



## **Friends or Foes? Optimal Incentives for Reciprocal Agents**

Luca Livio  
ECARES, SBS-EM, Université libre de Bruxelles

January 2018

**ECARES working paper 2018-03**

# Friends or Foes? Optimal Incentives for Reciprocal Agents<sup>☆</sup>

Luca Livio<sup>a</sup>

<sup>a</sup>ECARES, SBS-EM, Université libre de Bruxelles, 50 Avenue Roosevelt, Brussels 1050, Belgium

---

## Abstract

Widely used performance-based contracts put (positive or negative) externalities on co-workers. These externalities have been proven to shape an organization's working climate especially when workers exhibit social preferences. However, it is a priori unclear whether a more friendly or a more competitive working environment should be encouraged. In this paper I consider a theoretical model in which a self-interested principal has to motivate a set of agents. Agents are symmetric, potentially risk-averse and exhibit reciprocity concerns towards each other. The optimal incentive scheme is derived solving a psychological game with asymmetric information about effort choices. I show that the optimal incentive design depends on the interplay between the agents' attitudes towards risks and their preferences for reciprocity. In particular, the optimal scheme implements (i) a relative performance compensation scheme which induces an exchange of unkindness if agents are relatively little risk averse and (ii) a joint performance compensation scheme which induces an exchange of kindness if agents are sufficiently risk averse. My findings can explain some puzzling empirical results.

*Keywords:* Gift Exchange, Group Production, Incentives, Moral Hazard, Reciprocity, Team, Tournament.  
*JEL:* D82, D86, J41, M12, M52, M54

---

---

<sup>☆</sup>I am particularly indebted to Georg Kirchsteiger for his careful supervision and continuous support and to Bram De Rock, Martin Dufwenberg, Patrick Legros, Mathieu Parenti, and Alexander Sebald for helpful discussions. I would also like to thank Gani Aldashev, Tommaso Aquilante, Charles Bellemare, Estelle Cantillon, Alessandro De Chiara, Philippe Jehiel, Sabine Kröger, Marco Mariotti, Luca Merlino, Tom Potoms, Claudio Schioppa, Alex Teytelboym, Stéphane Wolton, and audiences at Birmingham Business School, Central European University Budapest, ECARES, Goethe University Frankfurt, Maastricht University, Mannheim University, ShanghaiTech University, Universität Innsbruck, Université Laval, ECORES 2015 Summer School, 2016 IMEBESS Conference, for useful comments and discussions. Usual disclaimer applies.

*Email address:* llivio@ulb.ac.be (Luca Livio)

## 1. Introduction

Human resource managers adopt a huge variety of compensation policies in a bid to align employees' interests and to achieve organizational goals (Murphy 1999, Indjejikian and Nanda 2002, Merchant and Otley 2007, Oyer and Schaefer 2011). Whether the compensation policy of an organization should emphasize friendship and cooperation or competition among its employees remains a debated controversy in the management literature. Such discussion involves psychological aspects that have typically been disregarded by the standard agency theory.<sup>1</sup> Some scholars and practitioners believe that competition promotes efficiency and innovation because it stimulates individuals to outperform each other by working harder. According to Dick Grote, well-known chairman and CEO of the Grote's consultancy, one of the business leaders' main arguments in favor of relative incentives is the motivating power entailed by competition (Grote and Grote 2011). In the US, around 60% of the corporations adopt some sort of relative incentive systems such as stack-ranking based tournaments that pit employees against each other (Berger et al. 2013, Chen et al. 2012). From production-line workers to fund managers, many employees strive to compete for bonuses, promotions and even luxury prizes (Backes-Gellner and Pull 2013, Cerdin and Pargneux 2009, Kempf and Ruenzi 2008). Other scholars believe that intra-team competition is destructive and that the needs of organizations are better met by employing some joint-reward structures such as group reward and gain sharing (Banks et al. 2014, Kruse et al. 2010. See also Bloom and Van Reenen 2010 for a review). In such systems, a worker's compensation increases in the performance indicators of all members of the reference group. This allows employers to motivate non-trivial effort provision and to enhance cooperation, support and friendship among colleagues.<sup>2</sup>

According to social psychologists, friendship, cooperation and competitive behaviors are directly related to the psychological construct of reciprocity. The main idea of reciprocity is both simple and appealing: Rewarding those one believes have behaved nicely to her (positive reciprocity) and punishing those one believes have behaved badly to her (negative reciprocity).<sup>3</sup> As shown by experimental economists and labor psychologists, reciprocity is likely to play a major role in most human organizations such as partnerships and hierarchical firms.<sup>4</sup>

In this paper I introduce preferences for reciprocity into an otherwise standard principal-multi agents model of moral hazard.<sup>5</sup> In particular I concentrate on the case in which agents exhibit reciprocity concerns towards each other (horizontal reciprocity).<sup>6</sup> Solving a psychological game I show that the principal can reduce the cost of achieving high effort provision implementing two different types of interdependent com-

---

<sup>1</sup>In particular, early contributions to agency theory typically focused on how the optimal design of incentives depends on the interplay between the risk-aversion of the employees and the information the principal achieves given a certain production technology (Holmström 1979, Holmström 1982, Mookherjee 1984). Notable exceptions are the contributions of Lazear (1989) and Itoh (1991).

<sup>2</sup>In a field experiment, Petty et al. (1992) compared one division of an electric utility company that set up a gain sharing plan with another that did not. The gain sharing division performed better than the other one in 11 out of 12 aggregate objective performance measures, providing an estimated savings between \$ 875,000 and \$ 2 million. In addition, employees' perception of team-spirit was positively affected by gain-sharing also for workers charged with individual tasks.

<sup>3</sup>The basic concept of reciprocity is a fundamental principle in most human societies throughout history and it is embedded in several laws and religious texts (Gouldner 1960). In recent years, scholars in many disciplines referred to reciprocity to explain a wide range of phenomena.

<sup>4</sup>For instance, contributions in organizational behavior, management, and leadership recognized reciprocity as one of the key variables to predict many organizational outcomes (Cropanzano and Mitchell 2005, Liden et al. 1997, Bosse et al. 2009). Experimental economists have called for reciprocity to explain many social dilemmas both in the lab and in the field. See Fehr and Schmidt 2003 and DellaVigna 2009 for reviews. Bowling et al. (2004) have shown that positive reciprocity can be seen as the psychological antecedent of friendship and support. Similarly, negative reciprocity may represent an antecedent for competitive behaviors. In behavioral psychology, an antecedent is a stimulus that cues an organism to perform a learned behavior. When an organism perceives an antecedent stimulus, it behaves in a way that maximizes reinforcing consequences and minimizes punishing consequences. Perugini et al. (2003) define reciprocity as a norm to which an individual tends to conform even when she is not observed or externally sanctioned.

<sup>5</sup>To model preferences for reciprocity and to solve the psychological game I draw extensively from Dufwenberg and Kirchsteiger (2004).

<sup>6</sup>Previous theoretical research in economics (Akerlof 1982, Dufwenberg and Kirchsteiger 2000, Englmaier and Leider 2012) concentrate on the case of vertical reciprocity, i.e., on a situation in which reciprocal employees repay generous compensation packages offered by firms with greater effort provision.

compensation policies:<sup>7</sup> (i) a relative-incentive scheme which sustains a negative-reciprocity equilibrium and competition among coworkers and (ii) a joint-incentive scheme which sustains a positive-reciprocity equilibrium and stimulates friendship among coworkers. In both cases, the principal is able to replace monetary with psychological incentives by inducing a psychological gift-exchange.<sup>8</sup> I highlight that the final choice of the principal depends on a non-trivial interplay between the agents' psychological preferences and their attitudes towards risks. In particular, the optimal incentive scheme implements a high-powered relative-performance policy if agents are not too risk-averse while it implements a low-powered joint-performance compensation if they are sufficiently risk-averse.

In the baseline model the principal employs two identical agents to exert some productive effort. Agents are reciprocal towards each other. I assume that inducing high effort from both agents is always valuable for the principal. Each agent independently and non-cooperatively makes a binary and non-verifiable effort choice. Production is technologically independent, performance indicators are uncorrelated and satisfy the monotone likelihood ratio property.<sup>9</sup>

Since agents are reciprocal, the incentives provided by an incentive scheme are not restricted to the monetary compensations explicitly envisaged. Interdependent compensation schemes provide some *implicit psychological motives* linked to the externalities produced by an agent's effort choice on his coworker's material payoff. On the one hand, the competition entailed by relative-performance incentive schemes implies that high effort provision by agent  $i$  is seen as a nasty action by agent  $j$ , whose disutility of exerting effort is now lowered by his willingness to retaliate against  $i$ , choosing high effort himself. On the other hand, under joint-performance incentive schemes working hard by one agent is perceived as a kind action by his colleague, who is now more willing to put effort into production so as to repay his colleague. These desirable psychological effects are not generated by schemes in which an agent's compensation solely depends on his own indicator (individual-performance schemes).<sup>10</sup> Therefore: (i) the principal finds it optimal to use interdependent compensation schemes even though performance indicators are uncorrelated, thereby contradicting the informativeness principle (Holmström 1982, Mookherjee 1984); (ii) the principal's net profit increases in the agents' concerns for reciprocity.<sup>11</sup>

The final choice between relative and joint incentives depends on the agents' aversion towards variation in their income level. When agents are risk neutral, a high-powered *relative performance compensation scheme* which induces negative reciprocity and competition among agents always proves superior. The reason for this result is subtle. With team incentives agent  $i$  gets a bonus only if the effort of both agents is signalled to be high, while with relative incentives the bonus is paid only if agent  $i$ 's effort is signalled to be higher than  $j$ 's one. Since it is more likely that the signal is positive for both agents than that the signal is positive only for one particular agent, the bonus needed to overcome the effort cost must be higher in case of relative incentives than in case of team incentives. This implies that the externality induced by high effort is greater under relative (negative externality) than under joint (positive externality) incentives. Therefore, an agent is more keen to reciprocate the negative externality of his coworker's high-effort in case of relative incentives than to reciprocate the smaller positive externality of his coworker's choice in case of joint incentives. Thus, the psychological motivating power of the best-relative incentive scheme is endogenously stronger than that of the best-joint incentive scheme.

When risk-aversion is added to the picture, agents must be compensated for bearing the risk of a variable monetary compensation. As risk-aversion increases, it becomes more and more costly to induce a

---

<sup>7</sup>In interdependent compensation policies, worker  $i$ 's monetary rewards does not only depend on  $i$ 's individual performance indicator, but also the performance indicators of  $i$ 's peers. Therefore, worker  $i$ 's effort choice exhibits externalities on his coworkers' material payoffs no matter the underlying production technology.

<sup>8</sup>The idea that workers are willing to give up part of their monetary pay if compensated with some non-material rewards is consistent with evidence from many field experiments. See for instance Hamilton et al. (2003).

<sup>9</sup>This standard assumption simply implies that the higher the value of the individual performance indicator, the more likely it is that the agent worked hard.

<sup>10</sup>This is because in such individual-performance schemes the effort of an agent does not entail any material consequence on his coworker.

<sup>11</sup>In Section 4 I show that when the returns of efforts are small enough, being more reciprocal can also entail some material advantage for the agents.

psychological-gift exchange through a high-powered relative incentive scheme. A joint-incentive solution becomes comparably more attractive, as it allows the principal to better balance between the benefits (coming from the agents' reciprocal concerns) and the costs (coming-from risk-aversion) of implementing interdependent compensation policies. Therefore, there exists a threshold of the parameter of relative risk aversion above which the principal prefers to pay joint rather than relative compensation schemes. This threshold (weakly) increases in the precision of the performance indicators and in the agents' inherent concerns for reciprocity.

In an extension of the baseline model I show that the interplay between risk attitudes and preferences for reciprocity continues to hold when agents are engaged in joint production and measures of the individual performances are available to the principal.

My paper contributes to the long-lasting debate on the optimal incentive design in a principal-multi agent setting. Classical contributions to agency theory mainly focused on the technological aspects of the production process.<sup>12</sup> Holmström (1982) shows the importance of interdependent incentives in filtering out common shocks when individual performance indicators are subject to some common uncertainty. An agent compensation optimally depends on his peers' signals only if this allows the principal to better infer the state of world, thereby reducing an agent's risk exposure. In this paper I complement this theory by arguing that interdependent compensation policies can be used to provide some additional psychological motivation to employees. Lazear (1989) and Itoh (1991) motivate the use of joint incentives to reduce the risk of sabotage and to stimulate cooperative efforts when agents' tasks are interdependent. Che and Yoo (2001) show that in a repeated interaction setting the choice between relative and team incentives crucially depends on the patience of the productive agents. Kim (2015) examines incentive contracts when the principal and the agents disagree on the likelihood that a task succeeds. The author shows that the direction of disagreement alters the effectiveness of different monetary incentive schemes.<sup>13</sup>

The degree to which organizations should emphasize cooperation or competition among the employees is an age-old controversy in the management literature (see Bloom et al. 2011 for a discussion). My results bring some food for thought to this debate suggesting that firms should take into account the interplay between the workforce risk and social attitudes when deciding whether to foster friendship or competition.<sup>14</sup>

This paper also contributes to the burgeoning literature on behavioral contract theory.<sup>15</sup> My work is closely related to the pioneering article by Itoh (2004). The author considers a principal-multi agents framework to analyze the effects of social comparisons among coworkers that exhibit distributional concerns.<sup>16</sup> He finds that if agents are sufficiently inequity averse, pure-team incentives are optimal as they provide insurance against the disutility of being worse or better off than the coworkers, i.e., team incentives in this setting are paid to avoid some extra costs that would result from an ex-ante uneven income distribution. Conversely, he shows that relative-incentive policies are preferred by the principal and have some positive motivating effects when agents are sufficiently status seeking (i.e., when they like to earn more than their coworkers). Whereas in Itoh's model the principal's choice exclusively depends on the type of distributional preferences the workforce is endowed with, in my paper it hinges on whether it is endogenously more efficient to design a positive or a negative gift-exchange game given the agents preferences towards risk.<sup>17</sup>

---

<sup>12</sup>Kandel and Lazear (1992) and Barron and Gjerde (1997) works on peer pressure and Rotemberg and Saloner (1993) and Rotemberg (1994) works on leadership and human relations at the workplace are notable exceptions.

<sup>13</sup>In particular, the optimal contract implements relative performance evaluation when the principal is more optimistic than the agents, and joint performance evaluation when she is less optimistic.

<sup>14</sup>A different perspective is taken by the theoretical contribution of Dur and Sol (2010). The authors consider a setting in which social interactions among colleagues generate conditional altruistic preferences. They show that a combination of team and relative incentives can be used as an instrument to achieve the first-best level of social interaction.

<sup>15</sup>Spurred by experimental findings, this strand of research aims at incorporating behavioral decision making that departs from the self-interest axiom into otherwise standard contract theory settings. See Köszegi (2014) for a review.

<sup>16</sup>Notable contributions related to Itoh's work are Rey-Biel (2008), Bartling and von Siemens (2010), Bartling (2011), Englmaier and Wambach (2010), Dur and Glazer (2008).

<sup>17</sup>Although in many situations reciprocal players and inequity averse players behave in a similar way, intention-based reciprocity and inequity aversion are distinct motives (Fehr and Fischbacher 2002). Moreover, Fehr and Schmidt (2003) argue that reciprocity plays a bigger role in driving individuals decision making than pure distributional concerns.

To model reciprocity I draw extensively on the literature on psychological games. Motivated by experimental results, in the last few decades economists and game theorists have formally conceptualized preferences for reciprocity. Rabin (1993) developed a model to incorporate reciprocity in normal form games adopting a psychological game theory setting (see Geanakoplos et al. 1989, Battigalli and Dufwenberg 2009). Dufwenberg and Kirchsteiger (2004) extend Rabin’s analysis to extensive form games.<sup>18</sup> Finally, Sebald (2010) presents a framework that allows to consider chance moves. In these models reciprocal agents hold beliefs about others’ intentions as well as beliefs about others’ possible alternatives. If for instance a player  $i$  believes that the intentions of another player  $j$  are benign (malign) relative to his possible alternatives, he perceives him as kind (unkind). Given this, reciprocal agents achieve some positive psychological payoff reacting kindly to perceived kindness and unkindly to perceived unkindness. I draw on Dufwenberg and Kirchsteiger (2004) and Sebald (2010) building a model in which each agent’s effort choice impact on the distribution of the expected material payoff of his coworker. Agent  $i$ ’s effort choice is considered *kind* if  $i$  expects to give to  $j$  a material payoff greater than the one that would have been entailed if  $i$  had taken the alternative effort choice.

Previous contributions such as Dufwenberg and Kirchsteiger (2000), Non (2012) and Englmaier and Leider (2012) have considered principal-single agent setting in which the agent exhibits reciprocity concerns towards the principal. They show that generous compensation can substitute for performance based pay, consistently with the literature on perceived organizational support (see Pazy and Ganzach 2008) and with evidence from stylized labor markets experiments as well as with empirical research based on representative survey data.<sup>19</sup> Englmaier and Leider (2012) show that if the agent is both sufficiently motivated by reciprocity and risk-averse, then generous flat contracts can be optimal even if effort is non contractible. In this paper I show that the principal can take advantage of the presence of reciprocity concerns by inducing a gift-exchange game between the firm’s employees irrespective of the agents’ risk aversion. In a companion project, (Livio 2015), investigate how the choice of the production technology by the principal is affected by the interplay between risk and reciprocal attitudes of the agents. Sebald and Walzl (2015) present a principal-reciprocal agent model showing that agents might engage into conflict in response to ego-threatening performance appraisals by the principal. The authors identify conditions for a positive welfare effect of increasing costs of conflict and a negative welfare effect of more capable agents.<sup>20</sup>

The rest of this paper is organized as follows. Section 2 presents the main features of the model. Section 3 characterizes the optimal incentive scheme available to the principal. Section 4 relaxes the assumption that inducing high effort is always profitable for the principal. Section 5 consider a setting in which agents are engaged in joint production. Section 6 discusses the main results of the paper. Section 7 concludes. Appendix A contains all the proofs of Propositions and Lemmas. Appendix B contains a robustness check and extensions of the model.

## 2. The model

In a principal-multiple agents setting three types of incentive schemes are feasible: (i) *relative-incentive schemes* (RI), that penalize an employee when his peer performs well; (ii) *joint-incentive schemes* (JI), that reward an employee when his peer performs well; (iii) *individual-incentive schemes* (II), that does not take into account the peer performance indicator. A special case of RIs is the tournament scheme, in which an agent is rewarded if only if his performance is signalled to be the best one among those of his reference group, while a special case of JIs is the team scheme, in which an agent is rewarded if and only if all the team members perform well.

---

<sup>18</sup>Falk and Fischbacher (2006) also propose a model of reciprocity, but their definition does not depend solely on intentions but also on distributional concerns.

<sup>19</sup>See among the others Fehr et al. (1993), Brown et al. (2004), Bellemare and Shearer (2009), Bellemare and Shearer (2011), and Englmaier et al. (2014) for experimental evidence and Dohmen et al. (2009) and Englmaier et al. (2015) for evidence from survey data. In these latter cases causality has not been assessed.

<sup>20</sup>In a related paper, Sebald and Walzl (2014) experimentally show that individuals tend to sanction others who subjectively evaluate their performance whenever this assessment falls short of the individuals’ self-evaluation. This occurs even if the individuals’ earnings are unaffected by the subjective performance appraisal.

When agents are reciprocal the incentives provided by a certain scheme may not be restricted to the monetary incentives explicitly envisaged by the incentive scheme. Indeed, both RI and JI provide implicit psychological incentives, as one agent's effort choice inflicts some externality on his coworker's material payoff. On the one hand, the competition entailed by RI implies that high effort provision by agent  $i$  is seen as an unkind action by agent  $j$ , whose disutility of exerting effort is now lowered by his willingness to retaliate against  $i$  choosing high effort himself. On the other hand, under JI working hard by one agent is perceived as a kind action by his peer, who is now more willing to put effort into production so as to reciprocate. These psychological effects are not generated by IL, as in this case the fates of the agents are not tied together. Straightforwardly, optimal incentive schemes for reciprocal agents are more likely to be interdependent than the ones for self-interested agents.

