac{\alpha}{\alpha + \beta} \\ \Leftrightarrow \gamma &\geq \frac{(1 + \phi)(\alpha^2 + \alpha + 2\alpha\beta + \beta^2 + \beta) + \phi(2\alpha + 1)\beta}{(1 + \phi)(\alpha^2 + \alpha + 2\alpha\beta + \beta^2 + \beta) + (2\alpha + 1)\beta} \equiv \gamma_{**}, \text{ where } 1 < \gamma_{**} \leq \phi. \end{aligned}$$

Assume now: $\phi > \gamma_*$:

For $\gamma < \frac{1}{\phi}$ whatever the history of signals at time 1 no teams are formed at time 0. For $\frac{1}{\phi} \leq \gamma < \gamma_*$, the condition for team formation is: $\gamma \geq \gamma_{**}$. For $\gamma_* \leq \gamma < \phi$, the condition for team formation is: $\gamma \geq 1$. For $\gamma \geq \phi$, teams are always formed.

We can prove that thresholds for team formation at time 0 are always lower than ϕ . We define $W \equiv (\alpha + \beta)(\alpha + \beta + 1)$, it is easy to see that the expected utility for forming a team at time 0 is higher than the expected utility of staying alone as long as: $\gamma W + \gamma \phi z_w + y_w \geq W + \phi Z_w + \gamma Y_w \Leftrightarrow \gamma \geq \frac{W + \phi Z_w - y_w}{W + \phi z_w - Y_w}$ where: $z_w + y_w = Z_w + Y_w = \alpha + \beta = W$ and $\frac{W + \phi Z_w - y_w}{W + \phi z_w - Y_w} \leq \phi$.

As a result, the condition for team formation when $\phi \geq \gamma_*$ is:

$\gamma \geq \text{Max}\{1; \text{Min}\{\gamma_*, \gamma_{**}\}\}$. The first proposition follows directly from the previous thresholds. ■

Proof of Proposition 2. Conditions for team formation at time 1:

Using simple algebra, one can derive the following conditions:

Under assumption $2a - 2b$ ($2c$), assuming a team has been formed at time 0, if at least one worker ignores his bad signal (and does not realize it, it occurs with probability $(1 - \rho)$), the condition for team formation at time 1 are modified compared to the rational model as follows:

For $\gamma < \frac{1}{\phi}$, workers will decide to work alone whatever the history of signals. For $\frac{1}{\phi} \leq \gamma < \gamma_B \equiv \frac{2q^*}{\phi(q_B+q^*)}$, workers will form a team at time 1 only if the signals received are both successes. For $\gamma_B \leq \gamma < \gamma_* \equiv \frac{2q_G}{\phi(q_B+q_G)}$, workers will form a team at time 1 if the signals received are either both failures or both successes. For $\gamma \geq \gamma_*$, agents will form a team at time 1 whatever the history of signals. Assuming workers did not form a team at time 0, the conditions for team formation at time 1 are the same as when workers did form a team at time 0 except that thresholds are multiplied by ϕ^2 .

Conditions for team formation at time 0:

We use the following notations: $p_{GG} \equiv P[X_{I,1,1} = G; X_{I,2,1} = G]$, $p_{GB} \equiv P[X_{I,1,1} = G; X_{I,2,1} = B]$, $p_{BG} \equiv P[X_{I,1,1} = B; X_{I,2,1} = G]$

and $p_{BB} \equiv P[X_{I,1,1} = B; X_{I,2,1} = B]$.

We consider first the case: $\frac{1}{\phi} \leq \phi \leq \gamma_B \leq \phi^2 \gamma_B < \gamma_*$, it implies:

$\phi < \sqrt{1 + \frac{\beta}{2\alpha^2 + 3\alpha + 2\alpha\beta + \beta + 1}}$. For $\gamma < \frac{1}{\phi}$, agents will decide to work alone at time 0. For $\frac{1}{\phi} \leq \gamma < \phi$, the expected payoff of forming a team is higher than working alone if:

$$E_{1,S} \left[U_T | I_0, \frac{1}{\phi} \leq \gamma < \phi < \gamma_* \right] \geq E_{1,S} \left[U_{I,1} | I_0, \frac{1}{\phi} \leq \gamma < \phi < \gamma_* \right]$$

$$\Leftrightarrow (1 - \pi) E \left[U_G | I_0, \frac{1}{\phi} \leq \gamma < \phi \leq \gamma_* \right]$$

$$+ \pi \left[\begin{aligned} & P_{GG} \left(\gamma G + \frac{\gamma \phi}{2} E_{1,S} \left[X_{I,1,2} + X_{I,2,2} | X_{I,1,1} = G; X_{I,2,1} = G \right] \right) \\ & + P_{BB} \left(\gamma B + E_{1,S} \left[X_{I,1,2} | X_{I,1,1} = B \right] \right) \\ & + P_{GB} \left(\frac{\gamma}{2} G + E_{1,S} \left[X_{I,1,2} | X_{I,1,1} = G \right] \right) \\ & + P_{BG} \left(\frac{\gamma}{2} G + E_{1,S} \left[X_{I,1,2} | X_{I,1,1} = B \right] \right) \end{aligned} \right]$$

$$\geq (1 + \phi) \frac{\alpha}{\alpha + \beta}$$

$$\Leftrightarrow \gamma \geq \frac{(1 + \phi)(\alpha^2 + \alpha + \beta^2) + \phi(2\alpha + 1)\beta - \pi\beta^2}{(1 + \phi)(\alpha^2 + \alpha + \beta^2) + (2\alpha + 1)\beta - \phi\pi\beta^2} \equiv \gamma_{**, \pi}, \text{ where } \frac{\partial \gamma_{**, \pi}}{\partial \pi} > 0 \text{ and } 1 <$$

$\gamma_{**, \pi} \leq \phi$.

The condition for team formation is then: $\gamma \geq \gamma_{**, \pi}$. For $\gamma \geq \phi$, teams are always formed.

The second case is: $\frac{1}{\phi} \leq \phi < \gamma_B \leq \gamma_* \leq \phi^2 \gamma_B$, it implies that:

$$\sqrt{1 + \frac{\beta}{2\alpha^2 + 3\alpha + 2\alpha\beta + \beta + 1}} \leq \phi < \sqrt{1 + \frac{1}{2\alpha + 2\beta + 1}}$$

For $\gamma < \frac{1}{\phi}$, agents will decide to work alone at time 0. For $\frac{1}{\phi} \leq \gamma < \phi$, the expected payoff of forming a team is higher than working alone if: $\gamma \geq \gamma_{**, \pi}$. As a

result, the condition for team formation at time 0 is the same as in case 1.

In case 3, we consider: $\frac{1}{\phi} \leq \gamma_B \leq \gamma_* \leq \phi \leq \phi^2 \gamma_B$. The latter condition implies: $\phi \geq \sqrt{1 + \frac{1}{2\alpha+1}}$. For $\gamma < \frac{1}{\phi}$, agents will decide not to form a team. For $\frac{1}{\phi} \leq \gamma < \gamma_B$, the threshold is computed as in cases 1 and 2 so that team formation occurs at time 0 for $\gamma \geq \gamma_{**, \pi}$. For $\gamma_B \leq \gamma < \gamma_*$, the condition for team formation is: $\gamma \geq \gamma_{**}$. For $\gamma_* \leq \gamma < \phi$, the condition for team formation is: $\gamma \geq 1$. For $\gamma \geq \phi$, teams are always formed. As a result, the condition for team formation is such that:

$$\left\{ \begin{array}{l} \gamma \geq \gamma_{**, \pi} \text{ for } \gamma_{**, \pi} < \gamma_B \\ \gamma \geq \text{Max} \{ \gamma_B; 1; \text{Min} \{ \gamma_*; \gamma_{**} \} \} \text{ for } \gamma_{**, \pi} \geq \gamma_B \end{array} \right\}$$

In case 4, we have: $\frac{1}{\phi} \leq \gamma_B \leq \phi < \gamma_* \leq \phi^2 \gamma_B$. Case 4 corresponds to:

$\sqrt{1 + \frac{1}{2\alpha+2\beta+1}} \leq \phi < \sqrt{1 + \frac{1}{2\alpha+1}}$. For $\gamma < \frac{1}{\phi}$ individuals will decide to work alone. For $\frac{1}{\phi} \leq \gamma < \gamma_B$, the condition for team formation is $\gamma \geq \gamma_{**, \pi}$. For $\gamma_B \leq \gamma < \phi$, the condition for team formation is $\gamma \geq \gamma_{**}$. As a result, the condition for the formation of a team at time 0 is:

$$\left\{ \begin{array}{l} \gamma \geq \gamma_{**, \pi} \text{ for } \gamma_{**, \pi} < \gamma_B \\ \gamma \geq \text{Max} \{ \gamma_B; \gamma_{**} \} \text{ for } \gamma_{**, \pi} \geq \gamma_B \end{array} \right\}$$

For $\gamma \geq \phi$, teams are always formed. The first and second parts of the proposition are a direct consequence of the previous derivation of the thresholds for team formation at time 0. The crucial element is: $\frac{\partial \gamma_{**, \pi}}{\partial \pi} > 0$.

