

# Managing Social Comparison Costs in Organizations

Oscar F. Contreras and Giorgio Zanarone\*

March, 2017

## Abstract

This paper studies how organizations manage the social comparison costs that arise when their members enjoy different status. We argue that social comparisons are more acute when status differences are formalized, and hence potentially more visible. We show that when an organization has tight relationships with its members, it can manage social comparison costs by adopting a homogeneous formal governance structure, while efficiently customizing the terms of employment through self-enforcing informal adjustments. We discuss the implications of our model for compensation policy, organization design, and firm boundaries.

**Keywords:** Social Comparisons, Organization Design, Relational Contracts, Formal Contracts.

**JEL Classification:** D03, D23, M52, M54.

---

\* CUNEF. C/ Leonardo Prieto Castro 2, Ciudad Universitaria - Madrid, 28040, Spain. E-mail: (Contreras) ocontreras@cunef.edu; (Zanarone) gzanarone@cunef.edu. We thank Florian Englmaier, Matthias Fahn, Guido Friebel, Willie Fuchs, Maija Halonen, Bentley MacLeod, Alan Schwartz, Ingo Weller, and audiences at Munich University, the SIOE conference, and the second Workshop on Relational Contracts, for useful comments. This study received financial support from the Spanish Ministry of Economy and Competitiveness through grant ECO2014-57131-R.

# 1. Introduction

Organizations are riddled with social comparisons: employees dislike having a lower status than their peers, and press for disparities to be removed. Examples of social comparisons, and their effects on employee satisfaction and firm policies, abound in business history. For instance, Nickerson and Zenger (2008) report that faculty at a North American business school opposed overload compensation for “star” professors, despite knowing that those professors received extra income from external activities. In a field experiment conducted by Cohn et al. (2014) at a German service firm, salespeople reacted to a salary cut by reducing their productivity, but much more so when the cut was unequal—that is, when it did not affect the rest of the sales team.<sup>1</sup> Finally, Baker, Gibbs and Holmstrom (1994) report evidence of wage compression policies, possibly to avoid social comparisons, at a US service firm. In particular, they show that managers in the higher wage quartiles had received systematically lower salary raises than their peers in lower quartiles.

In this paper, we develop a formal model to study how organizations manage *social comparison costs*, such as those described above, through their choice of governance structure. Our analysis is based on three building blocks. First, we argue that social comparisons in an organization are especially acute when the members’ relative status—in terms of pay and job position—is made *explicit*. This is consistent with Card et al. (2012) and Ockenfels et al. (2015), who show evidence that the satisfaction of employees decreases when they are explicitly informed that their compensation is lower than that of (some) peers.<sup>2</sup> Second, we argue that

---

<sup>1</sup> There are numerous empirical studies documenting a negative effect of pay inequality on employee satisfaction and performance, including recent work by Card et al. (2012), Ockenfels et al. (2015), and Breza et al. (2015). See Shaw (2014) for an interdisciplinary literature review.

<sup>2</sup> Card et al. (2012) study changes in the job satisfaction of faculty members at the University of California after the faculty is informed that a local online newspaper publishes compensation data for all university employees. Ockenfels et al. (2015) study

relative status is more explicit when each member's pay and job position are formalized in a contract, compared to the case where they are implicitly agreed. Formal contract terms may be more easily disclosed or exchanged, and may also become public as a byproduct of litigation or mandatory disclosure rules, such as those that apply to executive compensation in the U.S. (Gillan et al., 2009). Third, and following a well-established literature, we argue that formal contracts are more credible than informal ones because they can be enforced by third parties, such as courts, whereas informal contracts must be self-enforcing (see MacLeod, 2007, Malcomson, 2013, and Gil and Zananone, 2015, 2017, for up-to-date reviews of the theoretical and empirical literatures on informal contracting). Given these premises, organizations face a potential tradeoff between committing to formal terms of employment and avoiding social comparison costs by using less credible informal terms.

To analyze this trade-off, we model a simple organization that consists of a principal and two agents, each performing a contractible task in exchange for compensation. The agents are equally productive but have different outside options—for instance, because of varying personal constraints or firm-specific skills—and hence command different contract terms. However, explicit inequalities in compensation or task assignment may cause the low status agent to suffer disutility (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000) and to retaliate against the organization (Rabin, 1993; Hart and Moore, 2008).

We begin by studying a “spot organization” where the principal has arm's-length relationships with the two agents, and hence can only commit to task assignment and compensation through a formal contract. We show that the human resource policy of a spot

---

the link between relative compensation and the satisfaction of a German multinational's managers in German units, where local regulations impose pay transparency, versus U.S. units, where pay transparency is not required.

organization involves no distortion in job design, but exhibits *upwards wage compression*: the principal overpays the agent with low bargaining power, relative to her outside option, in order to retain her and to avoid retaliation.

Next, we analyze a “relational organization” where the principal expects to interact repeatedly with the two agents, and hence self-enforcing implicit agreements, as analyzed by Bull (1987), MacLeod and Malcolmson (1989), and Levin (2003), are potentially feasible. We show that if the relationship between the principal and the two agents is sufficiently tight, the optimal governance structure consists of a combination of formal and informal contract terms, and completely eliminates social comparison costs. The principal grants identical status to the two agents by assigning them the same job title and the same formal base salary, which is chosen to match the outside option of the agent with low bargaining power. Informally, the principal adjusts the salary of the agent with high bargaining power upwards through an implicit bonus, which is not communicated to the other agent and hence does not generate social comparisons.