However, the circumstances in which one type of interdependent incentive scheme is preferred to the other are not obvious a priori. In what follows, I present a simple model to highlight how the principal's choice depends on the interplay between the reciprocal preferences of the agents and their attitudes towards risks.

### 2.1. Production setting, information and sequence of events

There are two identical agents,  $a$  and  $b$ , who work for a firm. Each agent  $i \in \{a, b\}$  chooses effort levels  $e_i \in \{0, 1\}$  at a cost  $e_i * \phi$ , measured in monetary units. Throughout, I assume that agents' *hard working* is sufficiently valuable to the owner of the firm (the principal) that she always prefers to induce the agents to work hard (i.e.,  $e_i = 1, i \in \{a, b\}$ ). Hence, the principal's problem is to minimize the cost of motivating the agents.

Material incentives can be provided by the wages the principal offers on the basis of individual performance indicators that take the form  $x_i \in \{h, l\}$ . Production is technologically independent such that each agent affects only his own indicator. Indicators are assumed to be uncorrelated and the probability of getting a *good* indicator ( $x_i = h$ ) or a *bad* indicator ( $x_i = l$ ) solely depends on the individual effort of agent  $i$ :  $x_i = h$  (respectively,  $x_i = l$ ) with probability  $q \in (1/2, 1)$  if  $i$  worked (resp., shirked). Each  $x_i$  can be interpreted as the principal's (gross) profit generated by agent  $i$ 's activity. In general, each agent wage can depend on both performance indicators and must be non-negative.

To ensure that the principal's expected cost always increases in the agent's risk-aversion I assume that  $\phi > (2q - 1)$ .<sup>21</sup>

To focus on incentive provision I abstract from the participation decision of the employees and I assume that they already work for the firm.

The sequence of events is as follows:

1. The principal sets the incentive scheme, which consists of a schedule of wages contingent on (possibly) both performance indicators.
2. Both agents simultaneously and non-cooperatively choose their efforts.
3. The nature moves,  $x_a$  and  $x_b$  realize and transfers take place.

### 2.2. Preferences

**Material Utility.** The principal is risk neutral and solely interested in her expected profit.

Each agent  $i$  evaluates his salary  $w_i$  through a CRRA function  $u(w_i) = w_i^{1-\rho}$ , where  $\rho \in [0, 1)$  indicates the coefficient of relative risk-aversion.<sup>22</sup> At every end node of the game agent  $i$  achieves a material utility given by:

$$\pi(w_i, e) = w_i^{1-\rho} - e_i \phi.$$

<sup>21</sup>This assumption eases the computations needed to prove the main results. Relaxing this assumption does not affect the main results of the paper.

<sup>22</sup>Assuming a specific functional form allows us to explicitly characterize the solution to the principal problem. My results are robust to assuming that the function  $u(\cdot)$  exhibits a Constant Absolute Risk Aversion.

**Psychological Utility.** Agents are reciprocal towards each other. When agent  $i$  is called to take his effort decision, he takes into account that this affects not only his own material payoff but also the one of his coworker. Agent  $i$  believes to be kind towards  $j$  if and only if the expected material utility of  $j$  as a result of  $i$ 's effort choice is greater than what  $i$  thinks could be a fair value, given his system of beliefs. Following Dufwenberg and Kirchsteiger (2004), the reference point for  $i$ 's kindness is given by the simple average between  $j$ 's expected material utility when  $e_i = 1$  and  $j$ 's expected material utility when  $e_i = 0$ , that is:

$$\pi_j^{eq}(\beta_i(e_j)) = \frac{1}{2} [\pi_j(e_i = 1, \beta_i(e_j)) + \pi_j(e_i = 0, \beta_i(e_j))]$$

where  $\beta_i(e_j)$  represents  $i$ 's first-order belief about  $j$ 's effort choice, and  $\pi_j(e_i, \beta_i(e_j))$  is the  $i$ 's expectation of  $j$ 's material payoff given  $\beta_i(e_j)$  and  $e_i$ . Hence,  $i$ 's expected kindness towards  $j$  when  $i$  chooses effort  $e_i$  is given by:

$$k_{ij}(e_i, \beta_i(e_j)) = \pi_j(e_i, \beta_i(e_j)) - \pi_j^{eq}(\beta_i(e_j))$$

Similarly, agent  $i$ 's belief about  $j$ 's kindness to  $i$  is:

$$\lambda_{iji}(\beta_i(e_j), \chi_i(e_i)) = \pi_i(\beta_i(e_j), \chi_i(e_i)) - \pi_i^{eq}(\chi_i(e_i))$$

where  $\chi_i(e_i)$  represents  $i$ 's second-order belief about  $j$ 's belief about  $i$ 's effort choice. Agent  $i$ 's psychological utility is given by:

$$\gamma r_i(\lambda_{iji}, k_{ij}) = \gamma * \text{sign}(\lambda_{iji} * k_{ij}) * \sqrt{|\lambda_{iji} * k_{ij}|}$$

where the parameter  $\gamma$  is an element of the interval  $[0, 1]$  and represents the agents inherent concern for reciprocity. Note that the formulation above slightly differs from the multiplicative form  $(\gamma r_i(\lambda_{iji}, k_{ij}) = \gamma * \lambda_{iji} * k_{ij})$  typically assumed in the literature (e.g., Rabin, 1993, Dufwenberg and Kirchsteiger, 2004, Sebald, 2010, von Siemens, 2013). Dufwenberg and Kirchsteiger (2004) suggested the use of this alternative formulation in order to make  $i$ 's preferences invariant with respect to the choice of monetary units.<sup>23</sup> In my setting, this formulation along with the assumption that agents are identical allows me to reduce the roots of the principal's problem and to characterize its analytical solution.<sup>24</sup> Henceforth I refer to the case in which agents are selfish, i.e.,  $\gamma = 0$ , as the standard case. The belief dependent utility function of agent  $i$  is given by:

$$U_i(e_i, \beta_i(e_j), \chi_i(e_i)) = \pi_i(e_i, \beta_i(e_j)) + \gamma r_i(k_{ij}(e_i, \beta_i(e_j)), \lambda_{iji}(\beta_i(e_j), \chi_i(e_i)))$$

### 2.3. The Principal's Problem

Let  $w_{x_i x_j}^i$  be the salary paid to agent  $i$  when  $x_i$  and  $x_j$  realize. To determine the optimal incentive scheme the principal minimizes the expected wage bill:

$$q^2 \sum_{i \in \{a, b\}} w_{hh}^i + (1 - q)q \sum_{i \in \{a, b\}} (w_{hl}^i + w_{lh}^i) + (1 - q)^2 \sum_{i \in \{a, b\}} w_{ll}^i \quad (\text{EW})$$

subject to the following *incentive compatibility constraints*

$$U_i(1, 1, 1) \geq U_i(0, 1, 1) \quad (\text{IC}_i)$$

<sup>23</sup>This implies that agents are able to compare the material and the psychological utilities in monetary terms and that the parameter  $\gamma$  can be interpreted as the shadow price of the psychological utility in terms of the material utility.

<sup>24</sup>Note that agents' utilities represent the constraint of the principal's minimization problem. Given this formulation, in equilibrium the squared roots of two identical arguments multiply each other.



for all  $i \in \{a, b\}$ . These constraints ensure that the optimal incentive scheme leads to a sequential reciprocity equilibrium in which the outcome  $\{e_a = 1, e_b = 1\}$  is achieved.

As a result of the psychological part of the agent's utility functions in constraints (ICa) and (ICb), the principal's problem exhibits more than one solution candidate.<sup>25</sup> In particular, each of these candidates corresponds to a different kind of reciprocal exchange that can take place among employees and to a different kind of incentives.<sup>26</sup>

#### 2.4. Benchmark: The standard case

To focus on the role of reciprocity, I first recall the optimal contracting solution when agents are selfish. In the standard case the principal minimizes (EW) subject to the agents' incentive compatibility constraints:

$$q^2 u(w_{hh}^i) + q(1-q)(u(w_{hl}^i) + u(w_{lh}^i)) + (1-q)^2 u(w_{ll}^i) - \phi \geq (1-q)q(u(w_{hh}^i) + u(w_{ll}^i)) + (1-q)^2 u(w_{hl}^i) + q^2(u(w_{lh}^i))$$

for all  $i \in \{a, b\}$ .

If agents are risk neutral they do not mind about the level of risk they are asked to bear, provided that the expected salary compensates their effort costs. As a result, the principal is indifferent between the following individual (1), team (2) and tournament (3) incentive schemes:<sup>27</sup>

$$w_{hh}^i = w_{hl}^i = \frac{\phi}{(2q-1)} > w_{lh}^i = w_{ll}^i = 0 \quad (1)$$

$$w_{hh}^i = \frac{\phi}{(2q-1)q} > w_{hl}^i = w_{lh}^i = w_{ll}^i = 0 \quad (2)$$

$$w_{hl}^i = \frac{\phi}{(2q-1)(1-q)} > w_{hh}^i = w_{lh}^i = w_{ll}^i = 0 \quad (3)$$

When agents are risk averse, there is no scope for relative performance evaluation to filter out common shocks. Any kind of interdependent compensation policy would solely increase the risk borne by an agent (Holmström 1982, Mookherjee 1984). Therefore, the optimal incentive scheme is the following individual scheme:

$$w_{hh}^i = w_{hl}^i = \left(\frac{\phi}{2q-1}\right)^{\frac{1}{1-\rho}} > 0 = w_{lh}^i = w_{ll}^i. \quad (4)$$

### 3. Optimal incentive scheme

With this benchmark at hand I can focus on the optimal contracting structure the principal provides to reciprocal agents, i.e.,  $\gamma \in (0, 1]$ . To isolate effects I first assume the agents to be risk neutral (Section 3.1) and then I enrich the model by introducing risk aversion (Section 3.2).

<sup>25</sup>This is because the principal's minimization problem is non-concave. To see why, consider that the term  $\lambda_{iji} * k_{ij}$  in the psychological payoff function implies that in the constraints (ICa) and (ICb) the choice variables of the principal's problem (the salaries paid to  $i$  and  $j$ ) multiply each other. As a result, the admissible set forms a convex hull only if a sign restriction to the type of kindness exchanged by the two agents is made.

<sup>26</sup>It has to be noted that a given incentive scheme may sustain a multiplicity of reciprocity equilibria. For instance, for each level of risk aversion and for  $\gamma > \left(\frac{2q-1}{q}\right)^{\frac{1}{1-\rho}}$ , the best relative incentive scheme sustains two equilibria: the desired one in which both agents work hard and one in which both agents shirk. In the latter equilibrium, agent  $i$  believes that agent  $j$  will shirk and, as a result,  $i$  shirks and minimizes the likelihood to let  $j$  behind. Similarly, the best joint incentive scheme sustains an equilibrium in which agent  $i$  shirks to reciprocate the nasty behavior he expects from  $j$ , provided that the inherent reciprocity concern  $\gamma$  is big enough. The idea is that when  $\gamma$  is low, agents are mainly interested in the monetary compensation they achieve through effort. The principal can substitute only a small share of the monetary compensation by inducing a horizontal gift exchange. As a result, the only equilibrium sustained by an incentive scheme is one in which agents give up a relatively small part of their monetary compensation so as to behave reciprocally, given their beliefs. Conversely, when  $\gamma$  is big returns from behaving reciprocally are so high that an incentive scheme may have unwanted consequences and an agent may be willing to shirk and give up a substantial part of his monetary compensation so as to reciprocate the behavior he expects from his coworker. In what follows I make an arbitrary restriction and I focus only on those equilibria which sustains  $(e_a = e_b = 1)$ .

<sup>27</sup>It can be easily seen that any of the incentive schemes here below implies an expected wage bill of  $\frac{2q\phi}{(2q-1)}$ . For further details, see Bolton and Dewatripont (2005).

### 3.1. Risk-neutral and reciprocal agents

When agents are reciprocal and risk-neutral, the principal is no longer indifferent between alternative kinds of compensation policies. A tournament-type incentive scheme turns out to be the most valuable option.

To get an intuition, consider what happens when the standard incentive schemes (1), (2) and (3) are provided to reciprocal, risk-neutral agents. First notice that all those incentive schemes implement hard working from both agents also when they are reciprocal. Indeed, the material payoff an agent receives already suffices to satisfy his (IC) constraint. Any additional psychological payoff immediately translates into a psychological rent.

**Lemma 1.** *The standard incentive schemes (1), (2) and (3) induce the outcome  $\{e_a = 1, e_b = 1\}$  as an equilibrium of the subgame starting at stage 2.*

**Proof.** This and all following proofs are given in the Appendix. □

While it is clear that incentive scheme (1) does not entail any additional psychological payoff - an agent's salary being independent of the other agent's indicator - incentive schemes (2) and (3) do. In particular, psychological payoffs arise as a result of a mutual exchange of favor(s) under the team-incentive scheme (2) and as a mutual exchange of disfavor(s) under the tournament incentive scheme (3).

The higher the psychological rent entailed by an incentive scheme, the easier it is for the principal to replace explicit monetary with implicit psychological incentives. Lemma 2 shows that agents receive the highest psychological rent when the tournament scheme (3) is provided.

**Lemma 2.** *When the standard incentive schemes (1), (2) and (3) are provided, the equilibrium psychological payoff agent  $i$  expects to receive at stage 2 is:*

$$r_i(k_{ij}, \lambda_{iji}) = \begin{cases} 0 & \text{under scheme (1);} \\ \frac{\gamma\phi}{2} & \text{under scheme (2)} \\ \frac{\gamma q\phi}{2(1-q)} & \text{under scheme (3).} \end{cases}$$

To see why scheme (3) provides greater psychological incentives than scheme (2), recall that the bonus paid to an agent under the tournament scheme (3) is larger than the bonus paid to an agent under the team-incentive scheme (2). This implies that the (negative) expected externality inflicted by  $i$ 's hard working under scheme (3) is greater than the (positive) expected externality induced by  $i$ 's hard working under the low-powered scheme (2). It follows that an agent is more keen to reciprocate the negative externality of his coworker's high effort when scheme (3) is paid than to reciprocate the smaller positive externality of his coworker's choice when scheme (2) is paid. Compared to schemes (1) and (2), the relative-incentive scheme (3) maximizes the distance between the kindness  $i$  expects to give to  $j$  by choosing high effort and what  $i$  could have expected to give to  $j$  by shirking. Therefore, when designing the optimal scheme for reciprocal agents the principal can take greater advantage of the presence of reciprocity concerns by designing relative-incentive scheme, characterized by a substantial bonus paid in a relatively unlikely state of the world, rather than a joint-incentive scheme, characterized by a lower bonus paid in the most likely state of the world.

This intuition is confirmed by solving the principal's problem. In a first stage, I can characterize the three optimal candidate schemes, which are listed in the lemma below:

**Lemma 3.** *When agents are reciprocal and risk neutral with respect to wealth variations, the principal problem has three local solution candidates:*

1. *The standard individual performance scheme (1);*
2. *a team incentive scheme in which:*

$$w_{ihh} = \frac{\phi}{q(2q-1)(1+\gamma)} > w_{ihl} = w_{ilh} = w_{ill} = 0; \quad (5)$$

3. and a tournament incentive scheme in which:

$$w_{ihl} = \frac{\phi}{(2q-1)(1-q+q\gamma)} > w_{ihh} = w_{ilh} = w_{ill} = 0; \quad (6)$$

for all  $i \in \{a, b\}$ .

Comparing the expected wage bills entailed by each solution candidate, I am able to find the best incentive option for the principal:

**Proposition 1.** *When agents are reciprocal and risk neutral with respect to wealth variations, the principal optimally provides incentives through the tournament scheme (6).*

Consistently with Lemma 2, Proposition 1 highlights that the principal optimally implements the relative-incentive scheme (6). This allows her to maximize the magnitude of the expected externality of  $i$ 's effort choice on  $j$ 's material payoff, thereby efficiently replacing monetary with psychological incentives.

Compared to the standard case, the principal is no longer indifferent between providing risk-neutral agents with individual, joint or relative incentive schemes. The optimal incentive scheme provided to reciprocal, risk-neutral agents entails a tournament that induces an exchange of unkind behaviors among and competition between coworkers. Each agent's effort is seen as a nasty behavior from his colleague, as it reduces the latter's chance to collect a bonus. Even though an unfriendly working environment is established, each agent enjoys a positive psychological payoff, as he manages to pay his coworker back in own coin.

### 3.2. Risk-averse agents

Section 3.1 shows that the optimal contracting option to motivate reciprocal, risk-neutral agents. The situation changes significantly when I consider reciprocal, risk-averse agents. On the one hand, interlinked compensation policies allow the principal to provide implicit incentives for hard working, thereby reducing the cost of motivating the workforce. On the other hand, they force each agent to bear some extra risk unrelated with his effort decision. This additional loss is not too relevant when agents are only moderately risk-averse. Therefore, in such scenario the principal continues to prefer a relative-incentive compensation policy, even though a pure tournament is not necessarily desirable. This means that the optimal incentive scheme is still based on relative incentives but it no longer looks like a pure tournament. In contrast, when agents are sufficiently risk-averse, a relative-incentive policy is inefficient as it would charge agents with too much risk. A good balance between the costs of providing agents with insurance and making the most of their reciprocal preferences is achieved implementing a less powered joint-incentive compensation policy.

To show this result I first characterize the three solution candidates and I then compare the expected cost they entail to the principal.

The first locally optimal incentive scheme corresponds to the optimal standard incentive scheme (4).

**Observation 1.** *Scheme (4) represents the best individual-incentive scheme that can be provided by the principal. In equilibrium, agents do not achieve any psychological payoff.*

The second solution candidate implements a joint-incentive policy and induces a mutual exchange of positive kindness among coworkers. However, it takes the structure of a pure-team incentive scheme only if agents are sufficiently motivated by reciprocity.

**Lemma 4.** *The best joint-incentive scheme is such that:*

- If  $\gamma \in (0, 1 - q/q)$ , then:

$$w_{hh}^i = \left( \frac{\phi}{(2q-1) \left( q(1+\gamma) + (1-q-\gamma q) \left( \frac{1-q-q\gamma}{(1+\gamma)(1-q)} \right)^{\frac{1-\rho}{\rho}} \right)} \right)^{\frac{1}{1-\rho}}$$

$$>$$

$$w_{hl}^i = w_{hh}^i * \left( \frac{1-q-q\gamma}{(1+\gamma)(1-q)} \right)^{\frac{1}{\rho}} > w_{lh}^i = w_{ll}^i = 0.$$

- if  $\gamma \in [1 - q/q, 1]$ , then:

$$w_{hh}^i = \left( \frac{\phi}{(2q-1)(1+\gamma)q} \right)^{\frac{1}{1-\rho}} > w_{hl}^i = w_{lh}^i = w_{ll}^i = 0.$$

If  $\gamma \rightarrow 0^+$ , providing agents with insurance is the dominating issue and the best mutual kindness incentive scheme converges to the individual incentive scheme (4): Agent  $i$  is allocated a bonus whenever his indicator is good, while the indicator of his peer has only a vanishing effects on his compensation. As  $\gamma$  increases, reciprocity gains importance and the principal increases the size of the gifts exchanged by the agents at the cost of providing them with lower insurance against risk of income variations: When the intensity of reciprocity is modest ( $\gamma \leq (1-q)/q$ ), each agent always receives a positive bonus when his own indicator turns out to be good but he is rewarded more when also his colleague performs well, i.e.  $w_{hh}^i > w_{hl}^i$ . The wedge between  $w_{hh}^i$  and  $w_{hl}^i$  increases as tastes for reciprocity become more intense; when reciprocity is the more relevant issue, namely when  $\gamma > (1-q)/q$ , the non-negativity constraint associated to  $w_{hl}^i$  bites and the principal provides pure team incentives.