The proof holds for assumption 2b when π is taken to be p and for assumption 2c when $\pi \equiv \Pi_\rho$. ■

Proof of Corollary 1. It follows from proposition 2; $\forall p > 0$ less efficient teams (characterized by $\gamma \geq 1$) will be formed than in the case $p = 0$. In addition, the teams that are formed will be split with higher probability in the case $p > 0$, this implies that the expected utility of the teams formed when $p > 0$ will be lower than in the case $p = 0$. As a consequence, the expected welfare of coworkers is always higher in the absence of self-serving attribution than in the case $p > 0$. ■

Proof of Proposition 3. Conditions for team formation at time 1 after a team has been formed at time 0 (resp. has not been formed):

Case \ Signals	(G, G)	(B, B)	(B, G); (G, B)
Observability (ω)	$\gamma\phi \geq 1$	$\gamma\phi \geq \frac{2q^*}{q^*+q_B} \equiv \phi\Gamma_*$	$\gamma\phi \geq \frac{2q_G}{q_G+q_B} \equiv \phi\gamma_*$
At least 1 worker is biased	$(\frac{\gamma}{\phi} \geq 1)$	$(\frac{\gamma}{\phi} \geq \frac{2q^*}{q^*+q_B})$	$(\frac{\gamma}{\phi} \geq \frac{2q_G}{q_G+q_B})$
Non Observability ($1 - \omega$)	$\gamma\phi \geq 1$	$\gamma\phi \geq 1$	$\gamma\phi \geq 1$
At least 1 worker is biased	$(\frac{\gamma}{\phi} \geq 1)$	$(\frac{\gamma}{\phi} \geq 1)$	$(\frac{\gamma}{\phi} \geq 1)$
Observability (ω)	$\gamma\phi \geq 1$	$\gamma\phi \geq 1$	$\gamma\phi \geq \frac{2q_G}{q_G+q_B}$
No biases	$(\frac{\gamma}{\phi} \geq 1)$	$(\frac{\gamma}{\phi} \geq 1)$	$(\frac{\gamma}{\phi} \geq \frac{2q_G}{q_G+q_B})$
Non Observability ($1 - \omega$)	$\gamma\phi \geq 1$	$\gamma\phi \geq 1$	$\gamma\phi \geq 1$
No biases	$(\frac{\gamma}{\phi} \geq 1)$	$(\frac{\gamma}{\phi} \geq 1)$	$(\frac{\gamma}{\phi} \geq 1)$

At time 0, case 1: $\frac{1}{\phi} \leq \Gamma_* \leq \gamma_* \leq \phi \leq \phi^2 \Gamma_* \leq \phi^2 \gamma_*$.

For $\gamma < \frac{1}{\phi}$: no teams are formed. For $\frac{1}{\phi} \leq \gamma < \Gamma_*$, teams are formed for $\gamma \geq \gamma_\omega \equiv \frac{(1+\phi)(\alpha+\beta)(\alpha+\beta+1) - \omega(2\alpha+1)\beta}{(1+\phi)(\alpha+\beta)(\alpha+\beta+1) - \omega(2\alpha+1)\beta\phi}$ where $1 < \gamma_\omega \leq \phi$ and $\frac{\partial \gamma_\omega}{\partial \omega} > 0$. For

$\Gamma_* \leq \gamma < \gamma_*$, teams are formed for:

$$\gamma \geq \gamma_{\omega,\pi} \equiv \frac{(1+\phi)(\alpha+\beta)(\alpha+\beta+1) - \omega\beta[(2\alpha+1)\beta + \pi\beta]}{(1+\phi(1-\omega))(\alpha+\beta)(\alpha+\beta+1) + \omega\phi[\alpha(\alpha+1) + (1-\pi)\beta^2]} < \phi \text{ where } \frac{\partial \gamma_{\omega,\pi}}{\partial \omega} > 0.$$

For $\gamma_* \leq \gamma < \phi$, teams are formed for $\gamma \geq 1$. For $\gamma \geq \phi$, teams are always formed.

Case 2: $\frac{1}{\phi} \leq \Gamma_* \leq \phi < \gamma_* \leq \phi^2 \Gamma_* \leq \phi^2 \gamma_*$ or

$$\frac{1}{\phi} \leq \Gamma_* \leq \phi \leq \phi^2 \Gamma_* < \gamma_* \leq \phi^2 \gamma_*$$

For $\gamma < \frac{1}{\phi}$: no teams are formed. For $\frac{1}{\phi} \leq \gamma < \Gamma_*$, teams are formed for $\gamma \geq \gamma_{\omega,\pi}$. For $\Gamma_B < \gamma < \phi$, teams are formed for $\gamma > \gamma_{\omega}$, where $\gamma < \phi$. For $\gamma \geq \phi$, teams are always formed.

Case 3: $\frac{1}{\phi} \leq \phi < \Gamma_* \leq \gamma_* \leq \phi^2 \Gamma_* \leq \phi^2 \gamma_*$ or

$$\frac{1}{\phi} \leq \phi < \Gamma_* \leq \phi^2 \Gamma_* < \gamma_* \leq \phi^2 \gamma_*$$

For $\gamma < \frac{1}{\phi}$: no teams are formed. For $\frac{1}{\phi} \leq \gamma < \phi$, teams are formed for $\gamma \geq \gamma_{\omega,\pi}$. For $\gamma \geq \phi$, teams are always formed. An increase in signals' availability (ω rises) changes workers' decisions in the following cases where $\omega_H > \omega_L$:

a) A team is formed at time 0 in the low signals availability case but not formed in the high signals availability case, this can occur in case 1, 2 and 3 since $\gamma_{\omega,\pi}$ and γ_{ω} are increasing in ω . The coworkers' expected payoffs is higher in the low signals availability since a profitable team ($\gamma \geq 1$) is formed in that case whereas it is not formed in the case of high signals availability.

b) Workers form a team at time 0 in both the high and low signals availability cases, but a team is formed at time 1 if no signals are available whereas a team may not be formed if signals are available in the case of asymmetric signals or in the case of two bad signals when learning biases do arise. In that case, coworkers' expected welfare is higher when signals availability is low.

It follows from a) and b) that expected coworkers' welfare is higher for a lower signals availability ω under assumption 2c. Under assumption 2a, proposition 3 follows directly from b). The situation is more complex under assumption 2b since workers are underestimating the probability of a team break at time 1. As a result, an increase in ω will lead to more restrictive conditions for team formation at time 0 ($\gamma_{\omega,p} < \gamma_{\omega,\Pi}$) so that as long as ω does not rise too much (the thresholds for team formation specified above are increasing in ω). For $p = 0$ or $\omega = 0$, $\gamma_{\omega,p} = \gamma_{\omega,\Pi}$ so that a rise in γ can only have a positive impact on coworkers' expected welfare. Coworkers' expected welfare rises for any ω_H such that $\gamma_{\omega_L,\Pi} > \gamma_{\omega_H,p}$. For $\gamma_{\omega_H,p} \geq \gamma_{\omega_L,\Pi}$, the coworkers expected payoffs are higher in the case of low signals availability ω_L . ■

Proof of Proposition 4. Conditions for team formation at time 1 after a team has been formed at time 0 (resp. after a team has not been formed at time 0):

Signals \ Case	No signal 1 Signal 2	No signal 1 Signal 2	No signal 1 No Signal 2	Signal 1 Signal 2
(G, G)	$\gamma\phi \geq \Delta_G$ $(\frac{\gamma}{\phi} \geq \Delta_G)$	$\gamma\phi \geq \Delta_G$ $(\frac{\gamma}{\phi} \geq \Delta_G)$	$\gamma\phi \geq 1$ $(\frac{\gamma}{\phi} \geq 1)$	$\gamma\phi \geq 1$ $(\frac{\gamma}{\phi} \geq 1)$
(B, B)	$\gamma\phi \geq \Gamma_*$ $(\frac{\gamma}{\phi} \geq \Gamma_*)$	$\gamma\phi \geq \Gamma_*$ $(\frac{\gamma}{\phi} \geq \Gamma_*)$	$\gamma\phi \geq 1$ $(\frac{\gamma}{\phi} \geq 1)$	$\gamma\phi \geq 1$ $(\frac{\gamma}{\phi} \geq 1)$
(B, G)	$\gamma\phi \geq \Delta_G$ $(\frac{\gamma}{\phi} \geq \Delta_G)$	$\gamma\phi \geq \Gamma_*$ $(\frac{\gamma}{\phi} \geq \Gamma_*)$	$\gamma\phi \geq 1$ $(\frac{\gamma}{\phi} \geq 1)$	$\gamma \geq \gamma_*$ $(\gamma \geq \phi^2 \gamma_*)$
(G, B)	$\gamma\phi \geq \Gamma_*$ $(\frac{\gamma}{\phi} \geq \Gamma_*)$	$\gamma\phi \geq \Delta_G$ $(\frac{\gamma}{\phi} \geq \Delta_G)$	$\gamma\phi \geq 1$ $(\frac{\gamma}{\phi} \geq 1)$	$\gamma \geq \gamma_*$ $(\gamma \geq \phi^2 \gamma_*)$

Where $\Delta_G \equiv \frac{2q_G}{q^*+q_G}$, $\Gamma_* \equiv \frac{2q^*}{q^*+q_B}$ and $\gamma_* \equiv \frac{2q_G}{\phi(q_B+q_G)}$.