Our model has counterintuitive implications for organizational design and policy. First, it predicts that as the organization moves from spot to relational—for instance, because its members expect to be in a long-term collaboration, or because they operate in a high-trust cultural environment—*formal* wage compression should *increase*, whereas *actual* wage compression should *decrease*.

Second, the model predicts that as the organization moves from spot towards relational, it is less likely to distort its internal architecture and boundaries as a means to manage social comparison costs. On one hand, we show that as argued by Nickerson and Zenger (2008), spot organizations may separate employees, and hence reduce social comparisons among them, by inefficiently splitting them among different departments, or by outsourcing to external partners

activities that would be optimally operated in-house. On the other hand, we show that relational organizations do not need to resort to these distortionary policies, because they can use homogeneous formal contract terms to eliminate social comparisons, while relying on informal private agreements to optimally differentiate among employees. Relatedly, we find that in contrast with the argument proposed by Nickerson and Zenger (2008), an organization's use of distortionary policies does not necessarily increase in the propensity of its members to engage in social comparisons. In fact, at intermediate levels the distortions *decrease* in such propensity because envy between the agents makes falling back on the spot organization less attractive for the principal, and hence increases his incentive to pay the informally agreed compensation.

Finally, our model provides a rationale for pay secrecy norms. While at times criticized (e.g., Futrell, 1978; Burkus, 2016), pay secrecy appears resilient, especially in the U.S. (Edwards, 2005; Hill, 2016; Ockenfels et al., 2015). In our model, secrecy in informal compensation adjustments, combined with homogeneity in formal base salaries, allows an organization to optimally customize pay without triggering potentially disruptive social comparisons among its employees.<sup>3</sup>

The rest of the paper is organized as follows. Section 2 discusses our contributions to the literature on organizational design. Section 3 presents our baseline model of social comparison costs in organizations. Section 4 analyzes social comparison costs in a spot organization. Section 5 analyzes social comparison costs in a relational organization. Section 6 analyzes how social comparisons affect the organization's choice of compensation policy, firm boundaries and internal architecture. Section 7 concludes.

---

<sup>3</sup> For an alternative explanation of pay secrecy as a device to reduce labor mobility, see Danziger and Katz (1997).

## 2. Relation to the Literature

Our paper belongs to a small but growing literature in organizational economics and strategy, which analyzes how non-standard preferences and “fairness” concerns affect organizations. In a seminal paper, Hart and Moore (2008) argue that incomplete formal contracts serve as reference points for what the parties can expect to bargain, thus limiting frustration and conflict in the relationship. Hart and Moore (2008) use their model to analyze the tradeoff between rigid and flexible pricing terms and the optimal allocation of authority in employment contracts. In subsequent papers, Hart and coauthors build on the reference point idea to analyze asset ownership (Hart, 2009), firm scope (Hart and Holmstrom, 2010), and the optimal degree of contractual incompleteness (Halonen and Hart, 2013). Our model differs from this literature in two important ways. First, we study a different rationale for formal contracts—namely, homogenizing the perceived relative status of an organizations’ members. Second, and most important, we explore the interaction between formal and informal contract terms in managing social comparison costs.<sup>4</sup>

In the strategic management literature, Zanarone et al. (2015) analyze a model where suppliers derive satisfaction from punishing uncompromising clients, and thus can credibly threaten to reveal confidential information on their clients to negotiate price increases. Zanarone et al. (2015) use their model to study how fixed price contracts and information disclosure policies may be used to discourage excess information acquisition. More related to our paper, Nickerson and Zenger (2008) argue that social comparison costs are more likely to arise within firms than between, and study how the formal governance and the boundaries of firms may be

---

<sup>4</sup> See Fehr et al. (2015) for some experimental evidence on the interaction between formal and informal contracts in shaping reference points.

chosen to mitigate such costs. We contribute to their important insight by studying the complementary role of formal and informal contracts in managing social comparison costs, and methodologically, by embedding our analysis into a formal model that allows us to precisely identify the mechanisms underlying such complementarity.

A related literature analyzes incentive contracts in the presence of fairness concerns, both when performance is verifiable (e.g., Englmaier and Wambach, 2010; Englmaier and Leider, 2012) and when it is non-verifiable (e.g., Kragl and Schmid, 2009; Bartling and von Siemens, 2010; Kragl, 2015). One insight from this literature is that fairness concerns have an ambiguous effect on incentive compensation, as they may induce low performers to restore equity by increasing their effort. Consistent with this line of research, Bradler et al. (2016) show evidence that rewarding high performers increases the productivity of low performers. Unlike our paper, this literature does not explore the role of contracts as mechanisms to manage social comparison costs in organizations.