The third solution candidate follows a somehow similar pattern, as shown by the following lemma:

**Lemma 5.** *The best relative-incentive scheme is such that:*

- If  $\gamma \in (0, 1)$ , then:

$$w_{hl}^i = \left( \frac{\phi}{(2q-1) \left( (1-q+\gamma q) + q(1-\gamma) \left( \frac{(1-\gamma)(1-q)}{1-q+q\gamma} \right)^{\frac{1-\rho}{\rho}} \right)} \right)^{\frac{1}{1-\rho}}$$

$$>$$

$$w_{hh}^i = w_{hl}^i * \left( \frac{(1-q)(1-\gamma)}{1-q+q\gamma} \right)^{\frac{1}{\rho}} > w_{lh}^i = w_{ll}^i = 0.$$

- if  $\gamma = 1$ , then:

$$w_{hl}^i = \left( \frac{\phi}{(2q-1)(1-q+q\gamma)} \right)^{\frac{1}{1-\rho}} > w_{hh}^i = w_{lh}^i = w_{ll}^i = 0.$$

Again, when agents are only slightly reciprocal ( $\gamma \rightarrow 0^+$ ), the optimal scheme converges to scheme (4). When  $\gamma$  takes on modest values, each agent achieves the highest bonus if he is the only one who performed well, a lower bonus when both indicators reveal good evidence, and nothing otherwise. When  $\gamma$  is very high, a pure tournament scheme is set up and only the agent who eventually proves to be the best receives a positive bonus.

Observation 2 immediately follows from Lemmas 4 and 5:

**Observation 2.** *At the limit of  $\gamma \rightarrow 0^+$ , the total expected cost entailed by the best joint and by the best relative-incentive schemes converges to the total expected cost entailed by the best individual scheme (4).*

Moreover, simple computations lead to the following observation:

**Observation 3.** *The expected value of the monetary rewards paid to the agents strictly decreases in  $\gamma$  under both the best joint and the best relative-incentive schemes.*

Observation 3 highlights that the more agents are motivated by reciprocity concerns, the more efficient is the psychological gift-exchange game designed by the principal.

Finally, from Lemmas 4 and 5 it is clear that the incentive structure of interdependent schemes becomes simpler as the agents' concern for reciprocity increases. For  $\gamma$  sufficiently high, both the best joint and the best relative-incentive schemes pay a positive bonus only in one state of nature.

**Observation 4.** *The best joint and the best relative-incentive schemes pay a positive bonus in two states of nature when the agents' inherent concern for reciprocity  $\gamma$  is low and in only one state of nature when  $\gamma$  is sufficiently high.*

To characterize the solution to the principal's problem, I first compare the individual scheme (4) with the other two solution candidates. The following lemma holds:

**Lemma 6.** *The individual incentive scheme (4) is strictly dominated by both the best joint-incentive scheme and the best relative-incentive scheme for any  $\rho \in [0, 1)$  and  $\gamma \in (0, 1]$ .*

Lemma 6 highlights that the standard scheme 4 is never adopted by the principal to stimulate effort provision from reciprocal agents. This result comes from Observations 2 and 3 and from the continuity in  $\gamma$  of the total expected cost entailed by the best joint and the best relative incentive schemes.

Comparing the two best interdependent-incentive schemes I show that the final choice of the principal is crucially affected by the parameter of relative-risk aversion,  $\rho$ . The following proposition:

**Proposition 2.** *For all  $\gamma \in (0, 1)$  and  $\rho \in (0, 1)$  there exists a threshold value  $\rho^* \in (0, 1)$  such that:*

- *if  $\rho < \rho^*$  the principal strictly prefers the best relative-incentive scheme;*
- *if  $\rho = \rho^*$  the principal is indifferent between the best relative-incentive and the best joint-incentive schemes;*
- *if  $\rho > \rho^*$  the principal strictly prefers the best joint-incentive scheme.*

In a nutshell, when agents are (almost) risk neutral, the cost of providing them with insurance is negligible. Since a reciprocal agent is interested in the size of the consequences he is going to inflict on his coworker, the principal makes the most of reciprocity concerns by designing a relative-incentive scheme that makes agents competing against each other. As risk aversion increases, it becomes costly for the principal to provide agents with insurance using the best relative-incentive scheme, as it pays the highest bonus in states ( $\{x_a = h, x_b = l\}$  and  $\{x_a = l, x_b = h\}$ ) which are not the most likely when both agents exert high effort. When risk aversion overcomes the threshold value  $\rho^*$ , the best joint-incentive scheme allows to exploit the beneficial effects of reciprocity better than relative-incentive schemes. Since in equilibrium the former scheme pays the highest bonus in the most likely state,  $\{x_i = h, x_j = h\}$ , it efficiently combines monetary and psychological incentives simultaneously addressing the agents' demand for insurance.

### 3.3. Comparative statics

Proposition 2 shows that the optimal scheme for reciprocal and risk-averse agents implements some kind of interdependent compensation policy. The choice between a relative or a joint incentive scheme crucially depends on the workers' tastes towards risks. Figures 1 and 2 offer a graphical intuition of how the other parameters of the model affect the principal's choice. I represent the optimal incentive scheme as a function of combinations of the inherent concern for reciprocity,  $\gamma$ , (on the horizontal axis) and the parameter of relative risk aversion,  $\rho$ , (on the vertical axis) when the cost of effort  $\phi$  equals 2. In Figure 1

I consider a scenario in which  $x_i$  is a poor signal of  $i$ 's effort ( $q = 0.65$ ) whereas in Figure 2 I consider a scenario in which  $x_i$  is a good signal of  $i$ 's effort ( $q = 0.85$ ).

Figure 1 shows that the principal optimally chooses the best joint-incentive scheme for values of risk aversion higher than  $\rho^*$  (represented by the continuous line), as stated by Proposition 2. When  $\gamma > (1 - q)/q = 0.5384$ , the non-negativity constraint associated with  $w_{hh}^i$  in the best joint-incentive scheme bites. The principal finds it optimal to offer a pure-team scheme in which agents receive a bonus only if both efforts are signalled to be high.

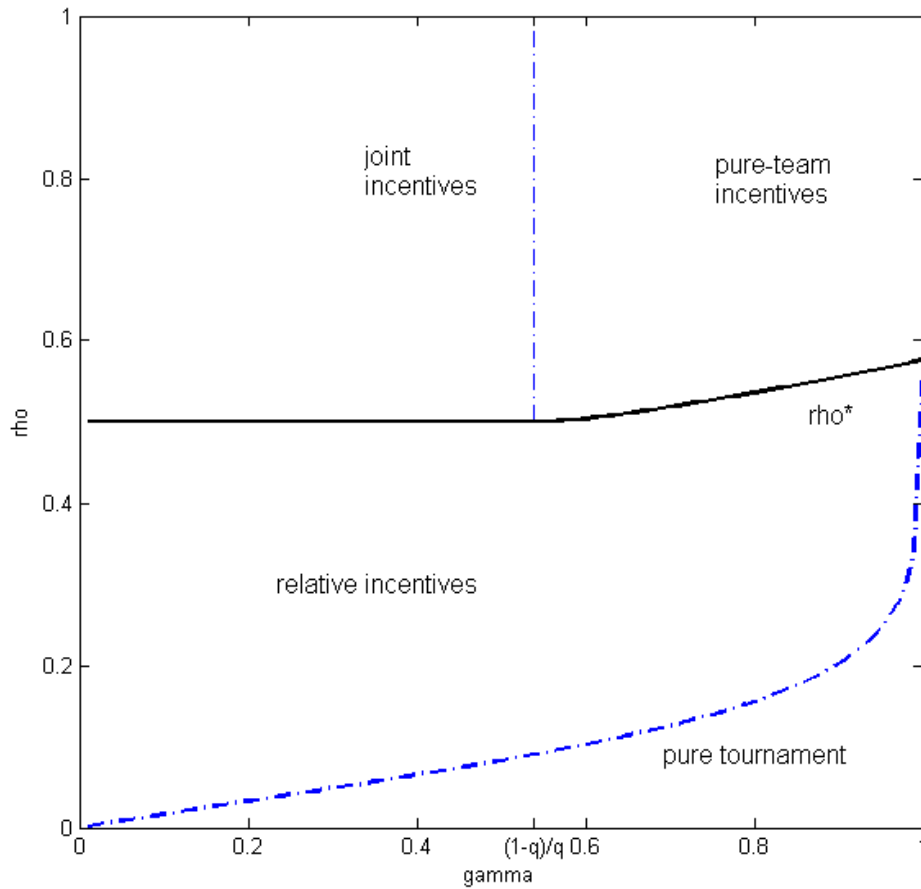


Figure 1: Optimal Incentive Scheme.  $\phi = 2, q = 0.65$

For value of  $\rho < \rho^*$ , the principal strictly prefers the best relative-incentive scheme. This takes the form of a pure tournament when  $\rho$  approaches 0 and when  $\gamma$  approaches 1.<sup>28</sup>

In Figures 1 and 2 it can also be seen that  $\rho^*$  (weakly) increases in  $\gamma$ . The best joint-incentive scheme has a comparative advantage over the best relative-incentive scheme in providing agents with insurance. When the agents' reciprocity concern,  $\gamma$ , increases, the material rent necessary to induce positive effort

<sup>28</sup>Lemma 5 states that the best relative-incentive scheme takes the form of a pure tournament if  $\gamma = 1$ , as for such value the non-negativity constraint associated with  $w_{hh}^i$  bites. This holds true also at the limit of  $\gamma$  approaching 1. In Figures 1 and 2 I consider as a pure tournament also those cases in which  $w_{hh}^i$  takes a value smaller than 0.000001.

provision decreases (Observation 3). This is because the psychological impact of a "gift" with a given size increases in  $\gamma$ . The principal can thus reduce the power of the explicit monetary incentives of both the best joint and relative incentive schemes. This in turn reduces the cost of providing the agents with insurance, making the comparative advantage of the best joint-incentive scheme less relevant. Therefore,  $\rho^*$  (weakly) decreases in  $\gamma$ .

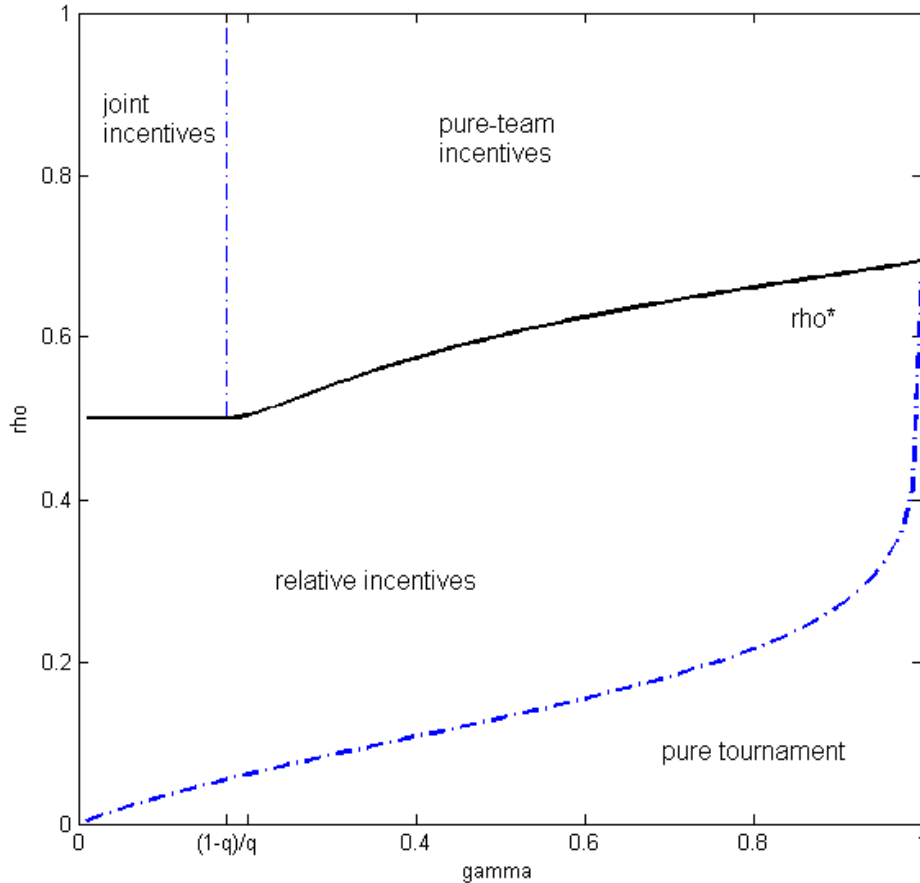


Figure 2: Optimal Incentive Scheme.  $\phi = 2, q = 0.85$

Similarly, the best relative-incentive scheme becomes comparably more attractive for higher values of the signal precision  $q$ . Indeed, a greater  $q$  decreases the power of explicit incentives as it decreases the risk borne by each agent  $i$  when choosing  $e_i = 1$ . Again this reduces the comparative advantage of the best joint-incentive scheme. This can be directly seen comparing Figures 1 and 2. In the latter figure the continuous line representing  $\rho^*$  always take (weakly) greater values than in the former figure. As a result, the area in which the best relative-incentive scheme dominates is bigger than in the former figure.

In general, the comparative advantage of the best-joint incentive scheme decreases when (i) agents become more reciprocal and (ii) the precision of individual performance indicator improves.

In contrast, the choice of the optimal compensation policy does not depend on the agents' effort cost, which solely affects the distance between the best relative-incentive and joint-incentive schemes for given values of  $q$  and  $\gamma$ . This result is shown in the Appendix. These results are summarized in the following

observation:

**Observation 5.** *The threshold value  $\rho^*$ :*

- (weakly) increases in the agents' inherent concern for reciprocity,  $\gamma$ ;
- (weakly) increases in the precision of the performance indicators,  $q$ ;
- does not vary with the cost of effort,  $\phi$ .

#### 4. High effort costs

The previous section shows that the principal reduces the cost of achieving high effort from both workers designing a psychological gift-exchange game, thereby replacing monetary with psychological incentives. It turns out that a reciprocal agent expects to earn less money than a standard agent. Crucial for this result is the assumption that hard working is so valuable for the principal that she always prefers to induce high effort provision no matter the costs. In this section we analyze what happens if we relax this assumption, i.e., we assume that the expected returns to hard working are not overly high. We show that principal can find it profitable to induce effort provision from reciprocal agents in instances in which it is prohibitively costly with standard agents.

Let the performance indicator  $x_i \in \{h, l\}$  be the (gross) profit generated by agent  $i$ 's activity. Without loss of generality, I normalize  $l$  to zero and I assume that only  $h$  takes on a positive value. Moreover, let me assume that  $\phi > h^{(1-\rho)}(2q-1)$ . When this assumption holds, it is never optimal for the principal to induce high-effort provision from standard agents, effort being too costly with respect to the principal's expected gains. No matter the risk-aversion of the agents, the principal finds it profitable to pay the following flat wage scheme:

$$w_{x_i x_j}^i = 0 \quad (7)$$

for all  $x_i, x_j$ . In equilibrium, the scheme (7) induces low-effort provision from both agents.

This is not necessarily the case when agents are reciprocal. As a matter of example, let me focus on a situation in which  $\rho > \rho^*$  for some  $\gamma$  and  $q$  such that  $\gamma > (1-q)/q$ . From Proposition 2 I know that the scheme that induces  $\{e_a = 1, e_b = 1\}$  at the smallest cost is the best joint-incentive scheme highlighted by Lemma 6. For these parameter values, it takes the form of a pure team-incentive scheme. Such incentive scheme proves superior to the flat incentive scheme (7) whenever:

$$h^{(1-\rho)}(2q-1) < \phi < h^{(1-\rho)}(2q-1)q(1+\gamma) \quad (8)$$

Therefore, the psychological gift-exchange that takes place between reciprocal agents makes the cost of inducing high-effort provision sustainable for the principal also in some instances in which it is overly costly for standard agents. When condition (8) is satisfied, reciprocal agents are asked to exert positive effort and expects to receive a positive compensation, whereas standard agents are asked to exert no effort and expects to receive the minimum salary.

In this example I considered a specific situation in which agents are sufficiently risk averse. This result can be easily generalized to other situation for a non-knife-edge subset of the parameter space.<sup>29</sup> The following proposition holds:

**Proposition 3.** *For each  $\rho \in (0, 1)$ , there exists a full dimension subset of the parameter space in which reciprocal agents expect to earn more and are asked to provide more effort than self-interested agents.*

For values of the effort costs  $\phi$  slightly greater than  $h^{(1-\rho)}$ , a slightly positive reciprocity concerns parameter  $\gamma$  suffices to make positive effort provision valuable for the principal.

Proposition 3 suggests that reciprocal agents can be more productive than standard agents, consistently with the empirical analysis of Dohmen et al. (2009)<sup>30</sup>.

<sup>29</sup>This also holds when agents are only slightly risk averse. As it shall be clear from Proposition 2, in this case the principal prefers a relative-incentive scheme.

<sup>30</sup>See Section 6 for a discussion.



## 5. Group Production

The previous sections consider a somewhat special case in which workers' production technologies are independent, i.e., group production exhibits no technological synergies. However, workgroups and joint-production have become increasingly popular in the recent years. Not only they are widely adapted by firms, but many success stories are associated with teams.<sup>31</sup>

In this section I extend the analysis to determine the optimal compensation scheme under the assumption that agents are engaged in joint production. Let me consider a scenario in which the principal asks  $a$  and  $b$  to jointly realize a project. Agents work simultaneously. Each agent is charged with a different part of the production process and exerts an effort  $e_i \in \{0, 1\}$ . Again, effort  $e_i$  costs each agent  $\phi * e_i$ . As in the previous sections, the principal observes neither  $e_a$  nor  $e_b$ . The realized value of the project is verifiable and can be contracted upon.

Once realized, the project can be either a success or a failure, i.e., it is worth  $y \in \{\bar{y}, \underline{y}\}$  to the principal, with  $\bar{y} > \underline{y} = 0$ . The probability that the project is a success crucially depends on the agents' efforts. If both agents exerts effort equal to 1, then the project has a probability  $\bar{p} < 1$  of being successful; if any of the two agents shirks and the other works hard, the project is successful with probability  $\hat{p} < \bar{p}$ ; if both agents shirk, the project is successful with probability  $\underline{p} \in (0, \hat{p})$ .

In what follows I characterize the optimal incentive option in two different information settings: First I show that the principal always induces a positive reciprocal exchange when only the global output can be contracted upon (Section 5.1). Second I show that the usual interplay between reciprocity and risk attitudes holds when additional signals about the agents' individual performance are available (Section 5.2).

### 5.1. No individual performance indicators available

When only the value of the project can be contracted upon, the optimal incentive scheme rewards the agent with a bonus only if the value of the project equals  $\bar{y}$ , as highlighted by the following lemma:

**Lemma 7.** *When only the global output can be contracted upon, the optimal incentive scheme entails the following payments:*

$$\bar{w} = \left( \frac{\phi}{(1 + \gamma)(\bar{p} - \hat{p})} \right)^{\frac{1}{1-\rho}}; \quad \underline{w} = 0$$

where  $\bar{w}$  and  $\underline{w}$  are the salary paid to the agents when the project is worth  $\bar{y}$  and  $\underline{y}$  respectively.

Agents being awarded a positive bonus only if the project is successful. As a result, the optimal incentive scheme always entails a positive reciprocal exchange: An agent's effort increases the material payoff of his teammate and it is therefore perceived as a kind behavior.

### 5.2. Individual performance indicators

Now suppose that individual performance indicators  $x_i \in \{h, l\}$ ,  $i \in \{a, b\}$  are available. To ease comparisons with respect to the previous sections, I assume that each  $x_i$  takes value  $h$  with probability  $q$  when  $e_i = 1$  and with probability  $(1 - q)$  when  $e_i = 0$ . Moreover, I assume that each individual indicator does not represent a sufficient statistic for  $y$  with respect to  $e_i$  while it does represents a sufficient statistics for  $x_j$  with respect to  $e_i$ . When agents are self-interested, these assumptions imply that the principal finds it optimal to make each agent's salary contingent on the value of the project and on his own performance indicators. The optimal incentive scheme paid to agent  $i$  only depends on  $y$  and  $x_i$  but not on agent  $j$ 's performance as this does not gather any additional information about  $i$ 's effort.