Assuming first: $\Delta_G \geq \Gamma_*$.

At time 0, Case 1: $\frac{1}{\phi} \leq \phi \leq \frac{\Gamma_*}{\phi} \leq \frac{\Delta_G}{\phi} \leq \gamma_* \leq \phi\Gamma_* \leq \phi\Delta_G \leq \phi^2\gamma_*$

For $\gamma < \frac{1}{\phi}$, no teams are formed. For $\frac{1}{\phi} \leq \gamma < \phi$, teams are formed for:

$$\gamma \geq \frac{(1+\phi - ((1-\omega_1)+(1-\omega_2) - 2(1-\omega_1)(1-\omega_2)))(\alpha+\beta)(\alpha+\beta+1) - (2\alpha+1)\beta\omega_1\omega_2}{(1+\phi - \phi((1-\omega_1)+(1-\omega_2) - 2(1-\omega_1)(1-\omega_2)))(\alpha+\beta)(\alpha+\beta+1) - \phi(2\alpha+1)\beta\omega_1\omega_2} \equiv \gamma_{*,\omega_1,\omega_2}$$

Where $1 < \gamma_{*,\omega_1,\omega_2} \leq \phi$ and $\gamma_{*,\omega_1,\omega_2} = 1$ if and only if $\omega_1 = \omega_2 = 0$. We

have: $\frac{\partial \gamma_{*,\omega_1,\omega_2}}{\partial \omega_i} > 0 \forall i \in \{1, 2\}$. For $\gamma \geq \phi$, teams are always formed.

Case 2: $\frac{1}{\phi} \leq \frac{\Gamma_*}{\phi} < \phi \leq \frac{\Delta_G}{\phi} \leq \gamma_* \leq \phi\Gamma_* \leq \phi\Delta_G \leq \phi^2\gamma_*$ For $\gamma < \frac{1}{\phi}$, no teams are formed. For $\frac{1}{\phi} \leq \gamma < \frac{\Gamma_*}{\phi}$, teams are formed for: $\gamma \leq \gamma_{*,\omega_1,\omega_2}$. For $\frac{\Gamma_*}{\phi} \leq \gamma < \phi$, teams are formed for:

$$\gamma \geq \text{Max} \{ \gamma_{**}(\alpha, \beta, \omega_1, \omega_2); \gamma_{**}(\alpha, \beta, \omega_2, \omega_1) \} \equiv M_{\omega_1, \omega_2}$$

Where $\gamma_{**}(\alpha, \beta, \omega_1, \omega_2) \equiv$

$$\frac{(1+\phi)(\alpha+\beta)(\alpha+\beta+1) - (\omega_1(1-\omega_2) + \omega_2(1-\omega_1))(\alpha\beta + \alpha + \alpha^2) - \omega_1\omega_2(2\alpha+1)\beta - \omega_1(1-\omega_2)\beta}{(1+(1-\omega_1)(1-\omega_2)\phi)(\alpha+\beta)(\alpha+\beta+1) + \phi(\omega_1(1-\omega_2) + \omega_2(1-\omega_1)) \left(\frac{(2\alpha+1)\beta + \beta^2}{2} \right) + \phi\omega_1\omega_2(\alpha + \alpha^2 + \beta^2)}$$

We can show that $1 < M_{\omega_1, \omega_2} < \phi$ and $M_{\omega, \omega} = 1$. For $\gamma \geq \phi$, teams are always formed.

Case 3: $\frac{1}{\phi} \leq \frac{\Gamma_*}{\phi} \leq \frac{\Delta_G}{\phi} < \phi \leq \gamma_* \leq \phi\Gamma_* \leq \phi\Delta_G \leq \phi^2\gamma_*$

For $\gamma < \frac{1}{\phi}$, no teams are formed. For $\frac{1}{\phi} \leq \gamma < \frac{\Gamma_*}{\phi}$, teams are formed for:

$\gamma \geq \gamma_{*,\omega_1,\omega_2}$. For $\frac{\Gamma_*}{\phi} \leq \gamma < \frac{\Delta_G}{\phi}$, teams are formed for: $\gamma \geq \gamma_{**,\omega_1,\omega_2}$. For $\frac{\Delta_G}{\phi} \leq$

$\gamma < \phi$, teams are formed for: $\gamma \geq \frac{(1+\phi)(\alpha+\beta)(\alpha+\beta+1) - \omega_1\omega_2(2\alpha+1)\beta}{(1+\phi)(\alpha+\beta)(\alpha+\beta+1) - \phi\omega_1\omega_2(2\alpha+1)\beta} \equiv \gamma_{***,\omega_1,\omega_2}$ where $1 < \gamma_{***,\omega_1,\omega_2} \leq \phi$ and $\gamma_{***,\omega_1,\omega_2} = 1$ if and only if $\omega_1 = \omega_2 = 0$. We have:

$\frac{\partial \gamma_{***,\omega_1,\omega_2}}{\partial \omega_i} > 0 \forall i \in \{1, 2\}$. For $\gamma \geq \phi$, teams are always formed.

Case 4: $\frac{1}{\phi} \leq \frac{\Gamma_*}{\phi} \leq \frac{\Delta_G}{\phi} \leq \gamma_* < \phi \leq \phi\Gamma_* \leq \phi\Delta_G \leq \phi^2\gamma_*$

For $\gamma < \frac{1}{\phi}$, no teams are formed. For $\frac{1}{\phi} \leq \gamma < \frac{\Gamma_*}{\phi}$, teams are formed for:

$\gamma > \gamma_{*,\omega_1,\omega_2}$. For $\frac{\Gamma_*}{\phi} \leq \gamma < \frac{\Delta_G}{\phi}$, teams are formed for: $\gamma > M_{\omega_1, \omega_2}$. For

$\frac{\Delta_G}{\phi} \leq \gamma < \gamma_*$, teams are formed for: $\gamma > \gamma_{***,\omega_1,\omega_2}$. For $\gamma_* \leq \gamma < \phi$, teams are formed for: $\gamma > 1$. For $\gamma > \phi$, teams are always formed.

Proposition 4 follows directly from: $(\gamma_{*,\omega_1,\omega_2}, M_{\omega_1, \omega_2}, \gamma_{***,\omega_1,\omega_2}) = (1, 1, 1) \Leftrightarrow \omega_1 = \omega_2 = 0$.

For $\Delta_G < \Gamma_*$ the demonstration is identical. ■

Proof of Proposition 5. Directly follows from the conditions for team formation

at time 1. Teams are formed at time 1 after a team has been formed (has not been formed) at time 0 if $\gamma \geq \phi$ ($\frac{1}{\phi}$). As a result, at time 0 teams are formed for $\gamma \geq 1$ (positive synergies). ■

Proof of Proposition 6. We show here that it is impossible to design allocation rules at time 1 that will lead workers to form teams whenever $\gamma \geq \frac{1}{\phi} [\phi]$ after a team has [not] been formed at time 0. We show that it is true even if we assume biases are observable so that allocations can be made contingent on learning biases occurring at time 1.

At time 1, if only one coworker suffers from self-serving biases and the history of signals is (B, B) , the allocation rule most favorable for workers' cooperation is such that team formation arises for:

$\gamma\phi \geq \frac{2X_B+1}{2} > 1$ since $X_B \equiv \frac{q^*}{q^*+q_B} > \frac{1}{2}$ (if a team has been formed at time 0) and (if no teams have been formed at time 0) $\frac{\gamma}{\phi} \geq \frac{2X_B+1}{2} \equiv \Theta_b > 1$.

If worker 1 (resp. 2) suffers from learning biases and the history of signals is (B, G) (resp. (G, B)), the conditions most favorable for workers' cooperation are such that teams are formed for:

$\gamma\phi \geq \frac{q^*}{q^*+q_G} + \frac{q_G}{q_B+q_G} (\equiv \Theta_{gb}) > 1$ (if a team has been formed at time 0) and $\frac{\gamma}{\phi} \geq \frac{q^*}{q^*+q_G} + \frac{q_G}{q_B+q_G} > 1$ (if a team has not been formed at time 0).

If both workers suffer from learning biases and the history of signals is (B, B) , the conditions most favorable for workers' cooperation are such that teams are formed for:

$\gamma\phi \geq 2X_B (\equiv \Theta_{bb}) > 1$ (if a team has been formed at time 0) and $\frac{\gamma}{\phi} \geq 2X_B > 1$ (if a team has not been formed at time 0). As a result, even choosing the allocation rule that maximize workers' cooperation the conditions for team formation at time 1 are more demanding than in the efficient teams equilibrium since: $\Theta_{gb} > 1$, $\Theta_b > 1$ and $\Theta_{bb} > 1$. ■

Proposition 6'. We assume the ability of worker 1 (q_1) is common knowledge. We denote η the share of the group outcome obtained by worker 2. After bad news have been received about worker 2's ability and given that worker 2 suffers from biased recall the maximum level of cooperation is obtained for $\eta = \frac{\left(\frac{\alpha}{\alpha+\beta+1}+q_1\right)\frac{\alpha}{\alpha+\beta}}{\left(\frac{\alpha}{\alpha+\beta}+q_1\right)q_1+\left(\frac{\alpha}{\alpha+\beta+1}+q_1\right)\frac{\alpha}{\alpha+\beta}}$. Given η (allocation rule that maximizes workers' cooperation), the conditions for team formation at time 0 are as follows:

$$\left\{ \begin{array}{l} \gamma \geq \Gamma_p^{(\phi)} > 1 \text{ for } \phi < \phi_l \\ \gamma \geq \frac{\Phi}{\phi} > 1 \text{ for } \phi_l \leq \phi < \Phi \\ \gamma \geq 1 \text{ for } \phi \geq \Phi \end{array} \right\}$$

Where $\Phi = \frac{q_1}{\frac{\alpha}{\alpha+\beta+1}+q_1} + \frac{\frac{\alpha}{\alpha+\beta}}{\frac{\alpha}{\alpha+\beta}+q_1} > 1$.