Our paper also relates to the literature on the interaction between formal and informal contracts (e.g., Klein, 2000; Poppo and Zenger, 2002; Baker, Gibbons, and Murphy, 1994, 2011; Battigalli and Maggi, 2008; Kvaløy and Olsen, 2009; Ryall and Sampson, 2009; Zanarone, 2013).<sup>5</sup> In this literature, the parties rely on informal agreements because writing, verifying or enforcing formal contracts is costly. At the same time, the parties may want to use some formal contract terms to realign incentives and reduce renegeing temptations. We add to this literature in that we explore a novel reason for the joint use of formal and informal contracts—namely, managing social comparison costs—which *does not* depend on whether contract terms are verifiable.

---

<sup>5</sup> See Gil and Zanarone (2015) for a discussion of the predictive power and empirical relevance of this literature.

Finally, our paper relates to the literature on compensation policy and pay compression in organizations. Akerlof and Yellen (1990) analyze a model in which workers reduce effort when they perceive that they have received an unfair wage. MacLeod (2003) shows that when the performance of employees is subjectively evaluated, compensation is more compressed than when it is objectively evaluated because employees retaliate against the organization when they receive negative evaluations. Consistent with these theoretical works, empirical studies and case studies, such as Baker, Gibbs, and Holmstrom (1994), and Hall (2000), document wage compression policies in firms. We contribute to this literature by differentiating between formal and informal pay compression. Relatedly, we show that formal pay compression—and more generally, homogeneity in contract terms across the organization’s members—may be used to decrease social comparison costs without necessarily translating into actual compression and homogeneity.

### **3. The Model**

#### *3.1 Setup*

We consider an organization with one principal and two agents who have the possibility to interact at dates  $t = 1, \dots, \infty$ . We may interpret the principal and the two agents as an employer and his employees, as a franchisor and his franchisees, or alternatively, as a manufacturer and his suppliers. We assume that all parties are risk-neutral and discount future utilities using a common discount factor  $\delta \in (0,1)$ .



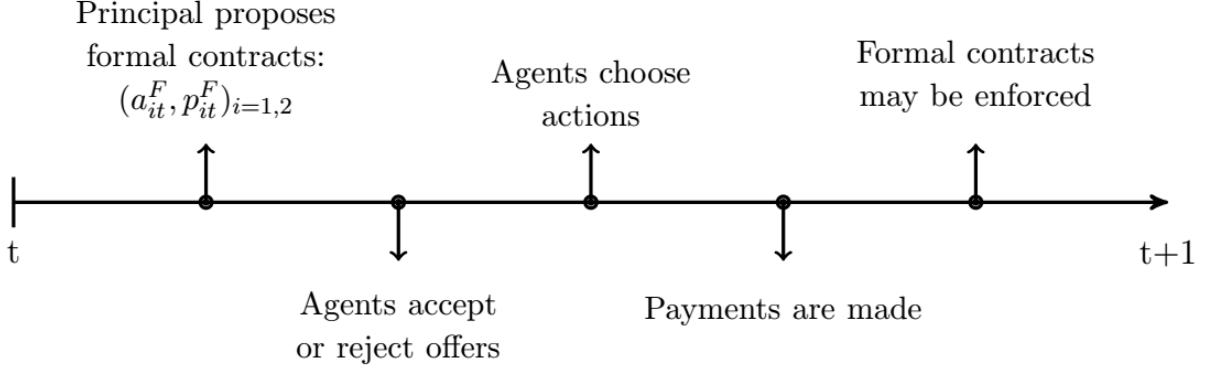


Figure 1: Timeline

The sequence of events within each period is illustrated in Figure 1. At the beginning of time  $t$ , the principal offers each of the agents a formal contract specifying an action  $a_{it}^F \in A \subset \mathbb{R}$  that agent  $i \in \{0,1\}$  is supposed to take and a payment  $p_{it}^F \in \mathbb{R}$  that the principal is supposed to make. Through the paper, we assume that actions and payments are both contractible, so that the formal contract can be enforced by a court of law.

After receiving the principal's offer, each agent decides independently whether to accept it. Let  $d_{it} \in \{0,1\}$  denote agent  $i$ 's decision. In the event of acceptance ( $d_{it} = 1$ ), agent  $i$  chooses an action  $a_{it}$  from the set  $A \subset \mathbb{R}$ , generating a direct benefit of  $\pi_A(a_{it})$  for himself, and a direct benefit of  $\pi_P(a_{it})$  for the principal. The total surplus generated by action  $a_{it}$  is thus given by  $S(a_{it}) \equiv \pi_P(a_{it}) + \pi_A(a_{it})$ . We assume that  $a^{FB} \equiv \operatorname{argmax}_{a \in A} S(a)$  exists, but otherwise impose no restrictions on the way  $\pi_P(\cdot)$  or  $\pi_A(\cdot)$  vary with  $a_{it}$ , allowing our model to capture a wide range of environments. For instance,  $\pi_P(\cdot), \pi'_P(\cdot) > 0$  and  $\pi_A(\cdot) < 0, \pi'_A(\cdot) < 0$  would capture the standard agency conflict between a principal and an agent who must perform costly task  $a_i$ . Alternatively,  $\pi_P(\cdot), \pi'_P(\cdot) > 0$  and  $\pi_A(\cdot)$  concave would capture more nuanced contractual frictions, such as that between a franchisor (principal) and two franchisees (agents) over the level of store-refurbishing and local advertising investment, or that between a

company's CEO (principal) and two divisional directors (agents) over the level of product coordination.