When agents are reciprocal, an independent incentive scheme is no longer optimal. Along the same line of the previous sections, the principal can replace monetary with psychological incentives by providing interdependent compensation policies. Similar to Section 3, it is possible to characterize a best joint and a

---

<sup>31</sup>See Che and Yoo (2001) for a discussion.

best-relative incentive scheme.<sup>32</sup> The final choice between joint or relative incentives depends again on the agents' risk aversion.

**Proposition 4.** *There exists a threshold value of  $q$ ,  $\bar{q}(\bar{p}, \hat{p}) \in (1/2, 1)$ , such that:*

- *if  $q \leq \bar{q}(\bar{p}, \hat{p})$ , then the optimal scheme implements a joint-incentive scheme that induces a positive reciprocal exchange irrespective of  $\rho$ ;*
- *$\partial \bar{q} / \partial \bar{p} \geq 0$  and  $\partial \bar{q} / \partial \hat{p} \leq 0$ ;*
- *if  $q > \bar{q}(\bar{p}, \hat{p})$ , then there exists a threshold value  $\tilde{\rho} \in (0, 1)$  such that the optimal scheme implements a relative-incentive scheme that induces a negative reciprocal exchange if  $\rho \leq \tilde{\rho}$  and a joint-incentive scheme otherwise.*

Compared to the individual production case (Proposition 2), Proposition 4 highlights that positive reciprocity is endogenously more efficient when agents are engaged in joint production. In particular, the principal finds it profitable to offer some joint-compensation package even when agents are risk-neutral, provided that the value of production is comparatively informative with respect to the individual performance indicators. Therefore, the greater the synergies in the production process, the more important it is to encourage friendship at the workplace. However, the principal can find it profitable to design a relative-incentive scheme that induces competition and negative reciprocity among coworkers. Indeed, when the individual performance indicators are comparatively more informative than the global output, the interplay highlighted in the previous sections still arises: Agents are offered some kind of relative-incentive scheme when they are not too risk averse and some kind of joint-incentive scheme when they are sufficiently risk-averse.

A similar result can be achieved in a setting in which agents can undertake some costly action to sabotage their coworkers' performance indicators. In such an environment inducing competition among coworkers may lead to undesired consequences as earlier noted by Lazear (1989). As a result, implementing the best-relative incentive scheme might be unprofitable also when agents are risk-neutral.

Proposition 4 also stresses that in the setting considered in this paper, individual performance indicators are collected not only to better assess each agent's individual contribution in order to provide better insurance. Indeed, they also make the reciprocal exchange of favor more efficient as agents are intrinsically interested in the material consequences of their actions to the other agent. Therefore, since  $x_i$  can be interpreted as an informative signal of  $i$ 's kindness towards  $j$ , the principal finds it profitable to make the salary of agent  $i$  contingent on  $x_i$ ,  $x_j$  and  $y$  even though  $x_i$  is sufficient for  $e_i$  with respect to  $x_j$ .

Contributions in the management literature already highlighted how the choice between a competitive and a joint incentive package depends on the synergies of the production technology. They argue that rewards schemes such as group incentives increase the group morale and they are preferable when joint-production involves a lot of synergies among the workers' activities. I complement this argument showing that the interplay between reciprocity and risk-aversion moderates this relationship.

## 6. Discussion

This paper highlights that:

1. the principal always find it profitable to implement interdependent compensation schemes and that
2. the final choice between a relative and a joint-incentive scheme depends on the interplay between the agents' other-regarding preferences (ORP) and their attitudes towards risks. When agents are risk-neutral, the principal prefers to set up a relative incentive scheme, as it better replace monetary with psychological incentives. When agents are risk-averse, the optimal incentive design implies solving

---

<sup>32</sup>These schemes exhibit similar features to the ones described by Lemmas 4 and 5. In particular, the expected salaries paid by the principal decrease in  $\gamma$  and the incentive structure is simpler when agents concerns for reciprocity are larger.

a threefold trade off involving risk, monetary, and psychological preferences: For risk aversion high enough, the principal sets up a joint-incentive scheme as it provides the best balance among those three forces.

As previously highlighted by other contributions (Itoh, 2004, Englmaier and Wambach, 2010, Bartling and von Siemens, 2010, Bartling, 2011) the first is a general result that in principle holds whenever agents are endowed with some kind of other-regarding preferences (e.g., status seeking, inequity aversion, envy). Conversely, the interplay between agents' reciprocity concerns and risk aversion highlighted in this paper only arises when agents exhibit preferences for reciprocity.<sup>33</sup>

My model thus generates a clear prediction: The use of joint-incentive policies should be empirically more prevalent among employees with higher degrees of risk aversion, especially in job settings in which individual performance indicators are poorly correlated (Proposition 2).<sup>34</sup> A further step could be to test this hypothesis using a data set that would allow to control for the precision of the performance indicators and for the intensity of the preferences for reciprocity, which can moderate the relationship (Observation 5), as well as for the degree of interdependency of the production function (Proposition 4).

As argued by Oyer and Schaefer (2011) and by a recent report by the VisionLink advisory group, the use of joint incentives is comparatively more frequent at the bottom tiers of a hierarchical structure than at the intermediate and top tiers.<sup>35</sup> This seems consistent with my prediction, risk aversion being typically higher among blue collars and clerks than among managers and CEOs.<sup>36</sup> Further empirical investigation can help shading light on this correlation.

Lemmas 4 and 5 and Proposition 2 also suggest that the shape of the optimal incentive scheme depends on the intensity of the agents' inherent concern for reciprocity. If agents are only mildly motivated by reciprocity, providing agents with insurance is the most relevant issue for the principal. When this is the case agents are always paid when their own performance indicator turns out to be good. The performance indicator of the coworker is solely used to adjust the final bonus: When agents are not too (respectively, sufficiently) risk averse, an additional bonus is paid when the coworker's indicator turns out to be bad (resp., good). If agents attach a high value to reciprocity, the principal implements a *pure-tournament* incentive scheme in which only the "winning" agent is awarded a bonus when agents are poorly risk-averse, and a *pure-team* incentive scheme in which rewards are allocated only if both agents perform well when agents' risk-aversion is high enough. Related to that, Observation 3 highlights that the power of material incentives decreases with the intensity of agents' reciprocity concerns: When the agents' psychological concerns increase, lower monetary bonuses suffice to overcome the cost of effort.<sup>37</sup> Moreover, Observation 4 remarks that the incentive structure of the optimal incentive scheme is simpler when agents' reciprocity concerns are strong enough. Therefore, simpler and less-powered incentive structures should be more frequently observed in context in which the workforce exhibit stronger reciprocal attitudes. This paper might thus help to understand why the power of monetary incentives is often lower than predicted by standard moral hazard models, complementing the standard multitasking argument of Holmström and Milgrom (1991).

This paper also predicts that reciprocal agents are more productive than standard ones (Observation 3 and Proposition 3). This is consistent with the empirical analysis of Dohmen et al. (2009). Analyzing data

---

<sup>33</sup>Indeed, such interplay emerges because of (i) the fact that the best-relative incentive scheme maximizes the distance between the kindness  $i$  expects to give to  $j$  and the reference point and (ii) the fact that agents bear more risk under the best-relative incentive scheme than under the best-joint incentive scheme.

<sup>34</sup>The interplay between risk aversion and reciprocity concerns holds even when individual performance indicators are correlated, provided that correlation is not too strong with respect to reciprocity concerns.

<sup>35</sup>See <http://www.vladvisors.com/compensation-knowledge-center/articles/a-38-selecting-the-right-performance-measures-for-your-incentive-plan>

<sup>36</sup>See e.g. Hartog et al. (2002).

<sup>37</sup>Note that a similar result can be achieved under the assumption that agents are inequity averse, as shown by Bartling (2011) and Englmaier and Wambach (2010). In models with inequity averse agents, there are two major effects: First, inequity averse agents must be compensated for ex-ante wage inequality whenever their participation constraints bind; second, agents modify their effort choices in a bid to reduce wage inequality. However, experimental and empirical evidence suggest that wage inequality *per se* has no impact on both agents' participation and effort decisions (Charness and Kuhn 2007, Maximiano et al. 2007, Abeler et al. 2010, Bartling and von Siemens 2011).

from the GSOEP survey,<sup>38</sup> the authors find that workers with stronger preferences for positive reciprocity significantly exert more effort and earn more money than non-reciprocal workers.<sup>39</sup> The authors argue that this correlation is consistent with the idea of a gift-exchange between the principal and the agent, in which a reciprocal worker reacts exerting more effort in exchange for a generous compensation.<sup>40</sup> The theoretical analysis by Englmaier and Leider (2012) provides a framework consistent with this idea. The authors consider a principal-single agent model in which offering a reciprocal agent with a generous compensation is optimal when risk-aversion and reciprocity concerns are sufficiently high. Proposition 3 offers a complementary explanation to this evidence in which horizontal rather than vertical reciprocity is considered and that holds irrespective of the level of risk-aversion.

It is worth recalling that the analysis displayed in the paper relies on two crucial assumptions: (i) agents equally care about positive and negative reciprocal exchanges; (ii) agents are identical.

With respect to point (i), evidence of asymmetric preferences for the two sides of reciprocity have been found in both psychology and experimental economics (Perugini et al. 2003, Fehr and Schmidt 2003).<sup>41</sup> This feature can be easily incorporated in my model, as shown in Appendix B. The interplay between risk and reciprocal attitudes still holds. I show that when preferences for different kinds of reciprocity are extremely skewed, then one kind of incentive scheme dominates the other no matter the level of risk-aversion.

When preferences for reciprocity are extremely polarized, a badly designed incentive scheme may induce no effort provision as a unique equilibrium of the game. This idea is consistent with the field experiment of Bandiera et al. (2005). The authors present a field experiment in which they alter the incentive structure for field workers in a UK farm. They find that compensation policies based on relative-incentive perform worse than incentives based on individual measures of performance. They argue that this was due to workers partially internalizing the negative externality their effort imposes on others under relative incentives, consistently with agents being more concerned about positive than negative reciprocity.<sup>42</sup> Therefore, asymmetric preferences for positive and negative reciprocity should carefully be taken into account by any empirical investigation of my model.

Moreover, empirical research in management and economics has shown that there are significant differences in the compensation packages offered to CEOs, managers and employees across countries (Canyon and Murphy 2000, Tosi and Greckhamer 2004, Bryan et al. 2014). The authors argued that such discrepancies might be explained by divergencies in the cultural traits of the different countries.<sup>43</sup> I suggest that those discrepancies might also be explained by differences in the distributions of positive and negative reciprocators among countries.

With respect to point (ii), this assumption is crucial to characterize a close form solution of the model. However, numerical simulations show that the main results still go through when this assumption is re-

---

<sup>38</sup>GSOEP is a longitudinal panel dataset of the population in Germany. It is a household based study which started in 1984 and which reinterviews adult household members annually.

<sup>39</sup>The authors find a non-significant correlation between effort and wage levels for negative-reciprocal types.

<sup>40</sup>The idea that positive reciprocators repay generous wages with high effort even when there is no way to enforce contracts has been demonstrated also experimentally (See e.g., Fehr et al. 1993, Berg et al. 1995, Brown et al. 2004, Falk 2007).

<sup>41</sup>Perugini et al. (2003) developed a six-items questionnaire to measure preferences for positive and negative reciprocity. This questionnaire is widely used by firms to measure employees reciprocal attitudes and improve their Management Control Systems. Empirical tests demonstrated that positive and negative reciprocity are only slightly correlated and that they amount to different psychological constructs (Dohmen et al. 2008).

<sup>42</sup>The authors also show that workers internalize the externality to a greater extent when a large share of their coworkers are their close friends.

<sup>43</sup>Tosi and Greckhamer (2004) analyse data about HR practices from the Towers Perrin Worldwide Total Remuneration Reports for the years 1997, 1998, 1999, 2000, and 2001. These reports provide information from twenty-three countries on the average compensation levels and the structure of compensation for four different positions representing four organizational levels (CEO, human resources director, accountant, and manufacturing employees) in industrial companies with approximate annual sales of \$250-\$500 million. Canyon and Murphy (2000) highlight significant differences in compensation packages provided to CEOs and top managers between US and UK. Both studies conclude that differences in cultural dimensions can contribute understanding cross-national variations in compensation policies. Bryan et al. (2014) examine CEO compensation in 43 countries over the 1996 to 2009 period. The authors merge elements of the cross-cultural psychology and contracting literatures to identify cultural factors that theory predicts should affect agency conflicts. The empirical analysis confirms that cultural factors are strongly significant after controlling for the previously-identified determinants of CEO compensation structure relating to legal environment and firm-specific characteristics.

laxed. In this case the principal's choice between a joint and a relative incentive scheme is driven by some statistics of the joint distribution of reciprocity and risk attitudes of the workforce. For instance, when agents are heterogeneous on the return of their effort on profits (e.g., when one agent is the *star* of the firm), then the principal should tend to pick the kind of interdependent policy "favored" by the star. E.g., if the star is very risk averse, then the principal is probably going to set up a joint-incentive scheme that induces positive reciprocity.

This raises the question on how agents with different risk and reciprocal attitudes sort into different firms characterized by heterogeneous corporate culture<sup>44</sup> or how the ORP of the workforce affect organizational strategies other than the choice of the incentive scheme.<sup>45</sup> What could also be interesting is to investigate whether and to what extent employees' psychological preferences are affected by the incentive policy of their firm in a dynamic context. As also pointed out by Milgrom and Roberts (1994), this issue has received limited attention by economists,<sup>46</sup> despite the great amount of energies and initiatives devoted by Human Resource managers at shaping employees' preferences. I leave both these issues for future research.

Finally, in the model agents only exhibit preferences for reciprocity towards each other (horizontal reciprocity), whereas they do not exhibit any kind of other-regarding preferences with respect to the principal (vertical reciprocity). On the one hand, this assumption might be more plausible in situations in which agents only personally know their colleagues and maybe the members of the middle management, whereas they do not have any direct interaction with their company's CEO or top management. On the other hand, this assumption allows me to isolate the effect of horizontal reciprocity on the optimal compensation policy of the firm. Englmaier and Leider (2012) consider a principal-single agent model in which the agent exhibit vertical reciprocity concerns. They show that when agents are sufficiently risk averse, the principal solves the moral hazard problem by paying flat but generous incentive schemes. A generous incentive scheme induces the agent to reciprocate the principal's gift by exerting high effort and allows the principal to reduce the cost providing insurance to the agent.<sup>47</sup> In a setting in which agents care about both vertical and horizontal reciprocity, the robustness of the results presented in the previous sections would crucially depend on the relative importance of those two concerns and on their substitutability in the agents' utility functions.

## 7. Conclusion

Research in management, organizational behavior and experimental economics has stressed the key role played by reciprocity within firms and organizations. Only recently, the burgeoning literature in behavioral contract theory started incorporating preferences for reciprocity into otherwise standard contract theory models.

In this paper I solve for the optimal incentive scheme in a basic principal-multi agents setting under the assumption that agents are reciprocal towards each others. I show that the principal can limit the cost of motivating non-trivial effort provision offering agents with interdependent compensation policies. These policies imply that the effort exerted by an employee has a direct externality on the expected wage of his peers. A psychological gift-exchange game takes place among the employees, who consider the effort choices of their colleagues as kind or unkind depending on how peers performance indicators impact on salaries. I find that when agents are only modestly risk averse, the principal finds it optimal to design a relative-incentive scheme in which an employee struggles to signal himself as the best performer

---

<sup>44</sup>Kosfeld and von Siemens (2011) study a competitive labor market in which workers differ in their willingness to cooperate voluntarily. They show that there always exists a separating equilibrium in which workers self-select into firms that differ in their monetary incentives as well as their level of worker cooperation. However, they neither consider a psychological game setting nor extend their analysis to the case in which agents are risk-averse.

<sup>45</sup>In a related work-in-progress project, I investigate how the interplay highlighted in this paper also affect the principal's choice of the production technology.

<sup>46</sup>Rotemberg (1994) is a notable exception.

<sup>47</sup>An incentive scheme is said to be generous as the agent's participation constraint is slack at the optimum, the agent being guaranteed a material utility level bigger than what he can achieve by refusing the principal's offer. A similar intuition can be found in Dufwenberg and Kirchsteiger (2000).

of the group. Conversely, when agents are sufficiently risk averse a joint-incentive scheme always proves superior. This results suggest that the complementarity between explicit monetary and implicit psychological incentives can be exploited in two completely different ways, namely, either by inducing friendship or competition among coworkers. In a companion project (Livio 2015), I investigate how the principal's choice between a group and an individual production technology is affected by the interplay between risk and reciprocal attitudes of the agents.

My results may serve as guidelines for empirical investigations both in controlled field and lab experiments and using survey data. In particular: (i) the use of relative-incentive schemes should be empirically more prevalent among employees with a lower degree of risk aversion; (ii) the power of explicit monetary incentives should be lower the higher the degree of social preferences of the workforce; (iii) the productivity of a team of agents should increase with the degree of reciprocity concerns of the team's members.

The analysis carried in this paper relies on the assumption that agents interact in a one-shot game.<sup>48</sup> One criticism could be that, in the long run, life in an extra-competitive working environment in which negative reciprocity is daily experienced may generate a lot of stress and anger, whereas a friendly working environment in which positive reciprocity is experienced may generate happiness and satisfaction, thereby making joint incentives more appealing.<sup>49</sup> However, (i) a too friendly environment may induce too much social interactions among coworkers, thereby creating incentives for time and resource wasting and (ii) psychologists (see e.g., Lerner and Keltner 2001) have shown that people tend to react to anger by making more optimistic judgements of future events, which has been recognized to be an important source of creativity in innovative firms. Integrating these dynamics in the baseline set up may gather additional, interesting insights to both academics and practitioners.

## Appendix A

**Proof of Lemma 1.** In order for an incentive scheme to induce a psychological equilibrium of the subgame starting at stage 2 in which both agents exerts positive effort, it must be that the overall utility of playing  $e_i = 1$  is greater than the overall utility of playing  $e_i = 0$  given  $\beta_i(e_j = 1) = \chi_i(e_i = 1) = 1 \forall i, j \in \{a, b\}$ . In what follows I check each incentive scheme separately.

It is immediate to see that scheme (1) does not entail any reciprocal exchange, the salary paid to agent  $i$  being independent of  $x_j$  and therefore of  $e_j$ . Therefore, the principal's problem collapses to the standard one and there exists an equilibrium of the subgame starting at stage 2 in which scheme (1) induces high effort provision from both agents.

✓

When the team-incentive scheme (2) is offered and  $i$  believes that  $j$  exerts high effort,  $i$  believes that the state  $\{x_i = h, x_j = h\}$  occurs with probability  $q^2$  if  $i$  works hard and with probability  $q(1 - q)$  if  $i$  shirks. If the principal pays scheme (2), high effort by  $i$  represents a kind action towards  $j$  as it increases the latter expected material payoff. Mathematically, the kindness  $i$  expects to give to  $j$  choosing  $e_i = 1$  when he has a first order belief of  $\beta_i(e_j = 1) = 1$  is:

$$k_{ij}(1,1) = q^2 w_{hh}^j - \phi - \frac{1}{2} \left( q^2 w_{hh}^j - \phi + q(1 - q) w_{hh}^j - \phi \right)$$

<sup>48</sup>I do not think that considering a repeated-interaction setting *per-se* would dramatically impact on my results. Indeed, I expect repeated interaction to generate additional implicit incentives to enforce the gift-exchange game described in the analysis carried out here.