We have: $\Gamma_p^{(\phi)} = \frac{1+\phi-\frac{\beta p}{\alpha+\beta+1}}{1+\phi-\frac{\phi\beta p}{\alpha+\beta+1}}$, so that $1 < \Gamma_p^{(\phi)} \leq \phi$.

We denote ϕ_l the level of learning by doing such that $\Gamma_p^{(\phi_l)} = \frac{\Phi}{\phi_l}$.

As a result, the *ETO* is not attainable for the following values of parameters:

$$\left\{ \begin{array}{l} 1 \leq \gamma < \Gamma_p^{(\phi)} \text{ and } 1 \leq \phi < \phi_l \\ 1 \leq \gamma < \frac{\Phi}{\phi} \text{ and } \phi_l \leq \phi < \Phi \end{array} \right\} \blacksquare$$

Robustness of proposition 6 to any synergy function. We have assumed until now that the team outcome was: $f(a, X_{I,1,t}, X_{I,2,t}) = a(X_{I,1,t} + X_{I,2,t}), \forall a \in \mathfrak{R}^m$ for $m \geq 1$, but the impossibility result captured in proposition 9 is valid for any function f as long as there exists positive synergies. We consider functions f such that there exists strictly positive synergies (condition *i*) and f is continuous with respect to $X_{I,i,t}$ (condition *ii*).

- (i) $f(a, X_{I,1,t}, X_{I,2,t}) > X_{I,1,t} + X_{I,2,t}$ and $\frac{\partial f(a, X_{I,1,t}, X_{I,2,t})}{\partial X_{I,i,t}} > 1, \forall i \in \{1, 2\}$.
- (ii) $\frac{\partial f(a, X_{I,1,t}, X_{I,2,t})}{\partial X_{I,i,t}}$ is continuous $\forall i \in \{1, 2\}$.

In the presence of positive synergies and in the absence of biased self-attribution, the *ETO* is attainable when the allocation rule is based on relative ability. In the absence of positive synergies $f(a, X_{I,1,t}, X_{I,2,t}) = X_{I,1,t} + X_{I,2,t}$ and in the presence of self-serving learning, the *ETO* will not be achieved in the sense that there exists a vector of parameter values $a \in A \subset \mathfrak{R}^m$ for the function f such that no teams will be formed at time 0. If we assume that f is continuous with respect to synergies, that is (ii) holds, we can conclude that even when synergies are strictly positive there exists $\forall f$ satisfying (*i* – *ii*) a range of parameter values of the synergy function such that the *ETO* cannot be achieved. ■

Proof of Corollary 2. Similar to the proof of corollary 1. ■

Proof of Corollary 3. As a result of proposition 9, there exists no budget balanced contracts that can implement the efficient team equilibrium. Given that in our framework a contract is budget balanced if and only if it is renegotiation-proof (this is a direct consequence of Lemma 1 in Bartling and von Siemens (2004) for $\beta = 0$), for $\eta_1 + \eta_2 = 1$ there exists no renegotiation-proof contracts implementing the *ETO*.

The second part of the corollary is proven below. We consider now the case $\eta_1 + \eta_2 < 1$. We introduce a third party able to observe without biases the outcome of the two other workers. The third agent is rewarded η_3 of the group outcome ($\eta_1 + \eta_2 + \eta_3 = 1$). We consider the case in which a team has been formed at time 0, the same reasoning applies when a team has not been formed at time 0. Assume that $(X_{I,1,1}, X_{I,2,1}) = (B, B)$, $\sigma_1 = (G, B)$ and $\sigma_2 = (B, G)$, this occurs with probability $p^2 p_{BB}$. The third agent contract can be designed at time 0 in order to penalize workers breaking the team. A proportion $(1 - x) \left[(1 - x') \right]$ of the payoffs of the first and second workers is given to the third party if a team is maintained (broken) in the second period, where $(x, x') \in [0, 1]^2$. However, the introduction of a third party being able to penalize team breaks may not be renegotiation-proof.

A necessary condition to obtain the *ETO* is that forming a team is preferred to working alone when $(X_{I,1,1}, X_{I,2,1}) = (B, B)$, $\sigma_1 = (G, B)$ and $\sigma_2 = (B, G)$, where $h_1 + h_2 = 1 = x + xh_1 + xh_2$

$$\left\{ \begin{array}{l} \gamma\phi(q_B + q_G)xh_1 \geq x'q_G \\ \gamma\phi(q_B + q_G)(1 - x - xh_1) \geq x'q_G \end{array} \right\} (Q)$$

A necessary condition to get an *ETE* is $\frac{x'}{xh_1} = \frac{x'}{1-x-xh_1} \Leftrightarrow h_1 = h_2 = \frac{1}{2}$. Then, from the first inequality in (Q) we get that $x' < x$ is a necessary condition for an *ETE*. However, for this contract to be renegotiation-proof, we need that the third party is better-off when a team is actually formed (if not there would be an incentive to write a new contract a time 1), that is: $\gamma\phi(1-x)2q_B \geq (1-x')2q_B$
 $\Leftrightarrow \gamma\phi \geq \frac{1-x'}{1-x} > 1$. If it is not the case, workers can propose an amount of money for the third party to get a proportion of the individual outcome $(1-x'') < (1-x')$. We can conclude that there exists no long term commitment contracts that can implement the *ETO*. ■

Proof of Proposition 7. Notice that for $p = 0$, it is clear that the optimal choice for worker 2 is $\bar{\eta} \leq \frac{1}{2}$. For $p > 0$, the unbiased agent (worker 2) select $\bar{\eta}$ such that worker 1 form a team. It is clear that with probability $p p_{BB}$, worker 2 will ensure a team is formed (after a team has been formed at time 0) if:

$$\begin{aligned} \gamma\phi(q_B + q_G)\eta \geq q_G &\Rightarrow \bar{\eta} = \frac{q_G}{\gamma\phi(q_B + q_G)} \text{ and worker 1 form a team for:} \\ \gamma\phi(1 - \bar{\eta})2q_B \geq q_B &\Leftrightarrow \gamma\phi \geq \frac{q_B + 3q_G}{2(q_B + q_G)}. \text{ For } \gamma\phi < \frac{2q_G}{q_B + q_G}, \text{ the expected pay-} \\ \text{offs for the unbiased worker are higher for } \bar{\eta} > \frac{1}{2} &\text{ than for equal splitting if:} \\ p\gamma\phi(p_{BB}2(1 - \bar{\eta})q_B + p_{GB}(q_B + q_G)(1 - \bar{\eta}) + p_{GG}q_G2(1 - \bar{\eta})) & \\ + (1 - p)\gamma\phi(p_{BB}2(1 - \bar{\eta})q_B + p_{BG}(q_B + q_G)(1 - \bar{\eta}) + p_{GG}q_G2(1 - \bar{\eta})) &\geq \\ p(p_{BB}q_B + p_{GB}q_B + \gamma\phi p_{GG}q_G) & \\ + (1 - p)(\gamma\phi p_{BB}q_B + p_{GB}q_B + \gamma\phi p_{GG}q_G) & \\ \Leftrightarrow p > \frac{\left(\frac{2q_G}{q_B + q_G} - \gamma\phi\right)(p_{BB}q_B + p_{GG}q_G) + (1 - \gamma\phi)(p_{GB}q_B + p_{GB}q_G)}{(\gamma\phi - 1)p_{BB}q_B} &\equiv \bar{p} > 0. \blacksquare \end{aligned}$$

Proof of Proposition 8. First, we show an *ETE* is only possible in a truthful telling equilibrium (*TTE*). Assume the payoff at time 0' is $(X_{I,1,1}, X_{I,2,1}) = (B, B)$ and both agents suffer from self-serving learning (i.e: $\sigma_1 = (G, B)$ and $\sigma_2 = (B, G)$). To have team formation for any $\gamma\phi \geq 1$ ($\frac{\gamma}{\phi} \geq 1$) we need that workers' beliefs converge through thanks to the revelation game. As long as agents' beliefs do not converge, proposition 6 shows that an *ETE* is no attainable. The only way beliefs can converge in the case mentioned above is when both workers recover the true signal (B, B) . That is, the only equilibrium compatible with the *ETE* is a *TTE*.