At the end of the period, the parties choose whether and what payments to make and if someone failed to perform according to the formal contract, whether to have it enforced. Throughout the paper, we assume that neither the principal nor the agents face liquidity constraints.

If agent  $i$  rejects the principal's offers ( $d_{it} = 0$ ), he receives an outside option with value  $\underline{U}_i$  and the principal receives a payoff of zero. We assume that  $S(a^{FB}) > \underline{U}_i$  for  $i \in \{1,2\}$ , so that the agents' outside options are less efficient than the joint-surplus maximizing action. This assumption is standard in the literature, and without it, it would never be optimal for the principal to employ any agent. We make two additional assumptions on the agents' outside options.

**Assumption 1.** The outside option of agent 1 is strictly greater than the outside option of agent 2—that is,  $\underline{U}_1 > \underline{U}_2$ .

**Assumption 2.** No action yields enough direct benefits to compensate an agent for losing his outside option—that is,  $\pi_A(a) < \min\{\underline{U}_1, \underline{U}_2\}$  for any  $a \in A$ .

Assumption 1 states that although the agents are equally productive within the organization, they face different opportunity costs of being employed by the principal, maybe because of varying personal constraints or skills in alternative jobs. This would be the case, for example, if the agents must move to a different city to enjoy their outside options and they face different relocation costs because of their personal situations (married or single, with or without children, etc.). As will become clear from the discussion below, this assumption ensures that agents

perceive differential employment conditions as unrelated to performance or skills within the organization, thus maximizing the likelihood that social comparison costs may arise.

Assumption 2 implies that an agent will not work for the principal unless he can secure a strictly positive payment. Note that this assumption does not preclude the case in which agents “like” their jobs, it just requires that they do not like it enough to be willing to work without a payment. As will become clear in section 4, this assumption simplifies the task of characterizing optimal contracts by ensuring that no agent would want to deviate from the formally agreed actions without the principal’s approval.

### *3.2 Preferences and Social Comparison Costs*

We assume that in addition to their own payoffs, the agents care about the payoffs of their peers. We make two important assumptions about these social comparison processes. First, we assume that the agents compare their “formal payoffs” (that is, the payoffs that would arise if the formal contracts were enforced:  $\pi_A(a_{it}^F) + p_{it}^F$ ) rather than their actual payoffs (that is,  $\pi_A(a_{it}) + p_{it}$ , which may differ from  $\pi_A(a_{it}^F) + p_{it}^F$  if the formal contracts are not enforced). The psychological rationale for this assumption is that agents care about their perceived status within the organization (e.g., Adams, 1965; McAdams, 1992, 1995). In turn, an agent’s status is more strongly affected by formal contract terms because formally specified tasks and (to some extent) compensation are more visible to its peers.<sup>6</sup> As will become clear later, our main results would remain qualitatively unchanged if the agents cared also about actual payoff differentials, as long as formal payoff differentials are more important.

---

<sup>6</sup> Alternatively, we could imagine that the agents care about their reputation outside of the organization, which depends on the information available to “outsiders”. So long as formal contract terms are more easily observed outside of the organization than informal ones, the two interpretations are identical.

Our second important assumption is that social comparisons affect the organization in two different ways. First, we assume that the agent with less favorable formal contract terms will feel frustrated, resulting in a direct utility loss (e.g., Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000).<sup>7</sup> We refer to this as the “disutility” effect. Second, we assume that the agent with less favorable formal contract terms will retaliate against the principal by engaging in acts of sabotage—that is, by taking actions detrimental to the organization, including, among others, the withdrawal of non-contractible cooperation or “shading” (e.g., Rabin, 1993; Hart and Moore, 2008, and ensuing literature). We refer to this as the “sabotage” effect. Acts of sabotage are not verifiable by a court and therefore cannot be prevented by the principal.<sup>89</sup>

To formally define social comparison costs, let  $(a_{it}^F, p_{it}^F)$  be the formal contract offered to agent  $i$  in period  $t$ , let  $a_{it}$  be the actual action that he takes, and let  $p_{it}$  be the actual payment that he receives from the principal. Denoting the agents’ propensity to social comparisons by  $\alpha^d \in (0,1)$  and their ability to sabotage the principal by  $\alpha^s \in [0, \infty)$ , agent  $i$ ’s utility in period  $t$  is given by:

$$u_{it} = \pi_A(a_{it}) + p_{it} - \alpha^d \max\{0, \pi_A(a_{jt}^F) + p_{jt}^F - \pi_A(a_{it}^F) - p_{it}^F\}, \quad j \in \{1,2\}, j \neq i,$$

whereas the principal’s utility in the same period is given by:

$$u_{Pt} = \pi_P(a_{1t}) + \pi_P(a_{2t}) - p_{1t} - p_{2t} - \alpha^s \alpha^d |\pi_A(a_{1t}^F) + p_{1t}^F - \pi_A(a_{2t}^F) + p_{2t}^F|.$$

---

<sup>7</sup> For simplicity, we assume that the agent with more favorable formal contract does not experience a utility gain. This asymmetry in social comparison effects is consistent with recent empirical evidence by Cohn et al. (2014).

<sup>8</sup> Arguably, the agent with less favorable formal contract terms may also retaliate against the agent with more favorable formal contract terms, causing a utility loss to the latter. We abstract from this additional social comparison cost, since adding it to the analysis does not change any of our main results.