<sup>49</sup>Mikula et al. (1998) negative reciprocity is associated with the emotional state of anger, whereas positive reciprocity is associated with states of happiness and satisfaction, as well as with anticipation of guilt if one did not reciprocate a positive behaviour.

where the term in parenthesis indicates  $j$ 's equitable payoff given  $i$ 's beliefs. The expression above can be rewritten as:

$$\begin{aligned} k_{ij}(1,1) &= \frac{q(2q-1)w_{hh}^j}{2} \\ &= \frac{\phi}{2} \end{aligned}$$

which is always positive. The idea is that  $i$  behavior is kind to  $j$  as  $i$ 's effort makes it more easy for  $j$  to collect a positive bonus. If  $i$  deviates and decides to shirk, the kindness  $i$  expects to give to  $j$  is:

$$k_{ij}(0,1) = -\frac{\phi}{2},$$

which is always negative as the payoff  $i$  expects to give to  $j$  is lower than the latter's equitable payoff. Similarly, the kindness  $i$  expects to receive from  $j$  is given by:

$$\lambda_{iji}(1,1) = \frac{q(2q-1)w_{hh}^i}{2} = \frac{\phi}{2}.$$

Therefore,  $i$  expects to receive a positive psychological payoff when exerting high effort and a negative psychological payoff when exerting low effort. Hence, when scheme (2) is offered and  $\beta_i(e_j = 1) = \chi_i(e_i = 1) = 1$ , agent  $i$  achieves an overall utility of:

$$\frac{q^2\phi}{(2q-1)q} - \phi + \gamma\frac{\phi}{2} \quad (\text{A.1})$$

when he chooses  $e_i = 1$  and of:

$$\frac{q(1-q)\phi}{(2q-1)q} - \gamma\frac{\phi}{2} \quad (\text{A.2})$$

It can be easily verified that (A.1) is (strictly) greater than (A.2) for any  $\gamma \geq (>)$  than zero.

When tournament scheme (3) is offered and  $\beta_i(e_j) = \chi_i(e_i) = 1 \forall i, j \in \{a, b\}, i \neq j$ ,  $i$  expected kindness towards  $j$  when choosing  $e_i = 1$  is given by:

$$\begin{aligned} k_{ij}(1,1) &= q(1-q)w_{hl}^j - \phi - \frac{1}{2} [q(1-q)w_{hl}^j - \phi + q^2w_{hl}^j - \phi] \\ &= -\frac{q(2q-1)w_{hl}^j}{2} \\ &= -\frac{q\phi}{2(1-q)} \end{aligned}$$

which is negative as high effort now entails to  $j$  an expected payoff lower than his equitable payoff. Conversely,  $i$  expected kindness towards  $j$  if  $i$  chooses  $e_i = 0$  is given by:

$$k_{ij}(0,1) = \frac{q\phi}{2(1-q)}.$$

which is always positive. Similarly, the kindness  $i$  expects to receive from  $j$  is given by

$$\lambda_{iji}(1,1) = -\frac{q(2q-1)w_{hl}^i}{2} = -\frac{q\phi}{2(1-q)}$$

Therefore,  $i$ 's overall utility is given by:

$$(1-q)q\frac{\phi}{(2q-1)(1-q)} + \gamma\frac{q\phi}{2(1-q)} - \phi \quad (\text{A.3})$$

when he chooses  $e_i = 1$  and equal to:

$$(1-q)^2 \frac{\phi}{(2q-1)(1-q)} - \gamma \frac{q\phi}{2(1-q)} \quad (\text{A.4})$$

Again, it can be easily verified that (A.3) is (strictly) greater than (A.4) for any  $\gamma \geq (>)$  than zero.  $\square$

**Proof of Lemma 2.** The fact that scheme (1) entails a psychological payoff equal to zero follow from the first part of the proof of Lemma 1.

The fact that scheme (2) entails a psychological payoff equal to  $\gamma\phi/2$  can be easily verified from equations (A.1). The overall psychological incentive provided by scheme (2) is given by  $\gamma\phi$  and can be verified from (A.1) and (A.2).

Similarly, the fact that scheme (3) entails a psychological payoff equal to  $(q\gamma\phi)/(2(1-q))$  can be easily verified from equations (A.3). The overall psychological incentive provided by scheme (3) is given by  $q\gamma\phi/(1-q)$  and can be verified by (A.3) and (A.4).  $\square$

### Optimal incentive schemes for reciprocal agents

In this part of the Appendix I characterize the solution to the general problem of the principal described in section 2.3. I derive a series of results (Observations 6 and 7 and Equations A.9, A.10, A.15 and A.16) that are useful to prove the remaining propositions and lemmas.

The principal minimizes:

$$\min_{w_{x_i x_j}^i, w_{x_j x_i}^i, i, j \in \{a, b\}, i \neq j} q^2 \sum_{i \in \{a, b\}} w_{hh}^i + (1-q)q \sum_{i \in \{a, b\}} (w_{hl}^i + w_{lh}^i) + (1-q)^2 \sum_{i \in \{a, b\}} w_{ll}^i \quad (\text{EW})$$

subject to the *incentive compatibility constraints*

$$U_i(1, 1, 1) \geq U_i(0, 1, 1) \quad (\text{IC}_i)$$

for all  $i \in \{a, b\}$ . (IC<sub>i</sub>) can be rewritten as:

$$q^2 u(w_{hh}^i) + (1-q)q(u(w_{hl}^i) + u(w_{lh}^i)) + (1-q)^2 u(w_{ll}^i) - \phi + \gamma \text{sign}(\lambda_{iji}(1, 1) * k_{ij}(1, 1)) \sqrt{|\lambda_{iji}(1, 1)|} \sqrt{|k_{ij}(1, 1)|} \geq q(1-q)u(w_{hh}^i) + (1-q)^2 u(w_{hl}^i) + q^2 u(w_{lh}^i) + q(1-q)u(w_{ll}^i) + \gamma \text{sign}(\lambda_{iji}(1, 1) * k_{ij}(0, 1)) \sqrt{|\lambda_{iji}(1, 1)|} \sqrt{|k_{ij}(0, 1)|}$$

for all  $i, j \in \{a, b\}, i \neq j$  and where  $u(w) = w^{1-\rho}$  and:

$$k_{ij}(1, 1) = \frac{(2q-1)}{2} [q(u(w_{hh}^j) - u(w_{hl}^j)) + (1-q)(u(w_{lh}^j) - u(w_{ll}^j))] = \lambda_{ijj}(1, 1)$$

and

$$k_{ij}(0, 1) = -\frac{(2q-1)}{2} [q(u(w_{hh}^j) - u(w_{hl}^j)) + (1-q)(u(w_{lh}^j) - u(w_{ll}^j))]$$

Therefore, the lagrangian associated with the principal's problem can be written as:

$$L = q^2 \sum_{i \in \{a, b\}} w_{hh}^i + (1-q)q \sum_{i \in \{a, b\}} (w_{hl}^i + w_{lh}^i) + (1-q)^2 \sum_{i \in \{a, b\}} w_{ll}^i - \mu_a [(2q-1) * [q(u(w_{hh}^a) - u(w_{hl}^a)) + (1-q)u(w_{hl}^a) - u(w_{ll}^a)] - \phi + \gamma (r_a(\lambda_{aba}(11), k_{ab}(11)) - r_a(\lambda_{aba}(11), k_{ab}(01)))] - \mu_b [(2q-1) * [q(u(w_{hh}^b) - u(w_{hl}^b)) + (1-q)u(w_{hl}^b) - u(w_{ll}^b)] - \phi + \gamma (r_b(\lambda_{bab}(11), k_{ba}(11)) - r_b(\lambda_{bab}(11), k_{ba}(10)))]$$



The set of first order conditions for agent  $i$  are as follows:

$$\begin{aligned} \frac{\partial L}{\partial w_{hh}^i} &= q^2 - \mu_i \left[ q(2q-1)u'(w_{hh}^i) + \gamma \left( \frac{\partial r_i(\lambda_{iji}(1,1), k_{ij}(1,1))}{\partial w_{hh}^i} - \frac{\partial r_i(\lambda_{iji}(1,1), k_{ij}(0,1))}{\partial w_{hh}^i} \right) \right] \\ &\quad - \mu_j \gamma \left( \frac{\partial r_j(\lambda_{jij}(1,1), k_{ji}(1,1))}{\partial w_{hh}^a} - \frac{\partial r_j(\lambda_{jij}(1,1), k_{ji}(0,1))}{\partial w_{hh}^i} \right) \geq 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial L}{\partial w_{hl}^i} &= q(1-q) - \mu_i \left[ (1-q)(2q-1)u'(w_{hl}^i) + \gamma \left( \frac{\partial r_i(\lambda_{iji}(1,1), k_{ij}(1,1))}{\partial w_{hl}^i} - \frac{\partial r_i(\lambda_{iji}(1,1), k_{ij}(0,1))}{\partial w_{hl}^i} \right) \right] \\ &\quad - \mu_j \gamma \left( \frac{\partial r_j(\lambda_{jij}(1,1), k_{ji}(1,1))}{\partial w_{hl}^i} - \frac{\partial r_j(\lambda_{jij}(1,1), k_{ji}(0,1))}{\partial w_{hl}^i} \right) \geq 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial L}{\partial w_{lh}^i} &= (1-q)q - \mu_i \left[ -q(2q-1)u'(w_{lh}^i) + \gamma_i \left( \frac{\partial r_i(\lambda_{iji}(1,1), k_{ij}(1,1))}{\partial w_{lh}^i} - \frac{\partial r_i(\lambda_{iji}(1,1), k_{ij}(0,1))}{\partial w_{lh}^i} \right) \right] \\ &\quad - \mu_j \gamma \left( \frac{\partial r_j(\lambda_{jij}(1,1), k_{ji}(1,1))}{\partial w_{lh}^i} - \frac{\partial r_j(\lambda_{jij}(1,1), k_{ji}(0,1))}{\partial w_{lh}^i} \right) \geq 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial L}{\partial w_{ll}^i} &= (1-q)^2 - \mu_i \left[ -(1-q)(2q-1)u'(w_{ll}^i) + \gamma \left( \frac{\partial r_i(\lambda_{iji}(1,1), k_{ij}(1,1))}{\partial w_{ll}^i} - \frac{\partial r_i(\lambda_{iji}(1,1), k_{ij}(0,1))}{\partial w_{ll}^i} \right) \right] \\ &\quad - \mu_j \gamma \left( \frac{\partial r_j(\lambda_{jij}(1,1), k_{ji}(1,1))}{\partial w_{ll}^i} - \frac{\partial r_j(\lambda_{jij}(1,1), k_{ji}(0,1))}{\partial w_{ll}^i} \right) \geq 0 \end{aligned}$$

where e.g.:

$$\frac{\partial r_i(\lambda_{iji}(1,1), k_{ij}(1,1))}{\partial w_{hh}^i} = \frac{\partial \lambda_{iji}(1,1)}{\partial w_{hh}^i} \text{sign}(\lambda_{iji}(1,1) * k_{ij}(1,1)) \frac{1}{2} \frac{\sqrt{|k_{ij}(1,1)|}}{|\lambda_{iji}(1,1)|^{\frac{3}{2}}} \lambda_{iji}(1,1)$$

$$\frac{\partial r_j(\lambda_{jij}(1,1), k_{ji}(1,1))}{\partial w_{hh}^i} = \frac{\partial k_{ji}(1,1)}{\partial w_{hh}^i} \text{sign}(\lambda_{jij}(1,1) * k_{ji}(1,1)) \frac{1}{2} \frac{\sqrt{|\lambda_{jij}(1,1)|}}{|k_{ji}(1,1)|^{\frac{3}{2}}} k_{ji}(1,1)$$

Therefore, the solution depends on the type of reciprocal exchange that is induced in equilibrium by the principal, i.e., by the sign of  $k_{ij}(1,1)$ ,  $i \in \{a, b\}$  (which in equilibrium is equal to  $\lambda_{jij}(1,1)$ ), as well as by the kindness of the effort choice agent  $i$  gives up, i.e.,  $k_{ij}(0,1)$ .

If the principal does not want to induce a psychological gift exchange, i.e.,  $\lambda_{iji}(1,1) = k_{ij}(1,1) = k_{ij}(0,1) = 0$ , then solving the system of first-order conditions lead to the standard incentive scheme (4), which shows Observation 1.

If the principal wants to induce a positive psychological exchange, i.e.,  $\lambda_{iji}(1,1) = k_{ij}(1,1) = k_{ij}(0,1) > 0$ ,

and given that agents are identical, I can rewrite the system of first-order conditions as follows:

$$\frac{\partial L}{\partial w_{hh}^i} \geq 0 \Rightarrow = q^2 - \mu [q(2q - 1)(1 + \gamma)] u'(w_{hh}^i) \geq 0 \quad (\text{A.5})$$

$$\frac{\partial L}{\partial w_{hl}^i} \geq 0 \Rightarrow = q(1 - q) - \mu [(2q - 1)(1 - q - \gamma q)] u'(w_{hl}^i) \geq 0 \quad (\text{A.6})$$

$$\frac{\partial L}{\partial w_{lh}^i} \geq 0 \Rightarrow = (1 - q)q - \mu [(2q - 1)(-q + \gamma(1 - q))] u'(w_{lh}^i) \geq 0 \quad (\text{A.7})$$

$$\frac{\partial L}{\partial w_{ll}^i} \geq 0 \Rightarrow = (1 - q)^2 + \mu [(2q - 1)(1 + \gamma)(1 - q)] u'(w_{ll}^i) \geq 0 \quad (\text{A.8})$$

It can be easily seen that  $\mu > 0$ . Moreover, the following observation holds:

**Observation 6.** From conditions (A.5)-(A.8) I can see that  $w_{hh}^i = w_{ll}^i = 0$  if  $\gamma \in (0, 1 - q/q)$  and  $w_{hl}^i = w_{lh}^i = w_{ll}^i = 0$  if  $\gamma \in [1 - q/q, 1]$ . This implies that joint-incentives are used by the principal to induce the most efficient positive psychological exchange.

If  $\gamma \in (0, 1 - q/q)$ , from (A.5)-(A.6) I achieve:

$$\frac{1}{u'(w_{hh}^i)} = \frac{\mu(2q - 1)(1 + \gamma)}{q}$$

$$\frac{1}{u'(w_{hl}^i)} = \frac{\mu(2q - 1)(1 - q - \gamma q)}{q(1 - q)}$$

that leads to:

$$w_{hl}^i = \left( \underbrace{\frac{(1 - q - \gamma q)}{(1 + \gamma)(1 - q)}}_{\in(0,1)} \right)^{\frac{1}{\rho}} w_{hh}^i \quad (\text{A.9})$$

If  $\gamma \in [(1 - q)/q, q/(1 - q)]$ , from Observation 6 I know that only  $w_{hh}^i$  takes on positive values. Replacing this into the (ICi):

$$(2q - 1)(1 + \gamma)q(w_{hh}^i)^{1-\rho} = \phi. \quad (\text{A.10})$$

If the principal wants to induce a negative psychological exchange, i.e.,  $\lambda_{iji}(1, 1) = k_{ij}(1, 1) = k_{ij}(0, 1) < 0$ ,

and given that agents are identical, I can rewrite the system of first-order conditions as follows:

$$\frac{\partial L}{\partial w_{hh}^i} \geq 0 \Rightarrow = q^2 - \mu [(2q - 1)q(1 - \gamma)] u'(w_{hh}^i) \geq 0 \quad (\text{A.11})$$

$$\frac{\partial L}{\partial w_{hl}^i} \geq 0 \Rightarrow = q(1 - q) - \mu [(2q - 1)(1 - q + \gamma q)] u'(w_{hl}^i) \geq 0 \quad (\text{A.12})$$

$$\frac{\partial L}{\partial w_{lh}^i} \geq 0 \Rightarrow = (1 - q)q + \mu [(2q - 1)(q + \gamma(1 - q))] u'(w_{lh}^i) \geq 0 \quad (\text{A.13})$$

$$\frac{\partial L}{\partial w_{ll}^i} \geq 0 \Rightarrow = (1 - q)^2 + \mu [(2q - 1)(1 - \gamma)(1 - q)] u'(w_{ll}^i) \geq 0 \quad (\text{A.14})$$

It can be easily seen that  $\mu > 0$ . Moreover, the following observation holds:

**Observation 7.** From conditions (A.11)-(A.14) I can see that  $w_{lh}^i = w_{ll}^i = 0$  if  $\gamma \in (0, 1)$  and  $w_{hh}^i = w_{lh}^i = w_{ll}^i = 0$  if  $\gamma = 1$ . This implies that relative-incentives are used by the principal to induce the most efficient negative psychological exchange.

If  $\gamma \in (0, 1)$ , then from conditions (A.11) and (A.12) we achieve:

$$\frac{1}{u'(w_{hh}^i)} = \frac{\mu(2q - 1)(1 - \gamma)}{q}$$

$$\frac{1}{u'(w_{hl}^i)} = \frac{\mu(2q - 1)(1 - q + \gamma q)}{q(1 - q)}$$

that leads to:

$$w_{hh}^i = \left( \underbrace{\frac{(1 - \gamma)(1 - q)}{(1 - q + \gamma q)}}_{\in(0,1)} \right)^{\frac{1}{\rho}} w_{hl}^i \quad (\text{A.15})$$

If  $\gamma = 1$ , the non-negativity constraints associated with  $w_{hh}^i, w_{lh}^i, w_{ll}^i$  bite, while only  $w_{hl}^i$  takes on positive values. Replacing this into the (ICi):

$$(2q - 1)(1 - q + \gamma q)(w_{hl}^i)^{1-\rho} = \phi$$

which leads to

$$w_{hh}^i = \left( \frac{\phi}{(2q - 1)(1 - q + \gamma q)} \right)^{\frac{1}{1-\rho}} \quad (\text{A.16})$$

**Proof of Lemma 3.** Considering scheme (4) and equations (A.9), (A.10), (A.15) and (A.16) for  $\rho = 0$  shows the lemma.  $\square$

**Proof of Proposition 1.** From Lemma 3 I know that scheme (1) entails a total expected cost of:

$$E(W(1)) = 2q \frac{\phi}{(2q - 1)}$$

Similarly, we know that schemes (2) and (3) entail a total expected cost of:

$$E(W(2)) = 2q^2 \frac{\phi}{(2q-1)(1+\gamma)q}$$

and

$$E(W(3)) = 2(1-q)q \frac{\phi}{(2q-1)(1-q+q\gamma)}$$

respectively. Simple algebra shows that  $E(W(3))$  is strictly smaller than  $E(W(1))$  and  $E(W(2))$  for any  $\gamma > 0$ .  $\square$

**Proof of Lemma 4.** If  $\gamma \in (0, (1-q)/q)$ , replacing equation (A.9) into (ICi) and making it binding leads to:

$$(2q-1)[q(1+\gamma)u(w_{hh}^i) + (1-q-\gamma q)u(w_{hl}^i)] = \phi$$

$$(2q-1) \left[ q(1+\gamma) + (1-q-\gamma q) \left( \frac{(1-q-q\gamma)}{(1+\gamma)(1-q)} \right)^{\frac{1-\rho}{\rho}} \right] (w_{hh}^i)^{(1-\rho)} = \phi$$

That together with Observation 6 proof the first part of the lemma.  $\checkmark$

If  $\gamma \in [(1-q)/q, 1]$ , replacing (A.10) into (ICi) and making it binding leads to:

$$w_{hh}^i = \left( \frac{\phi}{(2q-1)(1+\gamma)q} \right)^{\frac{1}{1-\rho}}.$$

That together with Observation 6 shows the second part of the lemma.  $\square$

**Proof of Lemma 5.** If  $\gamma \in (0, 1)$ , replacing equation (A.15) into (ICi) and making it binding leads to:

$$(2q-1)[q(1-\gamma)u(w_{hh}^i) + (1-q+\gamma q)u(w_{hl}^i)] = \phi$$

$$(2q-1) \left[ (1-q+q\gamma) + (1-\gamma)q \left( \frac{(1-\gamma)(1-q)}{(1-q+q\gamma)} \right)^{\frac{1-\rho}{\rho}} \right] (w_{hl}^i)^{(1-\rho)} = \phi$$

That together with Observation 7 shows the first part of the lemma.  $\checkmark$

If  $\gamma = 1$ , replacing (A.16) into (ICi) we achieve:

$$w_{hh}^i = \left( \frac{\phi}{(2q-1)(1-q+q\gamma)} \right)^{\frac{1}{1-\rho}}.$$

That together with Observation 7 shows the second part of the lemma.  $\square$

From Observation 1 and Lemmas 4 and 5 I can easily compute the expected wage bill paid by the principal under each solution candidate, that are, respectively:

$$E(W(BI)) = 2q \left( \frac{\phi}{(2q-1)} \right)^{\frac{1}{1-\rho}} \text{ for any } \gamma \in (0, 1].$$

$$E(W(BJ)) = \begin{cases} 2 \left( q + (1-q) \left( \frac{1-q-q\gamma}{(1+\gamma)(1-q)} \right)^{\frac{1}{\rho}} \right) q \left( \frac{\phi}{(2q-1) \left( q(1+\gamma) + (1-q-\gamma q) \left( \frac{1-q-q\gamma}{(1+\gamma)(1-q)} \right)^{\frac{1-\rho}{\rho}} \right)} \right)^{\frac{1}{1-\rho}} & \text{if } \gamma \in [0, (1-q)/q); \\ 2q^2 \left( \frac{\phi}{(2q-1)(1+\gamma)q} \right)^{\frac{1}{1-\rho}} & \text{if } \gamma \in [(1-q)/q, 1]. \end{cases}$$

and

$$E(W(BR)) = \begin{cases} 2 \left( 1 - q + q \left( \frac{(1-q)(1-\gamma)}{1-q+q\gamma} \right)^{\frac{1}{\rho}} \right) q \left( \frac{\phi}{(2q-1) \left( (1-q+\gamma q) + q(1-\gamma) \left( \frac{(1-\gamma)(1-q)}{1-q+q\gamma} \right)^{\frac{1-\rho}{\rho}} \right)} \right)^{\frac{1}{1-\rho}} & \text{if } \gamma \in (0, 1); \\ 2(1-q)q \left( \frac{\phi}{(2q-1)(1-q+q\gamma)} \right)^{\frac{1}{1-\rho}} & \text{if } \gamma = 1. \end{cases}$$

Then notice the following observations:

**Observation 8.** It can be immediately verified that  $E(W(BI))$ ,  $E(W(BJ))$  and  $E(W(BR))$  are continuous and differentiable in  $\rho$  for given  $q$ ,  $\gamma$  and  $\phi$ .