Second, we prove that a *TTE* implementing the efficient team outcome is not possible for $\gamma\phi < \frac{2q_G}{q_B + q_G}$ ($\frac{\gamma}{\phi} < \frac{2q_G}{q_B + q_G}$) when a team has been formed at time 0. A *TTE* is such that in equilibrium workers reveal their observed signals: $a_i = \sigma_i$ so that beliefs in equilibrium are such that $P[(X_{I,1,1}, X_{I,2,1}) = (a_{12}, a_{21})] = 1$. A *TTE* must satisfy the following conditions stating that workers cannot be worse-off by playing $a_i \neq \sigma_i$ whether a team has been formed or not at time 0, where $p' = P(X_{I,i,1} = B \mid \sigma_i = G) = p \frac{p_B}{p_G}, \forall i \in \{1, 2\}$.

$$\left\{ \begin{array}{l}
\forall (i, j) \in S, \\
[1] \quad \eta_{(BGBG)} \geq \eta_{(ijBG)} \\
[1'] \quad \eta_{(BGBG)} \leq \eta_{(GBij)} \\
[2] \quad (1-p)\eta_{(BBBB)} + p\eta_{(BBBG)} \geq (1-p)\eta_{(ijBB)} + p\eta_{(ijBG)} \\
[2'] \quad (1-p)\eta_{(BBBB)} + p\eta_{(BBBG)} \geq (1-p)\eta_{(ijBB)} + p\eta_{(ijBG)} \\
[3] \quad p'(1-p)\eta_{(GBBB)}2q_B + p'\eta_{(GBBG)}2q_B + (1-p')(1-p)\eta_{(GBGB)}(q_B+q_G) \\
\quad + (1-p)p\eta_{(GBGG)}(q_B+q_G) \geq p'(1-p)\eta_{(ijBB)}2q_B + p'\eta_{(ijBG)}2q_B \\
\quad + (1-p')(1-p)\eta_{(ijGB)}(q_B+q_G) + (1-p')p\eta_{(ijGG)}(q_B+q_G) \\
[3'] \quad p'(1-p)\eta_{(BBij)}2q_B + p\eta_{(GBij)}2q_B + (1-p')(1-p)\eta_{(BGij)}(q_B+q_G) \\
\quad + (1-p')p\eta_{(GGij)}(q_B+q_G) \geq p'(1-p)\eta_{(BBBG)}2q_B + p'\eta_{(GBBG)}2q_B \\
\quad + (1-p')(1-p)\eta_{(BGBG)}(q_B+q_G) + (1-p')p\eta_{(GGBG)}(q_B+q_G) \\
[4] \quad (1-p')\eta_{(GGGG)}2q_G + p'\eta_{(GGBG)}(q_B+q_G) \geq (1-p')\eta_{(ijGG)}2q_G \\
\quad + p'\eta_{(ijBG)}(q_B+q_G) \\
[4'] \quad (1-p')\eta_{(GGij)}2q_G + p'\eta_{(GBij)}(q_B+q_G) \geq (1-p')\eta_{(GGGG)}2q_G \\
\quad + p'\eta_{(GBGG)}(q_B+q_G)
\end{array} \right.$$

$$\Leftrightarrow \left\{ \eta_{(ijkl)} = \eta^*, \forall (i, j, k, l) \in S^2 \right\}.$$

The system above defined a contract (C_{TTE}^0) .

The lower bound for achieving the *ETE* is $\gamma\phi \geq \frac{2q_G}{q_B+q_G} \left(\frac{\gamma}{\phi} \geq \frac{2q_G}{q_B+q_G} \right)$ and corresponds to the case $\eta^* = \frac{1}{2}$. Since $\frac{2q_G}{q_B+q_G} > 1$, we get the impossibility result stated in proposition 10. ■

Proof of Proposition 9. It is easy to check that the following strategies form a pooling equilibrium of our game: workers always play $a_1 = (G, B)$ and $a_2 = (B, G)$ where the beliefs in equilibrium and out of the equilibrium are assumed to be the same as the prior beliefs. In equilibrium: $\eta(a_1, a_2) = \eta$, so that the *ETE* is attained for $\gamma\phi \geq M \equiv \text{Max} \left\{ \frac{\hat{q}_G}{\eta(q_B+\hat{q}_G)}; \frac{\hat{q}_G}{(1-\eta)(q_B+\hat{q}_G)} \right\} \left(\frac{\gamma}{\phi} > M \right)$ when a team has (not) been formed at time 0, where $\hat{q}_G = wq_G + (1-w)q_B$ with $w = \frac{p_G}{p_G+pp_B}$. This pooling equilibrium, being uninformative, is equivalent to an equilibrium obtained in the absence of a revelation mechanism in which at time 0 workers decide on an allocation rule η for sharing the group outcome at the end of the first and second periods. ■

Proof of Corollary 4. From proposition 9, we know that the lowest bound for which the absence of a revelation mechanism leads to the *ETE* is $\frac{2\hat{q}_G}{q_B+\hat{q}_G}$ in the case $\eta = \frac{1}{2}$. Then, for $\gamma\phi \geq \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} \geq \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right)$ when a team has (not) been formed at time 0, there exists no equilibria that dominate the no revelation mechanism equilibrium for $\eta = \frac{1}{2}$. However, for $\gamma\phi < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right)$, the revelation mechanism allows workers to achieve a more cooperative equilibrium in which both workers are better-off. As explained in the main text, for $\gamma\phi < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \left(\frac{\gamma}{\phi} < \frac{2\hat{q}_G}{q_B+\hat{q}_G} \right)$ and $\eta = \frac{1}{2}$ teams are formed when signals are symmetric ($w.p : p_{BB} + p_{GG}$) in the *TTE* whereas they are formed only when signals are symmetric and no biases have occurred ($w.p : (1-p)p_{BB} + p_{GG}$) in the absence of a revelation game. ■

Proof of Proposition 10. A *TTE* is such that $a_i = \sigma_i, \forall i \in \{1, 2\}$. The *TTE* associated to contract (C_{TTE}^1) leads to team formation for:

$(\sigma_1, \sigma_2) \in S^2 / \{(B, B); (B, G)\}; \{(G, B); (B, B)\}; \{(G, B); (B, G)\}\}$. The contract is defined as follows in the case a team has been formed at time 0:

$$\Leftrightarrow \left\{ \begin{array}{l} \gamma\phi\eta_{GBGB} (q_B + q_G) \geq q_G, \gamma\phi(1 - \eta_{GBGB}) (q_B + q_G) \geq q_B \\ \gamma\phi\eta_{G BBB} (2q_B) < q_B, \gamma\phi\eta_{G BBG} (2q_B) < q_B, \gamma\phi\eta_{B BBG} (2q_B) < q_B \\ \eta_{GBGG} = \eta_{GBGB}, \gamma\phi\eta_{BGBG} (q_B + q_G) \geq q_B \\ \gamma\phi(1 - \eta_{BGBG}) (q_B + q_G) \geq q_G \\ \eta_{BGBG} = \eta_{GGBG}, \gamma\phi\eta_{GGGG} (2q_G) \geq q_G \\ \gamma\phi(1 - \eta_{GGGG}) (2q_G) \geq q_G, \gamma\phi\eta_{B BBB} (2q_B) \geq q_B \\ \gamma\phi(1 - \eta_{B BBB}) (2q_B) \geq q_B, \forall (i, j, k, l) \notin S', \eta_{ijkl} = 0 \end{array} \right\} (C_{TTE}^1)$$

$$\Leftrightarrow \left\{ \begin{array}{l} \eta_{GBGB} \in \left[\frac{1}{\gamma\phi} - \frac{q_B}{\gamma\phi(q_B + q_G)}, 1 - \frac{q_B}{\gamma\phi(q_B + q_G)} \right], \eta_{GBGG} = \eta_{GBGB} \\ \eta_{G BBB} \in \left[0, \frac{1}{2\gamma\phi} \right], \eta_{G BBG} \in \left[0, \frac{1}{2\gamma\phi} \right], \eta_{B BBG} \in \left[0, \frac{1}{2\gamma\phi} \right] \\ \eta_{BGBG} \in \left[\frac{q_B}{(q_B + q_G)\gamma\phi}, \frac{q_B}{\gamma\phi(q_B + q_G)} + \frac{\gamma\phi - 1}{\gamma\phi} \right], \eta_{GGBG} = \eta_{BGBG} \\ \eta_{GGGG} \in \left[\frac{1}{2\gamma\phi}, 1 - \frac{1}{2\gamma\phi} \right], \eta_{B BBB} \in \left[\frac{1}{2\gamma\phi}, 1 - \frac{1}{2\gamma\phi} \right] \\ \forall (i, j, k, l) \notin S', \eta_{ijkl} = 0 \end{array} \right\} (C_{TTE}^1)$$

Where:

$$S' \equiv \left\{ \begin{array}{l} (G, B, G, B); (G, B, G, G); (B, G, B, G); (G, G, B, G); (G, G, G, G); \\ (B, B, B, B); (B, B, B, G); (G, B, B, G); (G, B, B, B) \end{array} \right\}$$

The contract (C_{TTE}^1) leads to team formation with probability:

$(p_{GG} + p_{BG} + p_{GB} + p_{BB}(1-p)^2)$. We consider an alternative *TTE* leading to team formation for $(\sigma_1, \sigma_2) \in S^2 / \{(G, B); (B, G)\}$. This equilibrium defines the following contract:

$$\Leftrightarrow \left\{ \begin{array}{l} \eta_{GBGB} \in \left[\frac{1}{\gamma\phi} - \frac{q_B}{\gamma\phi(q_B + q_G)}, 1 - \frac{1}{2\gamma\phi} \right] \Rightarrow \gamma\phi \geq \frac{q_G}{q_B + q_G} + \frac{1}{2} \\ \eta_{GBGG} = \eta_{G BBB} = \eta_{GBGB} \\ \eta_{G BBB} \in \left[0, \frac{1}{2\gamma\phi} \right] \\ \eta_{BGBG} \in \left[\frac{1}{2\gamma\phi}, 1 - \frac{q_G}{\gamma\phi(q_B + q_G)} \right] \Rightarrow \gamma\phi \geq \frac{q_G}{q_B + q_G} + \frac{1}{2} \\ \eta_{GGBG} = \eta_{B BBG} = \eta_{BGBG} \\ \eta_{GGGG} \in \left[\frac{1}{2\gamma\phi}, 1 - \frac{1}{2\gamma\phi} \right], \eta_{B BBB} \in \left[\frac{1}{2\gamma\phi}, 1 - \frac{1}{2\gamma\phi} \right] \\ \forall (i, j, k, l) \notin S', \eta_{ijkl} = 0 \end{array} \right\} (C_{TTE}^{1'})$$

This contract leads to team formation for $(\sigma_1, \sigma_2) \in S^2 / \{(G, B); (B, G)\}$ as long as $\gamma\phi \geq \frac{q_G}{q_B + q_G} + \frac{1}{2}$. To end the proof of proposition 10 we need to show that

there does not exist a contingent contract improving $(C_{TTE}^{1'})$ for $\gamma\phi \geq \frac{q_G}{q_B + q_G} + \frac{1}{2}$. We show below how contingent contracts cannot achieve team formation for all

$(\sigma_1, \sigma_2) \in Z$ when $\gamma\phi < \frac{2q_G}{q_B + q_G}$; we denote:

$Z \equiv \{(B, G); (B, G)\}; \{(G, B); (G, B)\}; \{(G, B); (B, G)\}\}$. This is sufficient to show that no contracts can dominate $(C_{TTE}^{1'})$ for $\gamma\phi \geq \frac{q_G}{q_B + q_G} + \frac{1}{2}$. For $\gamma\phi < \frac{2q_G}{q_B + q_G}$ team formation is obtained for all $(\sigma_1, \sigma_2) \in Z$ if the allocation rule obtained when $(\sigma_1, \sigma_2) = \{(B, G); (B, G)\}$ is different from the one following $(\sigma_1, \sigma_2) = \{(G, B); (G, B)\}$. For this to be the case we need either:

- Worker 1 plays a different action after $\sigma_1 = (B, G)$ and $\sigma_1 = (G, B)$ whereas worker 2 plays the same action (a). Given $(X_{I,1,1}, X_{I,2,1}) = \sigma_1 = (B, G)$, and imposing team formation for all $(\sigma_1, \sigma_2) \in Z_1 \subset Z$, where $Z_1 \equiv \{ \{(G, B); (B, G)\}; \{(B, B); (B, G)\} \}$, we need $\eta_{BGa} = \eta_{BBa} = \eta_{GBa}$. However, for $\eta_{BGa} = \eta_{GBa}$ no teams can be formed for both $(\sigma_1, \sigma_2) = \{(G, B); (G, B)\}$ and $(\sigma_1, \sigma_2) = \{(B, G); (B, G)\}$ when $\gamma\phi < \frac{2q_G}{q_B + q_G}$.

- Worker 2 plays a different action after $\sigma_2 = (B, G)$ and $\sigma_2 = (G, B)$ whereas worker 1 plays the same action. Same reasoning than above to show that team formation for all $(\sigma_1, \sigma_2) \in \{ \{(G, B); (B, G)\}; \{(G, B); (B, B)\} \}$ and all $(\sigma_1, \sigma_2) \in \{ \{(G, B); (G, B)\}; \{(B, G); (B, G)\} \}$ is not possible for $\gamma\phi < \frac{2q_G}{q_B + q_G}$.

- Worker 1 (2) plays a different action after $\sigma_1 = (B, G)$ ($\sigma_2 = (B, G)$) and $\sigma_1 = (G, B)$ ($\sigma_2 = (G, B)$). If we want teams to be formed for all $(\sigma_1, \sigma_2) \in Z_2$, where $Z_2 \equiv \{ \{(G, B); (B, G)\}; \{(B, B); (B, G)\}; \{(G, B); (B, B)\} \}$ then we need to impose

$\eta_{BGBG} = \eta_{GBBG} = \eta_{BGBG}$. However, for $\eta_{BGBG} = \eta_{GBBG} = \eta_{BGBG}$ team formation is not possible for all $(\sigma_1, \sigma_2) \in \{ \{(G, B); (G, B)\}; \{(B, G); (B, G)\} \}$ when team formation is imposed for all $(\sigma_1, \sigma_2) \in \{ \{(G, B); (B, G)\} \}$. If we impose team formation only for the set:

$(\sigma_1, \sigma_2) \in \{ \{(G, B); (B, B)\}; \{(B, B); (B, G)\} \}$, we can get team formation as well for $(X_{I,1,1}, X_{I,2,1}) \in \{(G, B); (B, G)\}$ as long as $\gamma\phi \geq \frac{q_G}{q_B + q_G} + \frac{1}{2}$. As a result, for $\gamma\phi < \frac{q_G}{q_B + q_G} + \frac{1}{2}$, the best contract is the best among (C_{TTE}^1) and (C_{TTE}^0) , and for $\gamma\phi \geq \frac{q_G}{q_B + q_G} + \frac{1}{2}$, the best contract is the best among $(C_{TTE}^{1'})$ and (C_{TTE}^0) . For $\gamma\phi < \frac{q_G}{q_B + q_G} + \frac{1}{2}$, (C_{TTE}^1) dominates (C_{TTE}^0) if:

$$\left\{ \begin{array}{l} \text{Probability of team formation at time 1 under } (C_{TTE}^1) \\ > \\ \text{Probability of team formation at time 1 under } (C_{TTE}^0) \end{array} \right\}$$

$$\Leftrightarrow p_B (p^2 - 2p) + 2p_G > 0$$

Similarly, for $\gamma\phi \geq \frac{q_G}{q_B + q_G} + \frac{1}{2}$, $(C_{TTE}^{1'})$ dominates (C_{TTE}^0) if $-p_B p^2 + 2p_G > 0$. ■

Proof of Corollary 5. The proof follows from proposition 13 since the best fixed allocation rules contract is (C_{TTE}^0) for $\gamma\phi < \frac{2q_G}{q_B + q_G}$. This contract is dominated by contingent contracts respectively by $(C_{TTE}^{1'})$ and (C_{TTE}^1) for $p < \sqrt{\frac{2p_G}{p_B}}$ and $p(2-p) < \frac{2p_G}{p_B}$ where $p(2-p)$ increases in p . The second part of the corollary follows from the existence of an uninformative *BPE* based on fixed allocation rules implementing the *ETO* for $\gamma\phi \geq \frac{2q_G}{q_B + q_G}$ (proposition 12). ■

Proof of Proposition 11. The first part of the proposition follows from simple algebra comparing the expected utility of a worker in the different cases. We consider the case of symmetric contracts so that the expected utility of the two coworkers is the same. This is the most favorable situation for the manager since it is the case in which the expected payoffs for the worse coworker are maximum. The latter makes conditions for hiring a manager less demanding. The symmetric case can be seen as a natural consequence of the assumption of equal bargaining power among coworkers.

The second part of the proposition is proved in the main text. ■

Appendix B

Subgame-perfect Nash equilibria of the standard game. In addition to the subgame-perfect Nash equilibrium stated in proposition 1, the other subgame-perfect equilibria are as follows:

- 1- No workers form teams at time 1 and teams are formed for $\gamma \geq \phi$ at time 0.
- 2- No workers form teams at times 0 and 1.
- 3- Teams are formed at time 1 as stated in lemma 1 and no teams are formed at time 0.

None of the equilibria mentioned above is strict since they involve weakly dominated strategies. In addition they involve strategies that prevent any cooperation in at least one of the two periods. ■

Bayesian Perfect equilibria: non-observability of other's performance outside the team.

We assume that workers cannot observe other's performance when they do not form a team. We consider the case in which the two workers decide simultaneously either to form a team or to work alone. The agreement of the two workers is necessary for team formation. As established in proposition 1, at time 0, teams are always formed for $\gamma \geq \phi$. Thresholds for team formation when no teams have been formed at time 0 are always higher than ϕ . That is for $\gamma < \phi$, if no teams are formed at time 0, no teams will be formed at time 1. We can conclude that the conditions for team formation at time 0 in the case of non-observability of other's performance outside the team are identical to the case of observable outcomes. Any *BPE* has to satisfy the following condition: if $\gamma < \phi$ a worker does not form a team at time 1 if no teams have been formed at time 0. Indeed, it is easy to check that whatever the history of signals considered the strategy consisting of forming a team for $\gamma < \phi$ cannot be a mutual best response to one's coworker strategy. This is so since for $\gamma < \phi$ a worker knows that his coworker accepts to form a team only if he receives bad news. The model is then solved as in the standard case. The conditions for team formation at time 0 only depend on the thresholds for team formation when a team has been formed at time 0, since these conditions do not change with respect to the model with observable signals: propositions 1 and 2 are not modified.