<sup>9</sup> As in Hart and Moore (2008), one could argue that by engaging in acts of sabotage against the organization, the agents should be able to offset part of their disutility from receiving less favorable formal contract terms. Adding this possibility would not change any of our results.

Note that we have assumed that the two agents suffer equally from a negative social comparison and have the same ability to sabotage the principal, and that the agents' and principal's utility functions are linear in the magnitude of the formal payoff differential. These assumptions could all be relaxed without qualitatively affecting our results.

Implicit in our formulation is also the assumption that there is no upper bound on sabotage. This assumption captures the idea that, by engaging in sabotage, agents can destroy not only current output, but also the organization's reputation and future production capacity. This assumption can also be relaxed without substantially affecting our results.

### *3.3 Optimal Contract without Social Comparison Costs*

For future reference, we begin by considering the case in which the agents do not engage in social comparisons, or equivalently, in which  $\alpha^d = 0$ . The next proposition characterizes the principal's optimal contract under these circumstances.

**Proposition 1.** *Suppose that  $\alpha^d = 0$ . Then, regardless of the value of  $\delta$ :*

- (i) *The following contract is optimal for the principal:  $a_i^F = a^{FB}$  and  $p_i^F = \underline{U}_i - \pi_A(a^{FB})$  for  $i = 1, 2$ .*
- (ii) *Any optimal contract yields the principal a per period payoff of  $2S(a^{FB}) - \underline{U}_1 - \underline{U}_2$ .*

**Proof.** See Appendix.

The proposition shows that without social comparison costs, a simple formal contract between the principal and the two agents achieves the first-best. The intuition behind this result is straightforward. Since actions and payments are both contractible, the principal can rely on

formal contracts to implement the first-best action, thereby maximizing surplus, and extract all rents from the two agents through the monetary payments. Because the agents have different outside options, they will receive different payments. But in the absence of social comparisons, these differences in formal contract terms do not generate any costs to the organization.

#### 4. Social Comparison Costs under Spot Contracting

In this section, we characterize the principal’s optimal contract under spot transactions—that is, when the principal and the agents meet only once or do not use history dependent strategies. Because we are interested in analyzing how social comparison costs affect contract terms, we assume that it is profitable to employ the two agents.<sup>10</sup>

Recall that actions and payments are both contractible. Consistent with actual court behavior, we assume that if a breach of contract is verified, the court will force the breaching party to pay damages. In what follows, we shall assume that damages are calculated using the expectation damages remedy, but our results would remain qualitatively unchanged under alternative breach remedies.<sup>11</sup> For simplicity, we abstract from litigation costs.

We start by showing that under spot transactions all parties will adhere to their respective formal contracts. To see this, imagine that the principal agrees with agent  $i$  on a formal contract specifying action  $a_i^F$  and payment  $p_i^F$ . Imagine further that after accepting the principal’s offer, agent  $i$  decides to breach the contract by choosing action  $a_i \neq a_i^F$ . In that case, not only will the

---

<sup>10</sup> Employing the two agents will be optimal if the total surplus generated by the efficient action is “sufficiently” large—that is, if  $S(a^{FB}) \equiv \pi_P(a^{FB}) + \pi_A(a^{FB})$  is “sufficiently” large.

<sup>11</sup> Under expectation damages, the breaching party must compensate the affected party to ensure that the latter is in the same position as if the contract had been fulfilled as intended. For a discussion of different breach remedies and their role in shaping investment decisions, see, e.g., Shavell (1980) and Ohlendorf (2009).

principal refuse to make the agreed-upon payment to the agent, but he will also be entitled to damages equal to:

$$D(a_i, a_i^F, p_i^F) \equiv \max\{0, \pi_P(a_i^F) - p_i^F - \pi_P(a_i)\}.$$
<sup>12</sup>

Accordingly, the utility of agent  $i$  following breach would be given by:

$$U_i = \pi_A(a_i) - D(a_i, a_i^F, p_i^F) - \alpha^d \max\{0, \pi_A(a_j^F) + p_j^F - \pi_A(a_i^F) - p_i^F\}.$$

But since  $\pi_A(a_i) < \underline{U}_i$  by assumption 2 and  $D(a_i, a_i^F, p_i^F) \geq 0$  by definition, it must be that  $U_i < \underline{U}_i$ , which means that the agent would have been better off rejecting the principal's offer and pursuing his outside option. Thus, if agent  $i$  accepts the principal's offers, he will take the action specified by the formal contract. Moreover, the principal will make the formally agreed payment to the agent, for otherwise the agent will use a court to secure it.

Given the previous discussion, the principal's optimal contract under spot transactions solves the following problem:

$$\max_{\{a_1^F, p_1^F, a_2^F, p_2^F\}} \pi_P(a_1^F) + \pi_P(a_2^F) - p_1^F - p_2^F - \alpha^s \alpha^d |\pi_A(a_1^F) + p_1^F - \pi_A(a_2^F) + p_2^F|.$$

subject to the agents' participation constraints:

$$\pi_A(a_i^F) + p_i^F - \alpha^d \max\{0, \pi_A(a_j^F) + p_j^F - \pi_A(a_i^F) - p_i^F\} \geq \underline{U}_i, \quad i = 1, 2.$$

The next proposition characterizes the principal's optimal contract under spot transactions.