**Observation 9.** It can be immediately verified that  $E(W(BJ))$  and  $E(W(BR))$  are continuous functions of  $\gamma$ .

**Proof of Observation 2.** This observation can be immediately verified by considering  $E(W(BJ))$  and  $E(W(BR))$  at  $\gamma \rightarrow 0$ .  $\square$

**Proof of Observation 3.** From Observation 9 we know that  $E(W(BJ))$  and  $E(W(BR))$  are continuous in  $\gamma$ . Taking derivative of both  $E(W(BJ))$  and  $E(W(BR))$  with respect to  $\gamma$  leads to:

$$\frac{\partial E(W(BJ))}{\partial \gamma} = \begin{cases} \frac{2q^2 \left[ (1+\gamma) \left( (1-q) \left( \frac{1-q-q\gamma}{(1+\gamma)(1-q)} \right)^{\frac{1}{\rho}} + q \right) - 1 \right]}{(1-\rho)(1+\gamma)(1-q-q\gamma) \left( \frac{\phi}{(1+\gamma)(2q-1) \left( (1-q) \left( \frac{1-q-q\gamma}{(1+\gamma)(1-q)} \right)^{\frac{1}{\rho}} + q \right)} \right)^{\frac{1}{\rho-1}}} & \text{if } \gamma \in (0, (1-q)/q) \\ \frac{2q^2}{(r-1)(\gamma+1) \left( \frac{\phi}{q(2q-1)(\gamma+1)} \right)^{\frac{1}{\rho-1}}} & \text{if } \gamma \in [(1-q)/q, 1] \end{cases}$$

It can be seen that both expressions are always strictly negative for all the admissible parameter values.

$$\frac{\partial E(W(BJ))}{\partial \gamma} = \begin{cases} \frac{2q^2 \left( (q\gamma - q + 1) \left( \frac{(q-1)(\gamma-1)}{q\gamma - q + 1} \right)^{\frac{1}{\rho}} + (1-q)\gamma + q - 1 \right)}{(\rho-1)(\gamma-1)(q\gamma - q + 1) \left( - \frac{\phi(q-1)}{(2q-1)(q\gamma - q + 1) \left( q \left( \frac{(q-1)(\gamma-1)}{q\gamma - q + 1} \right)^{\frac{1}{\rho}} - q + 1 \right)} \right)^{\frac{1}{\rho-1}}} & \text{if } \gamma \in (0, 1) \\ \frac{2(1-q)q^2}{(\rho-1)(q\gamma - q + 1) \left( \frac{\phi}{(2q-1)(q\gamma - q + 1)} \right)^{\frac{1}{\rho-1}}} & \text{if } \gamma = 1 \end{cases}$$

It can be seen that both expressions are always strictly negative for all the admissible parameter values.  $\square$

**Proof of Lemma 6.** The proof of the lemma follows from Observations 2, 3 and 9.  $\square$

Then, let us define  $DiffBRBJ(\rho, \gamma, q, \phi)$  as:

$$DiffBRBJ(\rho, \gamma, q, \phi) = E(W(BR)) - E(W(BJ))$$

**Lemma 8.**  $\lim_{\rho \rightarrow 1^-} DiffBRBJ(\rho, \gamma, q, \phi) > 0$  for all  $\gamma \in (0, 1]$ ,  $q \in (1/2, 1)$  and  $\phi > (2q - 1)$ .

**Proof.**

- For  $\gamma \in (0, (1-q)/q]$ ,  $DiffBRBJ(\cdot, \cdot, \cdot, \cdot)$  is positive if and only if:

$$\underbrace{\left(1 - q + q \left(\frac{(1-\gamma)(1-q)}{1-q+q\gamma}\right)^{\frac{1}{\rho}}\right)^{1-\rho} \left(q(1+\gamma) + (1-q-q\gamma) \left(\frac{1-q-q\gamma}{(1+\gamma)(1-q)}\right)^{\frac{1-\rho}{\rho}}\right)}_{=D} - \underbrace{\left(q + (1-q) \left(\frac{1-q-q\gamma}{(1+\gamma)(1-q)}\right)^{\frac{1}{\rho}}\right)^{1-\rho} \left((1-q+q\gamma) + q(1-\gamma) \left(\frac{(1-\gamma)(1-q)}{1-q+q\gamma}\right)^{\frac{1-\rho}{\rho}}\right)}_{=E} > 0$$

Taking Taylor expansion series centered at  $\rho = 1$  of D and E leads to:

$$\begin{aligned} & 1 + (\rho - 1) \left( -(1-q-q\gamma) \ln \left( \frac{1-q-q\gamma}{(1-q)(1+\gamma)} \right) - \ln \left( \frac{1-q}{1-q-q\gamma} \right) \right) + \mathcal{O}((\rho-1)^2) \\ & - 1 - (\rho - 1) \left( (\gamma-1)q \ln \left( \frac{(1-\gamma)(1-q)}{1-q+q\gamma} \right) - \ln \left( \frac{1}{1+\gamma} \right) \right) + \mathcal{O}((\rho-1)^2) \\ & \simeq (\rho - 1) [-(1-q-q\gamma) \ln(1-q-q\gamma) - (2q\gamma) \ln(1-q) - q(1+\gamma) \ln(1+\gamma)] \\ & + (\rho - 1) [q(1-\gamma) \ln(1-\gamma) + (1-q+q\gamma) \ln(1-q+q\gamma)] \end{aligned}$$

which is always positive in this region of the parameters values.

✓

- For  $\gamma \in ((1-q)/q, 1)$ ,  $DiffBRBJ(\cdot, \cdot, \cdot, \cdot)$  is positive if and only if:

$$\begin{aligned} & \left(1 - q + q \left(\frac{(1-\gamma)(1-q)}{1-q+q\gamma}\right)^{\frac{1}{\rho}}\right)^{1-\rho} q(1+\gamma) \\ & \left((1-q+q\gamma) + q(1-\gamma) \left(\frac{(1-\gamma)(1-q)}{1-q+q\gamma}\right)^{\frac{1}{\rho}}\right)^{1-\rho} \end{aligned}$$

At the limit of  $\rho \rightarrow 1^-$ , this expression can be written as:

$$(1+\gamma)q - 1 > 0$$

which is always satisfied in this interval.

✓

- Finally, for  $\gamma = 1$  it can be easily seen that  $DiffBRBJ(\cdot, \cdot, \cdot, \cdot)$  is positive whenever:

$$(1-q)^{1-\rho} q(1+\gamma) - q^{1-\rho} (1-q+q\gamma) > 0$$

which always holds for  $\rho$  sufficiently closed to 1.

□

**Proof of Proposition 2.** The existence of  $\rho^*$  when  $\gamma = 1$  can be easily shown considering  $DiffBRBJ(\rho, 1, q, \phi)$  which is given by:

$$2q(1-q) \left( \frac{\phi}{(2q-1)(1-q+q\gamma)} \right)^{\frac{1}{1-\rho}} - 2q^2 \left( \frac{\phi}{(2q-1)q(1+\gamma)} \right)$$

which is  $\geq 0$  whenever:

$$(1 - q)^{1-\rho} \geq \frac{1}{2q^\rho} \quad (\text{A.20})$$

It can be immediately seen that (A.20) holds with equality for a risk-aversion value  $\rho^* \in (0.5, 1)$ . If  $\gamma \in (0, 1)$ , the existence of  $\rho^*$  can be shown in four steps:

1. From Observation 8 it follows that  $\text{DiffBRBJ}(\cdot, \cdot, \cdot, \cdot)$  is continuous and differentiable in  $\rho$  for given  $q, \gamma, \phi$ .
2. We know from Proposition 1 that, for  $\rho = 0$ ,  $\text{DiffBRBJ}(0, \gamma, q, \phi) < 0$  for all  $\gamma \in (0, 1]$ ,  $q \in (1/2, 1)$  and  $\phi > (2q - 1)$ .
3. We know from Lemma 8 that  $\rho \rightarrow 1$ ,  $\text{DiffBRBJ}(\rho, \gamma, q, \phi) > 0$  for all  $\gamma \in (0, 1]$ ,  $q \in (1/2, 1)$  and  $\phi > (2q - 1)$ .
4. Moreover, taking derivative of  $\text{DiffBRBJ}(\rho, \gamma, q, \phi)$  with respect to  $\rho$  I achieve:
  - if  $\gamma \in (0, (1 - q)/q)$ :

$$\begin{aligned} \frac{\partial \text{DiffBRBJ}(\rho, \gamma, q, \phi)}{\partial \rho} = & \frac{q(\rho(H^{\frac{1}{\rho}}[q \ln I - q \ln H] + (1 - q) \ln I) + qH^{\frac{1}{\rho}} \ln H)}{I^{\frac{1}{\rho-1}}(\rho - 1)^2\rho} \\ & - \frac{q(\rho(F^{\frac{1}{\rho}}[(1 - q) \ln G + (1 - q) \ln F] + q \ln G) - (1 - q)F^{\frac{1}{\rho}} \ln F)}{G^{\frac{1}{\rho-1}}(\rho - 1)^2\rho} \end{aligned}$$

where  $F = \frac{1-q-q\gamma}{(1-q)(1+\gamma)}$ ,  $G = \frac{\phi}{(1+\gamma)(2q-1)((1-q)F^{1/\rho+q})}$ ,  $H = \frac{(1-q)(1-\gamma)}{1-q+q\gamma}$  and  $I = \frac{(1-q)\phi}{(1-q+q\gamma)(2q-1)(-qH^{1/\rho+1-q})}$ .

It can be shown through numerical simulations that this expression is always positive in this region of the parameters values.

- if  $\gamma \in [(1 - q)/q, 1)$ :

$$\begin{aligned} \frac{\partial \text{DiffBRBJ}(\rho, \gamma, q, \phi)}{\partial \rho} = & \frac{q(\rho(H^{\frac{1}{\rho}}[q \ln I - q \ln H] + (1 - q) \ln I) + qH^{\frac{1}{\rho}} \ln H)}{I^{\frac{1}{\rho-1}}(\rho - 1)^2\rho} \\ & - \frac{q^2 \ln L}{L^{\frac{1}{\rho-1}}(\rho - 1)^2} \end{aligned}$$

where  $L = \frac{\phi}{(1+\gamma)q(2q-1)}$ . It can be shown through numerical simulations that this expression is always positive in this region of the parameter values.

The existence of  $\rho^*$  for all  $\gamma \in (0, 1)$  follows from points 1 to 4. □

**Proof of Observation 5.** This lemma can be proved considering  $\text{DiffBRBJ}(\rho, \gamma, q, \phi) = 0$  and applying the implicit function theorem to compute  $\frac{\partial \rho^*}{\partial q}$  and  $\frac{\partial \rho^*}{\partial \gamma}$ . □

### High effort costs

**Proof of Proposition 3.** Proposition 3 states that there exists a full dimension subset of the parameter space in which reciprocal agents expect to earn more and are asked to provide more effort than self interested agents. This directly comes from the fact that when the effort cost is low compared to the returns

to effort, the global optimum of the principal's problem does implement either the best-joint or the best-relative incentive scheme whereas the best-individual incentive scheme is always dominated. Therefore, when the cost of effort is above  $h^{(1-\rho)}(2q-1)$ , inducing positive effort from both agents is overly costly if they are selfish but not necessarily when they are reciprocal. In what follows I provide conditions on  $\phi$  such that reciprocal agents expect to earn more and are asked to work more than standard agents. In the text a condition for the case in which  $\rho > \rho^*$  and  $\gamma \in ((1-q)/q, 1]$  is provided. I have three cases left to consider:

1. If  $\rho < \rho^*$  and  $\gamma \in (0, 1)$ , the optimal scheme for reciprocal agents entails relative incentives. Therefore, reciprocal agents expect to earn more and are asked to work more than standard agents if:

$$h^{1-\rho}(2q-1) < \phi < \left( \frac{qh}{\left( (1-q) + q \left( \frac{(1-q)(1-\gamma)}{(1-q+q\gamma)} \right)^{\frac{1}{\rho}} \right)} \right)^{(1-\rho)} (2q-1) \left( (1-q+q\gamma) + q(1-\gamma) \left( \frac{(1-\gamma)(1-q)}{(1-q+q\gamma)} \right) \right)$$

2. If  $\rho < \rho^*$  and  $\gamma = 1$ , the optimal scheme for reciprocal agents entails relative-incentives. Therefore, reciprocal agents expect to earn more and are asked to work more than standard agents if:

$$h^{1-\rho}(2q-1) < \phi < \left( \frac{qh}{1-q} \right)^{(1-\rho)} (2q-1)(1-q+q\gamma)$$

3. If  $\rho > \rho^*$  and  $\gamma \in (0, (1-q)/q]$ , the optimal scheme for reciprocal agents entails joint-incentives. Therefore, reciprocal agents expect to earn more and are asked to work more than standard agents if:

$$h^{1-\rho}(2q-1) < \phi < \left( \frac{qh}{\left( q + (1-q) \left( \frac{(1-q-q\gamma)}{(1+\gamma)(1-q)} \right)^{\frac{1}{\rho}} \right)} \right)^{(1-\rho)} (2q-1) \left( (1+\gamma)q + (1-q-q\gamma) \left( \frac{(1-\gamma)(1-q)}{(1-q+q\gamma)} \right) \right)$$

□

## Group production

**Proof of Lemma 7.** When only the value of the total production can be contracted upon, there are only two possible state of nature:  $\{\underline{y}, \bar{y}\}$ . Given that  $\gamma \in (0, 1]$ , the only solution candidate to the principal's problem induces an exchange of kind behaviors from the employees. For  $e_i = \beta_j(e_i = 1) = \chi_i(e_i = 1) = 1$ ,  $i, j \in \{a, b\}$  and  $i \neq j$ , agent  $i$ 's incentive compatibility constraint can be written as:

$$\begin{aligned} & \bar{p}\bar{u}(w^i) + (1-\bar{p})\underline{u}(w^i) - \phi + \gamma \sqrt{\left| \frac{(\bar{p}-\hat{p})(u(\bar{w}^i) - u(\underline{w}^i))}{2} \right|} \sqrt{\left| \frac{(\bar{p}-\hat{p})(u(\bar{w}^j) - u(\underline{w}^j))}{2} \right|} \quad (\text{IC}_i) \\ & \geq \hat{p}\bar{u}(w^i) + (1-\hat{p})\underline{u}(w^i) - \gamma \sqrt{\left| \frac{(\bar{p}-\hat{p})(u(\bar{w}^i) - u(\underline{w}^i))}{2} \right|} \sqrt{\left| \frac{(\hat{p}-\bar{p})(u(\bar{w}^i) - u(\underline{w}^i))}{2} \right|} \end{aligned}$$



Given that agents are identical, solving the principal problem leads to the following first-order conditions for agent  $i$ :

$$\begin{aligned}\frac{\partial L}{\partial \bar{w}} &= \bar{p} - \mu [(\bar{p} - \hat{p})(1 + \gamma)] u'(\bar{w}) \geq 0 \\ \frac{\partial L}{\partial \underline{w}} &= (1 - \bar{p}) + \mu [(\bar{p} - \hat{p})(1 + \gamma)] u'(\underline{w}) \geq 0\end{aligned}$$

where  $u'(w) = (1 - \rho)w^{-\rho}$ . The non-negativity constraint associated with  $\underline{w}$  bites. Replacing  $\underline{w} = 0$  into the (IC $_i$ ) and solving for  $\bar{w}$  shows the lemma.  $\square$

**Proof of Proposition 4.** When all  $\{x_a, x_b, y\}$  can be contracted upon by the principal, there are eight possible state of nature:

- $\{\bar{y}, h, h\}$  that occurs with probability  $(\bar{p}q^2)$  if both agent exert effort equal to 1, w.p.  $(\hat{p}q(1 - q))$  if only agent  $i$  exerts effort 1 and if only  $j$  exerts effort equal 1;
- $\{\bar{y}, h, l\}$  w.p.  $(\bar{p}q(1 - q))$  if  $\{e_i = e_b = 1\}$ , w.p.  $(\hat{p}(1 - q)^2)$  if  $\{e_i = 1, e_j = 0\}$  and w.p.  $(\hat{p}q^2)$  if  $\{e_i = 0, e_j = 1\}$ ;
- $\{\bar{y}, l, h\}$  w.p.  $(\bar{p}(1 - q)q)$  if  $\{e_i = e_b = 1\}$ , w.p.  $(\hat{p}q^2)$  if  $\{e_i = 1, e_j = 0\}$  and w.p.  $(\hat{p}(1 - q)^2)$  if  $\{e_i = 0, e_j = 1\}$ ;
- $\{\bar{y}, l, l\}$  w.p.  $(\bar{p}(1 - q)^2)$  if  $\{e_i = e_b = 1\}$ , w.p.  $(1 - \hat{p})q(1 - q)$  if  $\{e_i = 1, e_j = 0\}$  and if  $\{e_i = 0, e_j = 1\}$ ;
- $\{y, h, h\}$  w.p.  $((1 - \bar{p})q^2)$  if  $\{e_i = e_b = 1\}$ , w.p.  $((1 - \hat{p})q(1 - q))$  if  $\{e_i = 1, e_j = 0\}$  and w.p.  $((1 - \hat{p})q^2)$  if  $\{e_i = 0, e_j = 1\}$ ;
- $\{y, h, l\}$  w.p.  $((1 - \bar{p})q(1 - q))$  if  $\{e_i = e_b = 1\}$ , w.p.  $((1 - \hat{p})(1 - q)^2)$  if  $\{e_i = 1, e_j = 0\}$  and w.p.  $((1 - \hat{p})q^2)$  if  $\{e_i = 0, e_j = 1\}$ ;
- $\{y, l, h\}$  w.p.  $((1 - \bar{p})q(1 - q))$  if  $\{e_i = e_b = 1\}$ , w.p.  $((1 - \hat{p})q^2)$  if  $\{e_i = 1, e_j = 0\}$  and w.p.  $((1 - \hat{p})(1 - q)^2)$  if  $\{e_i = 0, e_j = 1\}$ ;
- $\{y, l, l\}$  w.p.  $((1 - \bar{p})(1 - q)^2)$  if  $\{e_i = e_b = 1\}$ , w.p.  $((1 - \hat{p})(1 - q)q)$  if  $\{e_i = 1, e_j = 0\}$  and if  $\{e_i = 0, e_j = 1\}$ .