The equilibria of the game are as follows: at time 1 the conditions for team formation are given by lemma 1 if a team has been formed at time 0. If no teams have been formed at time 0, we get the following equilibria at time 1:

- Equilibrium 1

For $\gamma < \phi$, workers will decide to work alone whatever the history of signals, beliefs in equilibrium are the prior beliefs: $\mu_j = p_B, \forall j \in \{1, 2\}$, where μ_j is the belief of worker j about i 's ($j \neq i$) receiving a low payoff (B) at time 1. For $\phi \leq \gamma < \phi^2 \gamma_* \equiv \frac{2\phi}{1+\frac{\alpha}{\alpha+1}}$, workers will form a team at time 1 if the signals received are both failures, beliefs in equilibrium are as follows: $\mu_j = 1, \forall j \in \{1, 2\}$. For $\gamma \geq \frac{2\phi}{1+\frac{\alpha}{\alpha+1}}$, agents will work as a team at time 1 whatever the history of signals, beliefs in equilibrium are the prior beliefs.

- Equilibrium 2

Assuming a team has not been formed at time 0, the conditions for team formation at time 1 are as follows: for $\gamma < \phi$, workers will decide to work alone whatever the history of signals, beliefs in equilibrium are the prior beliefs. For $\phi \leq \gamma < \frac{2\phi}{1 + \frac{\alpha}{\alpha+\beta} \frac{\alpha+\beta+1}{\alpha+1}}$, workers will form a team at time 1 if the signals received are both failures, $\mu_j = 1, \forall j \in \{1, 2\}$. For $\gamma \geq \frac{2\phi}{1 + \frac{\alpha}{\alpha+\beta} \frac{\alpha+\beta+1}{\alpha+1}}$, agents will work as a team at time 1 whatever the history of signals, beliefs in equilibrium are the prior beliefs.

- Equilibrium 3: mixed strategies equilibrium. ■

Appendix C

Specification of an alternative model: no learning by doing effect.

The only difference with the standard model is that $\phi = 1$ and a cost $C > 0$ is incurred by both coworkers when they decide to shift from one type of project to another.

Conditions for team formation at time 1:

Assuming a team has been formed at time 0, the condition for team formation at time 1 for worker i are as follows:

Consider the case $(X_{I,1,1}; X_{I,2,1}) = (G; G)$, the condition for team formation is identical for both workers and is given by: $\gamma \geq 1 - C \frac{\alpha+\beta+1}{\alpha+1} \equiv \gamma_g$. For $(X_{I,1,1}; X_{I,2,1}) \in \{(G; B); (B; G)\}$, a team will be formed if: $\gamma \geq \frac{\alpha+1}{\alpha+1/2} - C \frac{\alpha+\beta+1}{\alpha+1/2} \equiv \gamma_{bg}$. For $(X_{I,1,1}; X_{I,2,1}) = (B; B)$, the condition for team formation is the same for both agents: $\gamma \geq 1 - C \frac{\alpha+\beta+1}{\alpha} \equiv \gamma_b$. If a team has not been formed at time 0 the thresholds for team formation are: $\gamma'_g \equiv 1 + C \frac{\alpha+\beta+1}{\alpha+1}$, $\gamma'_{bg} \equiv \frac{\alpha+1}{\alpha+1/2} + C \frac{\alpha+\beta+1}{\alpha+1/2}$ and $\gamma'_b \equiv 1 + C \frac{\alpha+\beta+1}{\alpha}$.

Conditions for team formation at time 0:

Case 1: $C \geq q_G$ so that: $\gamma_b \leq \gamma_{bg} \leq \gamma_g \leq \gamma'_g \leq \gamma'_{bg} \leq \gamma'_b$. The condition for team formation in that case is $\gamma \geq 1$.

Case 2: $C < \frac{1}{2(\alpha+\beta+1)(2\alpha+1)}$ so that: $\gamma_b \leq \gamma_g \leq \gamma'_g < \gamma_{bg} \leq \gamma'_{bg} \leq \gamma'_b$

For $\gamma < \gamma_g$, no teams are formed.

For $\gamma_g \leq \gamma < \gamma'_g$, teams are formed for $\gamma \geq \frac{2(\alpha^2 + \alpha + \alpha\beta + \beta^2 + \beta/2) + 2C(\alpha\beta + \beta^2 + \beta)}{2(\alpha^2 + \alpha + \alpha\beta + \beta^2 + \beta/2)} \equiv$

$\Gamma_A > 1$

For $\gamma'_g \leq \gamma < \gamma_{bg}$, the condition for team formation is $\gamma \geq \gamma'_G$. That is for case 2, the condition for team formation at time 0 is such that: $\gamma \geq \text{Min} \left\{ \Gamma_A, \gamma'_g \right\} > 1$

Case 3: $\frac{1}{2(\alpha+\beta+1)(2\alpha+1)} \leq C < q_G$ so that: $\gamma_b \leq \gamma_g < \gamma_{bg} \leq \gamma'_g \leq \gamma'_{bg} \leq \gamma'_b$, the conditions for team formation at time 0 are such that:

$$\left\{ \begin{array}{l} \gamma \geq \Gamma_A \text{ for } \Gamma_A < \gamma_{BG} \\ \gamma \geq 1 \text{ otherwise} \end{array} \right\}$$

It follows from the thresholds for team formation at time 0 that team with positive synergies may not be formed some range of values of C and γ . This result is in line with proposition 1. ■

Appendix D

Specification of an alternative model: assumption 2c' (full awareness of biases). We modify our standard model assuming that agents are fully aware of their biases at time 1 (assumption 2c') and considering that with probability x no news $\{\emptyset\}$ are received at time 1.

Conditions for team formation at time 1 after a team has (not) been formed at time 0

Signals	Probability	Threshold
(B, B)	$\begin{bmatrix} p_{BB}(1-x)^2(1-\Pi) \\ p_{BB}(1-x)^2\Pi \end{bmatrix}$	$\begin{bmatrix} \gamma \geq \frac{1}{\phi}(\phi) \\ \gamma \geq \delta_*(\phi^2\delta_*) \end{bmatrix}$
$(\emptyset; B); (B; \emptyset)$	$p_{Bx}(1-x)$	$\gamma \geq \delta_*(\phi^2\delta_*)$
$(B, G); (G, B)$	$p_{BpG}(1-x)^2$	$\gamma \geq \gamma_*(\phi^2\gamma_*)$
(G, G)	$p_{GG}(1-x)^2$	$\gamma \geq \frac{1}{\phi}(\phi)$
$(\emptyset; G); (G; \emptyset)$	$p_{Gx}(1-x)$	$\gamma \geq \delta_{**}(\phi^2\delta_{**})$
$(\emptyset; \emptyset)$	x^2	$\gamma \geq \frac{1}{\phi}(\phi)$

Where p_B (p_{BB}) is the probability of receiving one (two) bad signals a priori. The same notation applies to good signals. We denote: $\delta_* \equiv \frac{2\hat{q}_B}{\phi(\hat{q}_B+q_B)}$ and $\delta_{**} \equiv \frac{2q_G}{\phi(\hat{q}_B+q_G)}$ where $\hat{q}_B = x'q_B + (1-x')q^*$, where $x' = P[X_{I,i,1} = B | \sigma_{ii} = \emptyset] = \frac{pp_B}{pp_B+x}$, $\forall i \in \{1, 2\}$ and σ_{ii} is the signal perceived at time 1 by worker i about his own ability.

Conditions for team formation at time 0:

Case 1: $\frac{1}{\phi} \leq \phi \leq \delta_* \leq \delta_{**} \leq \gamma_* \leq \phi^2\delta_* \leq \phi^2\delta_{**} \leq \phi^2\gamma_*$. For $\gamma < \frac{1}{\phi}$ no teams are formed. For $\frac{1}{\phi} \leq \gamma < \phi$, teams are formed for: $\gamma \geq \bar{\gamma}_{1,\Pi}$.

$$\bar{\gamma}_{1,\Pi} \equiv \frac{(1+\phi)W - [\beta x(1-x)(\alpha+\beta+1) + (1-x)^2\beta(2\alpha+1) + \Pi\beta^2(1-x)^2 + x(1-x)((\alpha+\beta)(\alpha+1) + \alpha(\alpha+\beta+1))]}{W + \phi(\alpha(\alpha+1)(1-x)^2 + x^2(\alpha+\alpha^2+\beta^2) + (1-\Pi)\beta^2(1-x)^2)}$$

where $W = (\alpha + \beta)(\alpha + \beta + 1)$, $1 < \bar{\gamma}_{1,\Pi} \leq \phi$ and $\frac{\partial \bar{\gamma}_{1,\Pi}}{\partial \Pi} > 0$.

For $\gamma \geq \phi$, teams are always formed.