**Proposition 2.** *The principal's optimal contract under spot transactions specifies first-best actions for the two agents (i.e.,  $a_1^{spot} = a_2^{spot} = a^{FB}$ ) and the following payments:*

---

<sup>12</sup> Consistent with actual court behavior, we assume that if one party's breach is advantageous for the other, a court of law will not allow the breaching party to sue for a reward. In this case, the damage payment is zero.

$$p_1^{spot} = \underline{U}_1 - \pi_A(a^{FB}),$$

$$p_2^{spot} = \begin{cases} \underline{U}_1 - \pi_A(a^{FB}) & \text{if } \alpha^s \alpha^d \geq 1, \\ \frac{\underline{U}_2 + \alpha^d \underline{U}_1}{1 + \alpha^d} - \pi_A(a^{FB}) & \text{if } \alpha^s \alpha^d \leq 1. \end{cases}$$

*In equilibrium, all parties adhere to their respective formal contracts.*

**Proof.** See Appendix.

The proposition shows that under spot transactions the principal will require that the two agents take the efficient action,  $a^{FB}$ . In other words, social comparison processes do not distort actions, even under spot transactions. Intuitively, this is because choosing an inefficient action unnecessarily destroys surplus, as the principal can adjust payments to capture any gain generated by implementing a more efficient action without increasing social comparison costs.

Although social comparisons do not distort actions, they do affect payments. As Proposition 2 shows, social comparison costs force the principal to “overpay” agent 2, the agent with the lowest outside option. The intuition is as follows. In the absence of social comparisons, the principal would simply set payments so as to leave both agents without rents. Clearly, this would require a higher payment to agent 1, the agent with the highest outside option. But notice that once we consider social comparisons, this will make agent 2 feel frustrated. The principal must then overpay agent 2 in order to cover the utility loss resulting from a negative social comparison and ensure that his participation constraint is satisfied. Furthermore, because of the sabotage effect, the principal may want to raise agent 2’s payment beyond the minimum level consistent with the latter’s participation constraint. This will be true when sabotage is too damaging for the organization (i.e.,  $\alpha^s \alpha^d \geq 1$ ), in which case the principal will prefer to pay the two agents equally to completely eliminate sabotage.



## 5. Social Comparison Costs under Relational Contracting

In this section we turn to the main part of our analysis, that is, how informal contracts—agreements that are enforced by the parties’ fear of losing a valuable ongoing relationship rather than by a court (e.g., Baker et al., 2002)—interact with formal contracts in managing social comparison costs.

Following Levin (2003), we define a *relational contract* as a complete plan for the relationship between the principal and the two agents —i.e., a plan that specifies the principal’s and agents’ behavior for every period and every possible history of the repeated game. Let  $h^t$  denote the history of the game at the beginning of period  $t$  and  $H^t$  the set of all such histories. Let the initial history  $h^1$  be equal to the empty set. Then, for each date  $t$  and history  $h^t \in H^t$ , a relational contract specifies: (i) the formal contract  $(a_i^F, p_i^F)$  and the informal contract  $(a_i^I, p_i^I)$  that the principal should offer to agent  $i \in \{1,2\}$ ; (ii) whether the agents should accept the principal’s offers; and, in the event of acceptance, (iii) the actual action that the agents should take; and (iv) the actual payments that the principal should make.

### 5.1 Self-Enforcing Relational Contracts

As standard in the literature (e.g., MacLeod and Malcomson, 1989; Levin, 2003), we say that a relational contract is self-enforcing if it constitutes an equilibrium of the repeated game. Since all information is publicly observable, we adopt *Subgame Perfect Equilibrium* as the equilibrium concept. In what follow, we restrict attention to stationary relational contracts, that is, relational contract in which behavior does not change along the equilibrium path. Accordingly, in what follows we drop the time subscript from all equilibrium variables in the model.

Following Abreu (1988), we assume that the parties enforce relational contracts using optimal punishments; that is, we assume that any deviation from equilibrium actions will lead to the deviant’s worst equilibrium outcome. We also assume that if the principal deviates from an informal agreement, even against only one of the agents, the two agents would communicate and coordinated to punish the principal—that is, we assume that the agents employ a multilateral punishment scheme.<sup>13</sup> Under such a scheme, the agents learn about any deviation by the principal, including one that does not affect them directly, and view it as evidence of “untrustworthiness” and a reason to stop cooperating.

For future reference, we note that the optimal punishment against the principal is a permanent reversion to spot transactions. This is true because: (i) the equilibrium under spot transactions is also an equilibrium of the repeated game; and (ii) the principal can unilaterally revert to spot transactions, which implies that no equilibrium can yield the principal less than the optimal spot contract.<sup>14</sup>

In the rest of this subsection, we discuss in detail the conditions under which a relational contract is self-enforcing. First, the agents must be willing to work for the principal. This requires that the agents’ utilities under relational contracting be at least as large as their outside options. Therefore, letting  $U_i^R$  denote agent  $i$ ’s (per period) utility under relational contracting, i.e.,

$$U_i^R \equiv \pi_A(a_i^I) + p_i^I - \alpha^d \max\{0, \pi_A(a_j^F) + p_j^F - \pi_A(a_i^F) - p_i^F\},$$

---

<sup>13</sup> For a comparison of the optimality of multilateral and bilateral punishment schemes in different environments, see Levin (2002).