Let  $w_{yx_i x_j}^i$  be the salary paid to agent  $i$  when  $y, x_i$  and  $x_j$  realize.

If in equilibrium the principal decides to induce positive reciprocity, the first-order condition associated

with the principal's problem can be simplified as follows:

$$\frac{\partial L}{\partial w_{\underline{y}hh}^i} \geq 0 \Rightarrow \bar{p}q^2 - \mu[(\bar{p} + \hat{p})q - \hat{p}]q(1 + \gamma)(1 - \rho)(w_{\underline{y}hh}^i)^{-\rho} \geq 0 \quad (\text{A.21})$$

$$\frac{\partial L}{\partial w_{\underline{y}hh}^i} \geq 0 \Rightarrow (1 - \bar{p})q^2 - \mu[q(1 - \bar{p}) - (1 - q)(1 - \hat{p})]q(1 + \gamma)(1 - \rho)(w_{\underline{y}hh}^i)^{-\rho} \geq 0 \quad (\text{A.22})$$

$$\frac{\partial L}{\partial w_{\underline{y}hl}^i} \geq 0 \Rightarrow \bar{p}q(1 - q) - \mu[(1 - q)[\bar{p}q - (1 - q)\hat{p}] + q\gamma[(1 - q)\bar{p} - q\hat{p}]](1 - \rho)(w_{\underline{y}hl}^i)^{-\rho} \geq 0 \quad (\text{A.23})$$

$$\frac{\partial L}{\partial w_{\underline{y}hl}^i} \geq 0 \Rightarrow (1 - \bar{p})q(1 - q) - \mu[(1 - q)[q(1 - \bar{p}) - (1 - q)(1 - \hat{p})] + \gamma q[(1 - q)(1 - \bar{p}) - q(1 - \hat{p})]](1 - \rho)(w_{\underline{y}hl}^i)^{-\rho} \geq 0 \quad (\text{A.24})$$

$$\frac{\partial L}{\partial w_{\underline{y}lh}^i} \geq 0 \Rightarrow \bar{p}(1 - q)q - \mu[q[(1 - q)\bar{p} - q\hat{p}] - \gamma(1 - q)[q\bar{p} - \hat{p}(1 - q)]](1 - \rho)(w_{\underline{y}lh}^i)^{-\rho} \geq 0 \quad (\text{A.25})$$

$$\frac{\partial L}{\partial w_{\underline{y}lh}^i} \geq 0 \Rightarrow (1 - \bar{p})(1 - q)q - \mu[q[(1 - q)(1 - \bar{p}) - q(1 - \hat{p})] - \gamma(1 - q)[(1 - \bar{p})q - (1 - q)(1 - \hat{p})]](1 - \rho)(w_{\underline{y}lh}^i)^{-\rho} \geq 0 \quad (\text{A.26})$$

$$\frac{\partial L}{\partial w_{\underline{y}ll}^i} \geq 0 \Rightarrow \bar{p}(1 - q)^2 - \mu[(1 - q)(\bar{p}(1 - q) - \hat{p}q)](1 + \gamma)(1 - \rho)(w_{\underline{y}ll}^i)^{-\rho} \geq 0 \quad (\text{A.27})$$

$$\frac{\partial L}{\partial w_{\underline{y}ll}^i} \geq 0 \Rightarrow (1 - \bar{p})(1 - q)^2 - \mu[(1 - q)[(1 - q)(1 - \bar{p}) - q(1 - \hat{p})]](1 + \gamma)(1 - \rho)(w_{\underline{y}ll}^i)^{-\rho} \geq 0. \quad (\text{A.28})$$

**Observation 10.** From conditions (A.21)-(A.28) I have that, for any  $\rho \in (0, 1)$ :

- From conditions (A.21) and (A.23) we can see that  $w_{\underline{y}hh}^i$  is always positive while  $w_{\underline{y}hl}^i$  is positive only if  $\gamma < \frac{(1-q)(q(\bar{p}+\hat{p})-\hat{p})}{p(q(\bar{p}+\hat{p})-\bar{p})}$ . Moreover,  $w_{\underline{y}hh}^i > w_{\underline{y}hl}^i$  for all  $\gamma \in (0, 1]$ ;
- $w_{\underline{y}hh}^i$  is positive if and only if the individual performance indicators are informative enough, i.e.,  $q > \frac{1-\hat{p}}{2-\hat{p}-\bar{p}}$ .  $w_{\underline{y}hl}^i$  is positive if and only if reciprocity concerns are not too high, i.e.,  $\gamma < \frac{(1-q)(q(2-\hat{p}-\bar{p})-(1-\hat{p}))}{q(q(2-\bar{p}-\hat{p})-(1-\bar{p}))}$ .  $w_{\underline{y}hh}^i > w_{\underline{y}hl}^i$  for all  $\gamma \in (0, 1)$ ;
- $w_{\underline{y}lh}^i$  and  $w_{\underline{y}lh}^i$  are positive if and only if the total output is sufficiently informative compared to the individual performance indicators, i.e., respectively, when  $\bar{p} > \frac{\hat{p}(q^2+\gamma(1-q)^2)}{(1-q)q(1+\gamma)}$  and  $(1 - \bar{p}) < \frac{(1-\hat{p})q^2+\gamma(1-q)^2}{q(1-q)(1+\gamma)}$ .  $w_{\underline{y}lh}^i$
- $w_{\underline{y}ll}^i$  is positive if and only if  $q < \frac{\bar{p}}{\bar{p}+\hat{p}}$ ;
- $w_{\underline{y}ll}^i$  is never positive.

From Observation 10 it follows that a joint-incentive scheme is designed whenever the principal decides to induce positive reciprocity in equilibrium.

Rearranging conditions (A.21) to (A.28) for  $\rho = 0$  and then replacing for  $w_{\underline{y}hh}^i$  into  $i$ 's incentive compatibility constraint I achieve the following lemma:

**Lemma 9.** If  $\rho = 0$ , the best-joint incentive scheme takes the form of a team scheme in which:

$$w_{\underline{y}hh}^i = \frac{\phi}{q[q(\bar{p} + \hat{p}) - \hat{p}](1 + \gamma)}$$

and all the other payments are set equal to zero.  
The expected cost for the principal is given by:

$$E(W(BJG))_{\rho=0} = \frac{q\bar{p}\phi}{[q(\bar{p} + \hat{p}) - \hat{p}](1 + \gamma)}.$$

If in equilibrium the principal decides to induce negative reciprocity, the first-order condition associated with the principal's problem can be simplified as follows:

$$\frac{\partial L}{\partial w_{\underline{y}hh}^i} \geq 0 \Rightarrow \bar{p}q^2 - \mu[(\bar{p} + \hat{p})q - \hat{p}]q(1 - \gamma)(1 - \rho)(w_{\underline{y}hh}^i)^{-\rho} \geq 0 \quad (\text{A.29})$$

$$\frac{\partial L}{\partial w_{\underline{y}hh}^i} \geq 0 \Rightarrow (1 - \bar{p})q^2 - \mu q[q(1 - \bar{p}) - (1 - q)(1 - \hat{p})]q(1 - \gamma)(1 - \rho)(w_{\underline{y}hh}^i)^{-\rho} \geq 0 \quad (\text{A.30})$$

$$\frac{\partial L}{\partial w_{\underline{y}hl}^i} \geq 0 \Rightarrow \bar{p}q(1 - q) - \mu[(1 - q)[\bar{p}q - (1 - q)\hat{p}] - q\gamma[(1 - q)\bar{p} - q\hat{p}]](1 - \rho)(w_{\underline{y}hl}^i)^{-\rho} \geq 0 \quad (\text{A.31})$$

$$\frac{\partial L}{\partial w_{\underline{y}hl}^i} \geq 0 \Rightarrow (1 - \bar{p})q(1 - q) - \mu[(1 - q)[q(1 - \bar{p}) - (1 - q)(1 - \hat{p})] - \gamma q[(1 - q)(1 - \bar{p}) - q(1 - \hat{p})]](1 - \rho)(w_{\underline{y}hl}^i)^{-\rho} \geq 0 \quad (\text{A.32})$$

$$\frac{\partial L}{\partial w_{\underline{y}lh}^i} \geq 0 \Rightarrow \bar{p}(1 - q)q - \mu[q[(1 - q)\bar{p} - q\hat{p}] + \gamma(1 - q)[q\bar{p} - \hat{p}(1 - q)]](1 - \rho)(w_{\underline{y}lh}^i)^{-\rho} \geq 0 \quad (\text{A.33})$$

$$\frac{\partial L}{\partial w_{\underline{y}lh}^i} \geq 0 \Rightarrow (1 - \bar{p})(1 - q)q - \mu[q[(1 - q)(1 - \bar{p}) - q(1 - \hat{p})] + \gamma(1 - q)[(1 - \bar{p})q - (1 - q)(1 - \hat{p})]](1 - \rho)(w_{\underline{y}lh}^i)^{-\rho} \geq 0 \quad (\text{A.34})$$

$$\frac{\partial L}{\partial w_{\underline{y}ll}^i} \geq 0 \Rightarrow \bar{p}(1 - q)^2 - \mu[(1 - q)(\bar{p}(1 - q) - \hat{p}q)](1 - \gamma)(1 - \rho)(w_{\underline{y}ll}^i)^{-\rho} \geq 0 \quad (\text{A.35})$$

$$\frac{\partial L}{\partial w_{\underline{y}ll}^i} \geq 0 \Rightarrow (1 - \bar{p})(1 - q)^2 - \mu[(1 - q)[(1 - q)(1 - \bar{p}) - q(1 - \hat{p})]](1 - \gamma)(1 - \rho)(w_{\underline{y}ll}^i)^{-\rho} \geq 0. \quad (\text{A.36})$$

**Observation 11.** From conditions (A.29)-(A.36) I have that, for any  $\rho \in (0, 1)$ :

- $w_{\underline{y}hl}^i$  is always positive, whereas  $w_{\underline{y}hh}^i$  is positive whenever  $\gamma < 1$ . Moreover,  $w_{\underline{y}hh}^i > w_{\underline{y}hl}^i$  for all  $\gamma \in (0, 1]$ ;
- $w_{\underline{y}hh}^i$  is positive if and only if the individual performance indicators are informative enough, i.e.,  $q > \frac{1 - \hat{p}}{2 - \bar{p} - \hat{p}}$  and if  $\gamma < 1$ .  $w_{\underline{y}hl}^i$  is positive whenever  $q > \frac{(1 - \hat{p})}{2 - \bar{p} - \hat{p}}$ .  $w_{\underline{y}hh}^i > w_{\underline{y}hl}^i$  for all  $\gamma \in (0, 1)$ ;
- $w_{\underline{y}lh}^i$  and  $w_{\underline{y}ll}^i$  are positive only if the total output is sufficiently informative compared to the individual performance indicators.
- $w_{\underline{y}ll}^i$  is positive if and only if  $q < \frac{\bar{p}}{\bar{p} + \hat{p}}$  and if  $\gamma < 1$ ;
- $w_{\underline{y}ll}^i$  is never positive.

From Observation 11 it follows that a relative-incentive scheme is designed whenever the principal decides to induce negative reciprocity.

Rearranging conditions (A.29) to (A.36) for  $\rho = 0$  and then replacing for  $w_{\underline{y}hh}^i$  into  $i$ 's incentive compatibility constraint I achieve the following lemma:

**Lemma 10.** *If  $\rho = 0$ , the relative-joint incentive scheme takes the form of a tournament scheme in which:*

$$w_{\bar{y}hl}^i = \frac{\phi}{(1-q)[q(\bar{p} + \hat{p}) - \hat{p}] + 1\gamma[q(\bar{p} + \hat{p}) - \bar{p}]}$$

and all the other payments are set equal to zero.

The expected cost for the principal is given by:

$$E(W(BRG))_{\rho=0} = \frac{(1-q)q\bar{p}\phi}{(1-q)[q(\bar{p} + \hat{p}) - \hat{p}] + 1\gamma[q(\bar{p} + \hat{p}) - \bar{p}]}$$

By comparing  $E(W(BJG))_{\rho=0}$  and  $E(W(BRG))_{\rho=0}$  I can see that the former is always smaller than the latter whenever:

$$q^2 2(\bar{p} + \hat{p}) - q 2(\bar{p} + \hat{p}) + \hat{p} p < 0$$

while one root,  $q_1 = \frac{2(\bar{p} + \hat{p}) - \sqrt{4(\bar{p}^2 - \hat{p}^2)}}{4(\bar{p} + \hat{p})}$  is always smaller than  $1/2$ , the other one,  $q_2 = \frac{2(\bar{p} + \hat{p}) + \sqrt{4(\bar{p}^2 - \hat{p}^2)}}{4(\bar{p} + \hat{p})}$

takes always value in the interval  $(1/2, 1)$  since  $4(\bar{p}^2 - \hat{p}^2)$  takes always value on  $(0, 4[\bar{p} + \hat{p}]^2)$ .

Setting  $\tilde{q}(\bar{p}, \hat{p}) = q_2$  shows the first point of Proposition 4.

Moreover, it can be easily seen that:

$$\begin{aligned} \frac{\partial \tilde{q}}{\partial \bar{p}} &= \frac{\hat{p}}{2(\bar{p} + \hat{p})\sqrt{\bar{p}^2 - \hat{p}^2}} > 0 \\ \frac{\partial \tilde{q}}{\partial \hat{p}} &= -\frac{\bar{p}}{2(\bar{p} + \hat{p})\sqrt{\bar{p}^2 - \hat{p}^2}} < 0. \end{aligned}$$

that shows the second point of Proposition 4.

Consistently with the analysis of Section 3, when  $\rho$  increases above zero relative incentives become comparatively less attractive as they charge more risks on the agents. Hence the principal never finds it profitable to induce negative reciprocity when  $q \leq \tilde{q}$ .

Conversely, for  $q > \tilde{q}$  the usual interplay arises.

Note that  $q > \tilde{q}$  implies that  $q > \frac{1-\hat{p}}{2-\hat{p}-\bar{p}}$  and that  $q > \frac{\bar{p}}{\hat{p}+\bar{p}}$ . Therefore, from Observations 10 and 11 I know that the non-negativity constraints associated with  $w_{y_l h}^i, w_{y_l l}^i$  bite for all  $y \in \{\underline{y}, \bar{y}\}$  irrespective of the type of reciprocal exchange the principal wants to induce.

Let  $\tilde{\gamma}$  and  $\tilde{\gamma}'$  be equal to  $\frac{(1-q)(q(2-\hat{p}-\bar{p})-(1-\hat{p}))}{q(q(2-\hat{p}-\bar{p})-(1-\bar{p}))}$  and  $\frac{(1-q)(q(\bar{p}+\hat{p})-\hat{p})}{p(q(\bar{p}+\hat{p})-\bar{p})}$ , respectively. From Observation 10 I know that in the best-joint incentive scheme  $w_{yhl}^i$  is strictly positive only if  $\gamma < \tilde{\gamma}$  and  $w_{\bar{y}hl}^i$  is strictly positive only if  $\gamma < \tilde{\gamma}'$ . Therefore, I have to consider three sub-cases:

- If  $\gamma \in (0, \tilde{\gamma})$ , then from conditions (A.21) to (A.24) I find:

$$w_{\underline{y}hh}^i = \left( \frac{\bar{p}(q(2 - \bar{p} - \hat{p}) - (1 - \hat{p}))}{\underbrace{[(\bar{p} - \hat{p})q - \hat{p}](1 - \bar{p})}_{=X}} \right)^{\frac{1}{\rho}} w_{\bar{y}hh}^i \quad (\text{A.37})$$

$$w_{\bar{y}hl}^i = \left( \frac{[(1 - q)(\bar{p}q - (1 - q)\hat{p}) + q\gamma((1 - q)\bar{p} - \hat{p}q)]}{\underbrace{(1 + \gamma)(1 - q)[(\bar{p} - \hat{p})q - \hat{p}]}_{=Y}} \right)^{\frac{1}{\rho}} w_{\bar{y}hh}^i \quad (\text{A.38})$$

$$w_{\underline{y}hl}^i = \left( \frac{\bar{p}[(1 - q)(q(1 - \bar{p}) - (1 - q)(1 - \hat{p}) + \gamma q((1 - q)(1 - \bar{p}) - q(1 - \hat{p})))]}{\underbrace{(1 + \gamma)(1 - \bar{p})(1 - q)[(\bar{p} - \hat{p})q - \hat{p}]}_{=Z}} \right)^{\frac{1}{\rho}} w_{\bar{y}hh}^i \quad (\text{A.39})$$

Replacing into the (ICi) leads to:

$$w_{\bar{y}hh}^i = \left( \frac{\phi}{[m + nX^{\frac{1-\rho}{\rho}}](1 + \gamma)q + oY^{\frac{1-\rho}{\rho}} + rZ^{\frac{1-\rho}{\rho}}} \right)^{\frac{1}{1-\rho}}$$

where:

$$\begin{aligned} m &= [q(\bar{p} + \hat{p}) - \hat{p}] \\ n &= [q(1 - \hat{p}) - (1 - q)(1 - \bar{p})] \\ o &= (1 - q)[(\bar{p}q - (1 - q)\hat{p}) + q\gamma((1 - q)\bar{p} - \hat{p}q)] \\ r &= (1 - q)[(1 - \bar{p})q - (1 - q)(1 - \hat{p}) + \gamma q((1 - q)(1 - \bar{p}) - q(1 - \hat{p}))]. \end{aligned}$$

The other non-negative payments can be achieved by replacing  $w_{\bar{y}hh}^i$  into equations (A.37), (A.39) and (A.39).