For Case 2: $\frac{1}{\phi} \leq \delta_* < \phi \leq \delta_{**} \leq \gamma_* \leq \phi^2\delta_* \leq \phi^2\delta_{**} \leq \phi^2\gamma_*$

Case 3: $\frac{1}{\phi} \leq \delta_* \leq \delta_{**} < \phi \leq \gamma_* \leq \phi^2\delta_* \leq \phi^2\delta_{**} \leq \phi^2\gamma_*$ and Case 4: $\frac{1}{\phi} \leq \delta_* \leq \delta_{**} \leq \gamma_* < \phi \leq \phi^2\delta_* \leq \phi^2\delta_{**} \leq \phi^2\gamma_*$, the only threshold for team formation at time 0 that depends on Π is: $\bar{\gamma}_{1,\Pi}$. It follows from the

previous information and from $\frac{\partial \bar{\gamma}_{1,\Pi}}{\partial \Pi} > 0$ a result similar to proposition 2: if rational coworkers decide not to form a team at time 0, coworkers suffering from self-serving biases will not form a team at time 0 neither. ■

11 References

- Abegglen, J., 1958, *The Japanese factory*, Cambridge, MA, MIT press.
- Babcock, L., Loewenstein G., Issacharoff S., and C. Camerer, 1995, Biased judgments of fairness in bargaining, *American Economic Review* 85, 5, 1337-1443.
- Babcock, L., G. Loewenstein, 1997, Explaining bargaining impasse: the role of self-serving biases, *The Journal of Economic Perspectives* 11, 1, 109-126.
- Barberis, N., Shleifer, A., and R. Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307-343.
- Bartling and von Siemens, 2004, Efficiency in team production with inequity averse agents, working paper, University of Munich.
- Bénabou, R., J. Tirole, 2002, Self-confidence and personal motivation, *Quarterly Journal of Economics* 117, 3, 871-915.
- Bénabou, R., J. Tirole, 2004, Willpower and personal rules, *Journal of Political Economy* 112, 4, 848-887.
- Beyer, S., 1990, Gender differences in the accuracy of self-evaluations of performance, *Journal of Personality and Social Psychology* LIX, 960-970.
- Box, G., G. Tiao, 1973, *Bayesian inference in statistical analysis*, Reading MA: Addison Wesley.
- Bradley, G. W., 1978, Self-serving biases in the attribution process: a reexamination of the fact or fiction question, *Journal of Personal and Social Psychology* 36, 56-71.
- Brown, J. D., 1986, Evaluations of self and others; self-enhancement biases in social judgements, *Social Cognition* 4, 353-376.
- Che, Y. and S. Yoo, 2001, Optimal incentives for teams, *American Economic Review* 91, 525-541.
- Cohen, D., L. Prusak, 2001, *In good company: how social capital makes organizations work*, Harvard Business School Press.
- Cooper, A., Woo, C., and W. Dunkelberg, 1988, Entrepreneurs' perceived chances for success, *Journal of business venturing* 3, 97-108.
- Corgnet, B., 2005b, Team formation and biased self-attribution: a companion version, working paper, Universidad Carlos III.

Corgnet, B., 2005c, Reputation and competition in team formation, working paper, Universidad Carlos III.

Daniel, K., Hirshleifer, D., and A. Subrahmanyam, 1998, Investor psychology and security market under-and overreactions, *Journal of Finance* 53, 1839-85.

Eaton, B. C., A. Hollis, 2003, Private information and teamwork, working paper.

Englmaier, F. and A. Wambach, 2002, Contracts and inequity aversion, working paper, University of Munich.

Epstein, 1983, The unconscious, the preconscious, and the self-concept. In J. Suls and A. Greenwald (Eds.), *Psychological perspectives on the self*, vol 2, 220-247, Hillsdale, NJ: Erlbaum.

Farrell, J., S. and S. Scotchmer, 1988, Partnerships, *Quarterly Journal of Economics* 103, 2, 279-297.

Felson, R. B., 1984, the effects of self-appraisals of ability on academic performance, *Journal of Personality and Social Psychology* 47, 944-952.

Fiske, S. T., S. E. Taylor, 1991, *Social cognition*, McGraw-Hill, 2nd edition.

Florida, R., M. Kenney, 1991, Transplanted organizations: the transfer of Japanese industrial organizations to the US, *American Sociological Review* 56, 381-398.

Gilbert, D. J., Pinel, E. C, Wilson, T. D, Blumberg, S. J., and T. Wheatley, 1998, Immune neglect: a source of durability bias in affective forecasting, *Journal of Personality and Social Psychology* 75, 617-638.

Gervais, S., I. Goldstein, 2004, The effects of biased self-perception in teams, working paper.

Greenwald, A. G., 1980, The totalitarian ego: fabrication and revision of personal history, *American Psychologist* 35, 603-618.

Griffin, D., A. Tversky, 1992, The weighing of evidence and the determinants of overconfidence, *Cognitive Psychology* 24, 411-435.

Haitani, K., 1990, The paradox of Japan's groupism: threat to future competitiveness? *Asian Survey* 30, 237-250.

Heine, S. J, Lehman, D. R., Markus, H. R., and S. Kitayama, 1999, Is there a universal need for positive self-regard? *Psychological Review* 106, 4, 766-794.

Heine, S. J., Kitayama, S., and D. R. Lehman, 2001, Cultural differences in self-evaluation: Japanese readily accept negative self-relevant information, *Journal of Cross-Cultural Psychology* 32, 4, 434-443.

- Holmstrom, B., 1982, Moral hazard in teams, *Bell Journal of Economics* 13, 324-340.
- Ishida, J., 2005, Incentives in academics: collaboration under weak complementarities, working paper.
- Itoh, H., 1991, Incentives to help in multi-agent situations, *Econometrica* 59, 3, 611-636.
- Kandel, E., E. P., Lazear, 1992, Peer pressure and partnerships, *Journal of Political Economy* 100, 801-817.
- Kitayama, S., Markus, H. R., Matsumoto, H., and Norasakkunkit V., 1997, Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan, *Journal of Personality and Social Psychology* 72, 1245-1267.
- Koike, K., 1988, *Understanding industrial relations in modern Japan*, New York, St Martin's.
- Korner, I., 1950, *Experimental Investigation of some aspects of the problem of repression: repressive forgetting*, NO.970. New York, NY: Bureau of Publications, Teachers' College, Columbia University.
- Kuiper, N. A., P. A. Derry, 1982, Depressed and nondepressed content self-reference in mild depression, *Journal of Personality* 50, 67-79.
- Kuiper, N. A., M. R. MacDonald, 1982, Self and other perception in mild depressives, *Social Cognition* 1, 233-239.
- Kuiper, N. A., Olinger, L. J., MacDonald M. R., and B. F. Shaw, 1985, Self-schema processing of depressed and nondepressed content: the effects of vulnerability on depression, *Social Cognition* 3, 77-93.
- Lee, P. M., 1997, *Bayesian statistics: an introduction*, 2nd edition, London: Arnold.
- Lewinsohn, P. M., Mishel, W., Chaplin, W., and R. Barton, 1980, Social competence and depression: the role of illusory self-perceptions, *Journal of Abnormal Psychology* 89, 203-212.
- Liebowitz, S. and J. Palmer, 1983, Assessing assessments of the relative quality of Economics departments, *Quarterly Review of Economics and Business* 28, 2, 88-113.
- Long, J. S and R. McGinnis, 1982, On adjusting productivity measures for multiple authorship, *Scientometrics* 4, 5, 379-387.

- McDowell, J. and J. Smith, 1992, The effect of gender-sorting on propensity to coauthor: implications for academic promotion, *Economic Inquiry* 30, 68-82.
- Miller, D. T. and M. Ross, 1975, Self-serving biases in the attribution of causality: fact or fiction? *Psychological Bulletin* 82, 213-225.
- Mishel, W., Ebbesen, E. B., and A. R. Zeiss, 1976, Determinants of selective memory about the self, *Journal of Consulting and Clinical Psychology*, XLIV, 92-103.
- Nisbet, R. E. and L. Ross, 1980, *Human inference: strategies and shortcomings of social judgments*, Englewood Cliffs, NJ: Prentice-Hall.
- Rabin, M., 2002, Inference by believers in the law of small numbers, *Quarterly Journal of Economics* 117, 3, 775-816.
- Rabin, M., and J. Schrag, 1999, First impressions matter: a model of confirmatory bias, *Quarterly Journal of Economics* 114, 1, 37-82.
- Rey Biel, P., 2004, Inequity aversion and team incentives, working paper, University College London.
- Rotemberg, J. J., 1994, Human relations in the workplace, *Journal of Political Economy* 102, 684-717.
- Ross, L., 1977, The intuitive psychologist and his shortcomings: distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* 10, 174-221, New York: Academic Press.
- Salanié, B., 1997, *The Economics of contracts: a primer*, Cambridge Massachusetts: MIT Press, Chapter 6.
- Schinski, M., Kugler, A., and W. Wick, 1998, Perceptions of the academic Finance profession regarding publishing and the allocation of credit in coauthorship situations, *Financial Practice and Education* 8, 1, 60-68.
- Silverman, I., 1964, Self-esteem and differential responsiveness to success and failure, *Journal of Abnormal and Social Psychology* 69, 115-119.
- Svenson, O., 1981, Are we all less risky and more skillful than our fellow drivers?, *Acta Psychologica* 47, 143-148.
- Taylor, S. E., and J. D. Brown 1998, Illusion and well-being: a social psychological perspective on mental health, *Psychological Bulletin* 103, 193-210.
- Zuckerman M., 1979, Attribution of success and failure revisited or, : the motivational bias is alive and well in attribution theory, *Journal of Personality* 47, 245-285.