<sup>14</sup> The optimal punishment against agent 1 is also a permanent reversion to spot transactions. This is true because agent 1’s utility under spot transactions equals his outside option (see Proposition 2). Note, however, that the optimal punishment against agent 2 is potentially more complicated. This also follows from Proposition 2, which shows that if sabotage is very damaging to the principal (i.e.,  $\alpha^s \alpha^d > 1$ ), then the equilibrium under spot transactions fails to push agent 2’s utility down to his outside option. In that case (and depending on the discount factor), there may exist equilibria of the repeated game in which the utility of agent 2 is lower than under spot transactions. Accordingly, the optimal punishment against agent 2 can potentially depend on the discount factor.

the following participation constraints must be satisfied:

$$U_i^R \geq \underline{U}_i, \quad i = 1, 2. \quad (PC_i)$$

Second, the agents must find it profitable to take the informal actions prescribed by the relational contract. This requires that the agents' continuation utilities from honoring the relational contract be at least as large as their continuation utilities from renegeing. Hence, defining  $U_i^{pun}$  as agent  $i$ 's (per period) utility under the worst punishment that can be credibly imposed on him, and defining  $u_i^{dev}(a_i^F, p_i^F)$  as his (current period) utility from an optimal deviation conditional on having previously accepted the formal contract  $(a_i^F, p_i^F)$ , the following incentive compatibility constraints must be satisfied:

$$\pi_A(a_i^l) + p_i^l + \frac{\delta}{1-\delta} U_i^R \geq u_i^{dev}(a_i^F, p_i^F) + \frac{\delta}{1-\delta} U_i^{pun}, \quad i = 1, 2. \quad (IC_i^{act})$$

Notice that an agent can deviate by complying with the formal contract signed at the beginning of the period or by breaching the formal contract and paying damages to the principal. Accordingly, ignoring any current utility loss resulting from a negative social comparison, agent  $i$ 's (current period) utility from an optimal deviation is given by

$$u_i^{dev}(a_i^F, p_i^F) = \max\{\pi_A(a_i^F) - p_i^F, \pi_A(a_i^{dev}) - D(a_i^{dev}, a_i^F, p_i^F)\},$$

where

$$a_i^{dev} \in \arg \max_{a \neq a_i^F} \{\pi_A(a) - D(a, a_i^F, p_i^F)\}.$$

Finally, it must be in the interest of the principal to make the informal payments and in the interest of the agents to accept them. Notice that since a deviation by the principal leads both agents to stop cooperating, his optimal deviation involves renegeing on the two agents

---

<sup>15</sup> For the sake of simplicity, we have omitted from both sides of the constraint any current disutility associated with a negative social comparison, i.e.,  $-\alpha^d \max\{0, \pi_A(a_i^F) + p_i^F - \pi_A(a_i^F) - p_i^F\}$ .

simultaneously. Therefore, letting  $U_P^R$  denote the principal's (per period) utility under relational contracting, i.e.,

$$U_P^R \equiv \pi_P(a_1^I) + \pi_P(a_2^I) - p_1^I - p_2^I - \alpha^S \alpha^d |\pi_A(a_1^F) + p_1^F - \pi_A(a_2^F) + p_2^F|$$

the principal will be willing to make the informal payments as long as the following incentive compatibility constraint is satisfied:

$$-\sum_{j=1,2} p_j^I + \frac{\delta}{1-\delta} U_P^R \geq \sum_{j=1,2} p_j^{dev}(a_j^I, a_j^F, p_j^F) + \frac{\delta}{1-\delta} U_P^{spot}, \quad i = 1,2 \quad (IC_P^{pay})$$

where  $U_P^{spot}$  denotes the principal's utility under spot transactions—which, as mentioned before, coincides with the worst punishment that can be credible imposed on him—and

$$p_i^{dev}(a_i^I, a_i^F, p_i^F) = \begin{cases} p_i^F & \text{if } a_i^I = a_i^F, \\ -D(a_i^I, a_i^F, p_i^F) & \text{if } a_i^I \neq a_i^F, \end{cases}$$

denotes the payment made by the principal when optimally deviating against agent  $i$ .

To understand the expression for  $p_i^{dev}(a_i^I, a_i^F, p_i^F)$ , notice that the principal's optimal deviation consists of making the current payments to the agents as low as possible. The smallest payments that the principal can set, however, depend on whether the agents have previously complied with the actions prescribed by their formal contracts. When  $a_i^I = a_i^F$ , the principal cannot pay less than what was formally agreed, or else the affected agent would use a court to secure such a payment. In this case, the lowest payment that the principal can make when deviating against agent  $i$  is  $p_i^F$ . In contracts, when  $a_i^I \neq a_i^F$ , not only can the principal refuse to make the agreed-upon payment to the agent, but he can (and will) demand a compensation of  $D(a_i^I, a_i^F, p_i^F)$  in damages, which the agent will agree to pay since the principal can credibly threaten to go to court.