- If  $\gamma \in [\tilde{\gamma}, \tilde{\gamma}')$ , then I obtain:

$$w_{\bar{y}hh}^i = \left( \frac{\phi}{[m + nX^{\frac{1-\rho}{\rho}}](1 + \gamma)q + oY^{\frac{1-\rho}{\rho}}} \right)^{\frac{1}{1-\rho}}$$

and

$$\begin{aligned} w_{\underline{y}hh}^i &= \left( \frac{\bar{p}(q(2 - \bar{p} - \hat{p}) - (1 - \hat{p}))}{\underbrace{[(\bar{p} - \hat{p})q - \hat{p}](1 - \bar{p})}_{=X}} \right)^{\frac{1}{\rho}} w_{\bar{y}hh}^i \\ w_{\bar{y}hl}^i &= \left( \frac{[(1 - q)(\bar{p}q - (1 - q)\hat{p}) + q\gamma((1 - q)\bar{p} - \hat{p}q)]}{\underbrace{(1 + \gamma)(1 - q)[(\bar{p} - \hat{p})q - \hat{p}]}_{=Y}} \right)^{\frac{1}{\rho}} w_{\bar{y}hh}^i \\ w_{\underline{y}hl}^i &= 0 \end{aligned}$$

- Finally, if  $\gamma \in [\tilde{\gamma}', 1]$ , then I achieve:

$$w_{\underline{y}hh}^i = \left( \frac{\phi}{[m + nX^{\frac{1-\rho}{\rho}}](1+\gamma)q} \right)^{\frac{1}{1-\rho}}$$

and

$$w_{\underline{y}hh}^i = \left( \frac{\bar{p}(q(2-\bar{p}-\hat{p})-(1-\hat{p}))}{\underbrace{[(\bar{p}-\hat{p})q-\hat{p}](1-\bar{p})}_{=X}} \right)^{\frac{1}{\rho}} w_{\underline{y}hl}^i$$

$$w_{\underline{y}hl}^i = 0$$

$$w_{\underline{y}hl}^i = 0$$

I then consider the best-relative incentive contact. From conditions (A.29) and (A.30) I know that in the best relative-incentive scheme  $w_{\underline{y}hh}^i$  and  $w_{\underline{h}h}^i$  are strictly positive only if  $\gamma < 1$ . Therefore, I have to consider two sub-cases:

- If  $\gamma \in (0, 1)$ , then from conditions (A.29) to (A.32) I find:

$$w_{\underline{y}hh}^i = \left( \frac{(1-q)(q(\bar{p}+\hat{p})-\hat{p})(1-\gamma)}{\underbrace{[(1-q)(\bar{p}+\hat{p})q-\bar{p}+\gamma q(q(\bar{p}+\hat{p})-\hat{p})]}_{=L}} \right)^{\frac{1}{\rho}} w_{\underline{y}hl}^i \quad (\text{A.40})$$

$$w_{\underline{y}hh}^i = \left( \frac{(1-q)(1-\gamma)\bar{p}(q(2-\bar{p}-\hat{p})-(1-\hat{p}))}{(1-\bar{p})\underbrace{[(1-q)(\bar{p}+\hat{p})q-\bar{p}+\gamma q(q(\bar{p}+\hat{p})-\hat{p})]}_{=M}} \right)^{\frac{1}{\rho}} w_{\underline{y}hl}^i \quad (\text{A.41})$$

$$w_{\underline{y}hl}^i = \left( \frac{\bar{p}((1-q)((2-\bar{p}-\hat{p})-(1-\hat{p}))+q\gamma(q(2-\bar{p}-\hat{p})-(1-\bar{p})))}{(1-\bar{p})\underbrace{[(1-q)(\bar{p}+\hat{p})q-\bar{p}+\gamma q(q(\bar{p}+\hat{p})-\hat{p})]}_{=N}} \right)^{\frac{1}{\rho}} w_{\underline{y}hl}^i \quad (\text{A.42})$$

Replacing into the (ICi) leads to:

$$w_{\underline{y}hl}^i = \left( \frac{\phi}{[sM^{\frac{1-\rho}{\rho}} + tL^{\frac{1-\rho}{\rho}}](1-\gamma)q + u + vN^{\frac{1-\rho}{\rho}}} \right)^{\frac{1}{1-\rho}}$$

where:

$$s = [q(\bar{p}+\hat{p})-\hat{p}]$$

$$t = [q(1-\hat{p})-(1-q)(1-\bar{p})]$$

$$u = (1-q)[((\bar{p}q-(1-q)\hat{p})-q\gamma((1-q)\bar{p}-\hat{p}q))]$$

$$v = (1-q)[(1-\bar{p})q-(1-q)(1-\hat{p})-\gamma q((1-q)(1-\bar{p})-q(1-\hat{p}))].$$

The other non-negative payments can be achieved from (A.40) to (A.42).

- If  $\gamma = 1$ , then:

$$w_{\underline{y}hl}^i = \left( \frac{\phi}{u + vN^{\frac{1-\rho}{\rho}}} \right)^{\frac{1}{1-\rho}}$$

and

$$w_{\underline{y}hh}^i = 0$$

$$w_{\underline{y}hh}^i = 0$$

$$w_{\underline{y}hl}^i = \left( \underbrace{\frac{\bar{p}((1-q)((2-\bar{p}-\hat{p})-(1-\hat{p}))+q\gamma(q(2-\bar{p}-\hat{p})-(1-\bar{p})))}{(1-\bar{p})[(1-q)(\bar{p}+\hat{p})q-\bar{p}+\gamma q(q(\bar{p}+\hat{p})-\hat{p})]}_{=N}} \right)^{\frac{1}{\rho}} w_{\underline{y}hl}^i$$

Comparing the expected costs entailed by the best-joint and the best-relative incentive schemes in the three possible subcases, it is possible to show the existence of  $\bar{\rho}$  through numerical simulations.  $\square$

## Appendix B

In Section 6 I referred to the possibility of extending the model to incorporate asymmetric preferences for the two sides of reciprocity. Contributions in economics and social psychology suggest that preferences for positive reciprocity might be stronger than for negative reciprocity (See Perugini et al. 2003 and Dohmen et al. 2008, among the others). Moreover, the distribution of positive and negative reciprocators across countries seems to be different among countries. To account for this, let us modify the definition of the psychological payoff as follows:

$$\hat{r}_i(\lambda_{iji}, k_{ij}) = \begin{cases} r_i(\lambda_{iji}, k_{ij}) & \text{if } \lambda_{iji} > 0 \\ \alpha r_i(\lambda_{iji}, k_{ij}) & \text{if } \lambda_{iji} < 0 \end{cases}$$

for some  $\alpha \in [0, 1)$ . The lower  $\alpha$ , the less willing an agent is to bear some material costs to reciprocate a mischief. Under the assumptions of the baseline model, the total expected cost entailed by the best-joint and best-relative incentive schemes are:

$$E(W(BJ)) = \begin{cases} 2 \left( q + (1-q) \left( \frac{1-q-q\gamma}{(1+\gamma)(1-q)} \right)^{\frac{1}{\rho}} \right) q \left( \frac{\phi}{(2q-1) \left( q(1+\gamma) + (1-q-\gamma q) \left( \frac{1-q-q\gamma}{(1+\gamma)(1-q)} \right)^{\frac{1-\rho}{\rho}} \right)} \right)^{\frac{1}{1-\rho}} & \text{if } \gamma \in [0, (1-q)/q); \\ 2q^2 \left( \frac{\phi}{(2q-1)(1+\gamma)q} \right)^{\frac{1}{1-\rho}} & \text{if } \gamma \in [(1-q)/q, 1]. \end{cases}$$

and

$$E(W(BR\alpha)) = \begin{cases} 2 \left( 1-q + q \left( \frac{(1-q)(1-\alpha\gamma)}{1-q+q\alpha\gamma} \right)^{\frac{1}{\rho}} \right) q \left( \frac{\phi}{(2q-1) \left( (1-q+\alpha\gamma q) + q(1-\alpha\gamma) \left( \frac{(1-\alpha\gamma)(1-q)}{1-q+q\alpha\gamma} \right)^{\frac{1-\rho}{\rho}} \right)} \right)^{\frac{1}{1-\rho}} & \text{if } \gamma \in (0, \alpha); \\ 2(1-q)q \left( \frac{\phi}{(2q-1)(1-q+q\alpha\gamma)} \right)^{\frac{1}{1-\rho}} & \text{if } \gamma \in [\alpha, 1]. \end{cases}$$

It can be seen that the expected cost of the best relative-incentive scheme considered in Appendix A represents the limiting case of the one here above for  $\alpha = 1$ . From a comparison similar to that carried out in Appendix A I achieve the following lemma:

**Lemma 11.** *There  $\exists$  a threshold value  $\alpha^*(q, \gamma) \in (0, 1)$  such that:*

- *if  $\alpha \in [0, \alpha^*(q, \gamma)]$ , then the optimal scheme always implements the best joint-incentive scheme irrespective of  $\rho$ ;*
- *if  $\alpha \in (\alpha^*(q, \gamma), 1)$ , then there exists a threshold value  $\tilde{\rho} < \rho^*$  such that principal prefers the best-relative incentive scheme if  $\rho < \tilde{\rho}$  and the best-joint incentive if  $\rho > \tilde{\rho}$ .*

**Proof.** The lemma follows from the fact that the total expected cost of implementing the best-relative incentive scheme is a perturbation of the one considered in Appendix A and from the continuity of  $E(W(BR\alpha))$  in  $\alpha$ .  $\square$

A lower  $\alpha$  makes the best-relative incentive scheme comparably less valuable to the principal. When  $\alpha$  is very low, i.e.,  $\alpha \leq \alpha^*$ , the best-relative incentive scheme is so penalized that the principal always prefers to set up the best joint-incentive scheme irrespective of the agents' risk-aversion. When  $\alpha$  is not too low,  $\alpha > \alpha^*$ , the usual interplay occurs.

## Bibliography

- Abeler, J., Altmann, S., Kube, S., and Wibrall, M. (2010). Gift exchange and workers' fairness concerns: When equality is unfair. *Journal of the European Economic Association*, 8(6):1299–1324.
- Akerlof, G. A. (1982). Labor Contracts as Partial Gift Exchange. *Quarterly Journal of Economics*, 97(4):543–69.
- Backes-Gellner, U. and Pull, K. (2013). Tournament compensation systems, employee heterogeneity, and firm performance. *Human Resource Management*, 52(3):375–398.
- Bandiera, O., Barankay, I., and Rasul, I. (2005). Social preferences and the response to incentives: Evidence from personnel data. *Quarterly Journal of Economics*, pages 917–962.
- Banks, G. C., Batchelor, J. H., Seers, A., O'Boyle, E. H., Pollack, J. M., and Gower, K. (2014). What does team-member exchange bring to the party? a meta-analytic review of team and leader social exchange. *Journal of Organizational Behavior*, 35(2):273–295.
- Barron, J. M. and Gjerde, K. P. (1997). Peer pressure in an agency relationship. *Journal of Labor Economics*, pages 234–254.
- Bartling, B. (2011). Relative performance or team evaluation? optimal contracts for other-regarding agents. *Journal of Economic Behavior & Organization*, 79(3):183 – 193.
- Bartling, B. and von Siemens, F. A. (2010). The intensity of incentives in firms and markets: Moral hazard with envious agents. *Labour Economics*, 17(3):598 – 607.
- Bartling, B. and von Siemens, F. A. (2011). Wage inequality and team production: An experimental analysis. *Journal of Economic Psychology*, 32(1):1–16.
- Battigalli, P. and Dufwenberg, M. (2009). Dynamic psychological games. *Journal of Economic Theory*, 144(1):1–35.
- Bellemare, C. and Shearer, B. (2009). Gift giving and worker productivity: Evidence from a firm-level experiment. *Games and Economic Behavior*, 67(1):233–244.
- Bellemare, C. and Shearer, B. (2011). On The Relevance And Composition Of Gifts Within The Firm: Evidence From Field Experiments. *International Economic Review*, 52(3):855–882.
- Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1):122–142.
- Berger, L., Klassen, K. J., Libby, T., and Webb, A. (2013). Complacency and giving up across repeated tournaments: Evidence from the field. *Journal of Management Accounting Research*, 25(1):143–167.
- Bloom, N., Kretschmer, T., and Van Reenen, J. (2011). Are family-friendly workplace practices a valuable firm resource? *Strategic Management Journal*, 32(4):343–367.
- Bloom, N. and Van Reenen, J. (2010). Why do management practices differ across firms and countries? *Journal of Economic Perspectives*, 24(1):203–24.
- Bolton, P. and Dewatripont, M. (2005). *Contract theory*. MIT press.
- Bosse, D. A., Phillips, R. A., and Harrison, J. S. (2009). Stakeholders, reciprocity, and firm performance. *Strategic Management Journal*, 30(4):447–456.
- Bowling, N. A., Beehr, T. A., Johnson, A. L., Semmer, N. K., Hendricks, E. A., and Webster, H. A. (2004). Explaining potential antecedents of workplace social support: reciprocity or attractiveness? *Journal of Occupational Health Psychology*, 9(4):339.
- Brown, M., Falk, A., and Fehr, E. (2004). Relational contracts and the nature of market interactions. *Econometrica*, pages 747–780.
- Bryan, S. H., Nash, R. C., and Patel, A. (2014). Culture and ceo compensation. In *AFA 2013 San Diego Meetings Paper*.
- Cerdin, J.-L. and Pargneux, M. L. (2009). Career and international assignment fit: Toward an integrative model of success. *Human Resource Management*, 48(1):5–25.
- Charness, G. and Kuhn, P. (2007). Does pay inequality affect worker effort? experimental evidence. *Journal of Labor Economics*, 25(4):693–723.
- Che, Y.-K. and Yoo, S.-W. (2001). Optimal incentives for teams. *American Economic Review*, 91(3):525–541.
- Chen, C. X., Williamson, M. G., and Zhou, F. H. (2012). Reward system design and group creativity: An experimental investigation. *Accounting Review*, 87(6):1885–1911.
- Canyon, M. J. and Murphy, K. J. (2000). The prince and the pauper? ceo pay in the united states and united kingdom. *Economic Journal*, 110(467):640–671.



- Cropanzano, R. and Mitchell, M. S. (2005). Social exchange theory: An interdisciplinary review. *Journal of management*, 31(6):874–900.
- DellaVigna, S. (2009). Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*, 47(2):315–72.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2008). Representative trust and reciprocity: Prevalence and determinants. *Economic Inquiry*, 46(1):84–90.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2009). Homo reciprocans: Survey evidence on behavioural outcomes. *Economic Journal*, 119(536):592–612.
- Dufwenberg, M. and Kirchsteiger, G. (2000). Reciprocity and wage undercutting. *European Economic Review*, 44(4-6):1069–1078.
- Dufwenberg, M. and Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2):268–298.
- Dur, R. and Glazer, A. (2008). Optimal contracts when a worker envies his boss. *Journal of Law, Economics, and Organization*, pages 120–137.
- Dur, R. and Sol, J. (2010). Social interaction, co-worker altruism, and incentives. *Games and Economic Behavior*, 69(2):293–301.
- Englmaier, F., Kolaska, T., and Leider, S. (2015). Reciprocity in organisations-evidence from the wers.
- Englmaier, F. and Leider, S. (2012). Contractual and Organizational Structure with Reciprocal Agents. *American Economic Journal: Microeconomics*, 4(2):146–83.
- Englmaier, F., Strasser, S., and Winter, J. (2014). Worker characteristics and wage differentials: Evidence from a gift-exchange experiment. *Journal of Economic Behavior & Organization*, 97(C):185–203.
- Englmaier, F. and Wambach, A. (2010). Optimal incentive contracts under inequity aversion. *Games and Economic Behavior*, 69(2):312–328.
- Falk, A. (2007). Gift exchange in the field. *Econometrica*, 75(5):1501–1511.
- Falk, A. and Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2):293–315.
- Fehr, E. and Fischbacher, U. (2002). Why social preferences matter—the impact of non-selfish motives on competition, cooperation and incentives. *Economic Journal*, 112(478):C1–C33.
- Fehr, E., Kirchsteiger, G., and Riedl, A. (1993). Does Fairness Prevent Market Clearing? An Experimental Investigation. *Quarterly Journal of Economics*, 108(2):437–59.
- Fehr, E. and Schmidt, K. M. (2003). Theories of Fairness and Reciprocity - Evidence and Economic Applications. Technical report.
- Geanakoplos, J., Pearce, D., and Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior*, 1(1):60–79.
- Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, 25(2):pp. 161–178.
- Grote, R. C. and Grote, D. (2011). *How to be good at performance appraisals: Simple, effective, done right*. Harvard Business Press.
- Hamilton, H., Nickerson, J. A., Owan, H., Pollack, T. D. R., Koenker, R., Fitzenberger, B., Rao, A., Macdonald, G., and Levitt, S. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, pages 465–497.
- Hartog, J., Ferrer-i Carbonell, A., and Jonker, N. (2002). Linking measured risk aversion to individual characteristics. *Kyklos*, 55(1):3–26.
- Holmström, B. (1979). Moral Hazard and Observability. *Bell Journal of Economics*, 10(1):74–91.
- Holmström, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, pages 324–340.
- Holmström, B. and Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, pages 24–52.
- Indjejikian, R. J. and Nanda, D. (2002). Executive target bonuses and what they imply about performance standards. *Accounting Review*, 77(4):793–819.
- Itoh, H. (1991). Incentives to help in multi-agent situations. *Econometrica*, 59(3):pp. 611–636.
- Itoh, H. (2004). Moral hazard and other-regarding preferences. *Japanese Economic Review*, 55(1):18–45.
- Kandel, E. and Lazear, E. P. (1992). Peer Pressure and Partnerships. *Journal of Political Economy*, 100(4):801–17.
- Kempf, A. and Ruenzi, S. (2008). Tournaments in mutual-fund families. *Review of Financial Studies*, 21(2):1013–1036.
- Kim, J. (2015). Managerial beliefs and incentive policies. *Journal of Economic Behavior & Organization*, 119:84–95.
- Kosfeld, M. and von Siemens, F. A. (2011). Competition, cooperation, and corporate culture. *The RAND Journal of Economics*, 42(1):23–43.
- Köszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature*, 52(4):1075–1118.
- Kruse, D. L., Blasi, J. R., and Park, R. (2010). Shared capitalism in the us economy: Prevalence, characteristics, and employee views of financial participation in enterprises. In *Shared Capitalism at Work: Employee Ownership, Profit and Gain Sharing, and Broad-based Stock Options*, pages 41–75. University of Chicago Press.
- Lazear, E. P. (1989). Pay equality and industrial politics. *Journal of Political Economy*, 97(3):pp. 561–580.
- Lerner, J. S. and Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, 81(1):146.
- Liden, R. C., Sparrowe, R. T., and Wayne, S. J. (1997). Leader-member exchange theory: The past and potential for the future. *Research in personnel and human resources management*, 15:47–120.
- Livio, L. (2015). Lonesome riders or team players? organizing production with reciprocal agents. Technical report.
- Maximiano, S., Sloof, R., and Sonnemans, J. (2007). Gift exchange in a multi-worker firm. *Economic Journal*, 117(522):1025–1050.
- Merchant, K. and Oley, D. (2007). A review of the literature on control and accountability. *I Handbook of Management Accounting Research*.
- Mikula, G., Scherer, K. R., and Athenstaedt, U. (1998). The role of injustice in the elicitation of differential emotional reactions. *Personality and Social Psychology Bulletin*, 24(7):769–783.
- Milgrom, P. and Roberts, J. (1994). Economics, organization and management/p. milgrom, j. roberts.
- Mookherjee, D. (1984). Optimal incentive schemes with many agents. *Review of Economic Studies*, 51(3):433–446.
- Murphy, K. J. (1999). Executive compensation. *Handbook of Labor Economics*, 3:2485–2563.
- Non, A. (2012). Gift-exchange, incentives, and heterogeneous workers. *Games and Economic Behavior*, 75(1):319 – 336.

- Oyer, P. and Schaefer, S. (2011). Personnel economics: Hiring and incentives. *Handbook of Labor Economics*, 4:1769–1823.
- Pazy, A. and Ganzach, Y. (2008). Pay contingency and the effects of perceived organizational and supervisor support on performance and commitment. *Journal of Management*.
- Perugini, M., Gallucci, M., Presaghi, F., and Ercolani, A. P. (2003). The personal norm of reciprocity. *European Journal of Personality*, 17(4):251–283.
- Petty, M., Singleton, B., and Connell, D. W. (1992). An experimental evaluation of an organizational incentive plan in the electric utility industry. *Journal of Applied Psychology*, 77(4):427.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, 83(5):pp. 1281–1302.
- Rey-Biel, P. (2008). Inequity aversion and team incentives. *Scandinavian Journal of Economics*, 110(2):297–320.
- Rotemberg, J. J. (1994). Human Relations in the Workplace. *Journal of Political Economy*, 102(4):684–717.
- Rotemberg, J. J. and Saloner, G. (1993). Leadership Style and Incentives. *Management Science*, 39(11):1299–1318.
- Sebald, A. (2010). Attribution and reciprocity. *Games and Economic Behavior*, 68(1):339–352.
- Sebald, A. and Walzl, M. (2014). Subjective Performance Evaluations and Reciprocity in Principal-Agent Relations. *Scandinavian Journal of Economics*, 116(2):570–590.
- Sebald, A. and Walzl, M. (2015). Optimal contracts based on subjective performance evaluations and reciprocity. *Journal of Economic Psychology*, 47(C):62–76.
- Tosi, H. L. and Greckhamer, T. (2004). Culture and ceo compensation. *Organization Science*, 15(6):657–670.
- von Siemens, F. A. (2013). Intention-based reciprocity and the hidden costs of control. *Journal of Economic Behavior & Organization*, 92:55–65.