On the other hand, the agents will be willing to accept the informal payments as long as the following incentive compatibility constraints are satisfied:

$$p_i^l + \frac{\delta}{1-\delta} U_i^R \geq p_i^{dev}(a_i^l, a_i^F, p_i^F) + \frac{\delta}{1-\delta} U_i^{pun}, \quad i = 1, 2, \quad (IC_i^{pay})$$

where, as mentioned earlier,  $U_i^{pun}$  denotes agent  $i$ 's (per period) utility under the worst punishment that can be credibly imposed on him and  $p_i^{dev}(a_i^l, a_i^F, p_i^F)$  denotes the payment received by the agent in case of a deviation.

Notice that the function  $p_i^{dev}(a_i^l, a_i^F, p_i^F)$ , the payment received by agent  $i$  when optimally deviating against the principal, is the same as the payment made by the principal when optimally deviating against the agent. This is because, conditional on a deviation, the principal will try to minimize the payments to be made to the agents whereas the agents will try to maximize them. The actual payments would therefore be those to which the parties are legally entitled to, which, as mentioned earlier, depend on whether the agents have previously complied with the actions prescribed by their formal contracts.

## 5.2. Optimal Relational Contracts

We now characterize the principal's optimal relational contract—that is, the one that maximizes the principal's utility among all the self-enforcing relational contracts.

Let  $\mathbf{a}^F \equiv \{a_1^F, a_2^F\}$ ,  $\mathbf{a}^l \equiv \{a_1^l, a_2^l\}$ ,  $\mathbf{p}^F \equiv \{p_1^F, p_2^F\}$  and  $\mathbf{p}^l \equiv \{p_1^l, p_2^l\}$ . Given the discussion in the previous section, the principal's optimal relational contract solves the following problem:

$$\max_{(\mathbf{a}^F, \mathbf{a}^l, \mathbf{p}^F, \mathbf{p}^l)} \pi_P(a_1^l) + \pi_P(a_2^l) - p_1^l - p_2^l - \alpha^S \alpha^d |\pi_A(a_1^F) + p_1^F - \pi_A(a_2^F) + p_2^F|$$

subject to the agents' participation constraints,  $(PC_i)$ , the agents' incentive compatibility constraints,  $(IC_i^{act})$  and  $(IC_i^{pay})$ , and the principal's incentive compatibility constraint,  $(IC_P^{pay})$ .<sup>16</sup>

The next result shows that when the members of the organization are patient enough, the principal can combine formal and informal contracts to completely eliminate social comparison costs from the organization.

**Proposition 3.** *There exists  $\bar{\delta}^* \in (0,1)$  such that if the parties are patient enough ( $\delta \geq \bar{\delta}^*$ ), any optimal relational contract has the following characteristics:*

- (i) *Informally agreed actions are efficient:  $a_1^I = a_2^I = a^{FB}$ ;*
- (ii) *Informally agreed payments extract the two agents' rents:  $p_i^I = \underline{U}_i - \pi_A(a^{FB})$  for  $i = 1,2$ ;*
- (iii) *The payoffs the agents would receive if the formal contracts were enforced are identical:  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F \leq \underline{U}_2$ .*

**Proof.** See Appendix.

This proposition can deliver sharp testable predictions about the terms of the formal contract under the (realistic) assumption that drafting and administering formal contracts is somewhat costly, and therefore the principal prefers to minimize contractual complexity if possible. To be more precise, assume the principal incurs a small cost for drafting a formal contract but once this cost is incurred, he can use the same contract for the two agents at no additional cost. Then, the next result immediately follows from part (iii) of Proposition 3.

---

<sup>16</sup> Note that we have omitted the principal's participation constraint. Since the optimal spot contract is itself a relational contract and it yields the principal a payoff that is strictly greater than his outside option, the principal's participation constraint does not bind.

*Corollary.* *If the parties are patient enough ( $\delta \geq \bar{\delta}^*$ ), then in any optimal relational contract the formal contract terms offered to the two agents are identical:  $a_1^F = a_2^F$  and  $p_1^F = p_2^F$ .*

Proposition 3 and its corollary imply that when the members of the organization are patient enough, the principal can use formal contracts to decrease social comparison costs while using informal payments to ensure differentiation in actual compensation levels. More precisely, the principal can offer identical formal contract terms to the two agents and use informal payments to ensure that each agent is compensated according to his outside option.

The intuition behind this result is straightforward. By offering identical formal contract terms to the two agents, the principal ensures that the agents will neither take actions that reduce the output of the organization (i.e., engage in sabotage) nor experience any disutility associated with a negative social comparison. In other words, offering identical formal contracts to the agents allows the principal to eliminate social comparison costs from the organization. The key part of the argument is that, although the principal can always eliminate social comparison costs via “egalitarian” formal contracts, when the members of the organization are patient enough the principal can do so without having to actually overpay any of the agents. Essentially, the principal can offer the two agents *any* pair of formal contracts, provided that they are identical, and then use the informally agreed payments to ensure that each agent is actually compensated according to his outside option. When the members of the organization are sufficiently patient, they will all have incentives to comply with the informal payments, allowing the principal to capture the gains from eliminating social comparison costs.

Define  $\psi \equiv \frac{\alpha^d(1+\alpha^s)}{(1+\alpha^d)}$ . The next proposition characterizes the principal’s optimal relational contract when the discount factor is relative small.

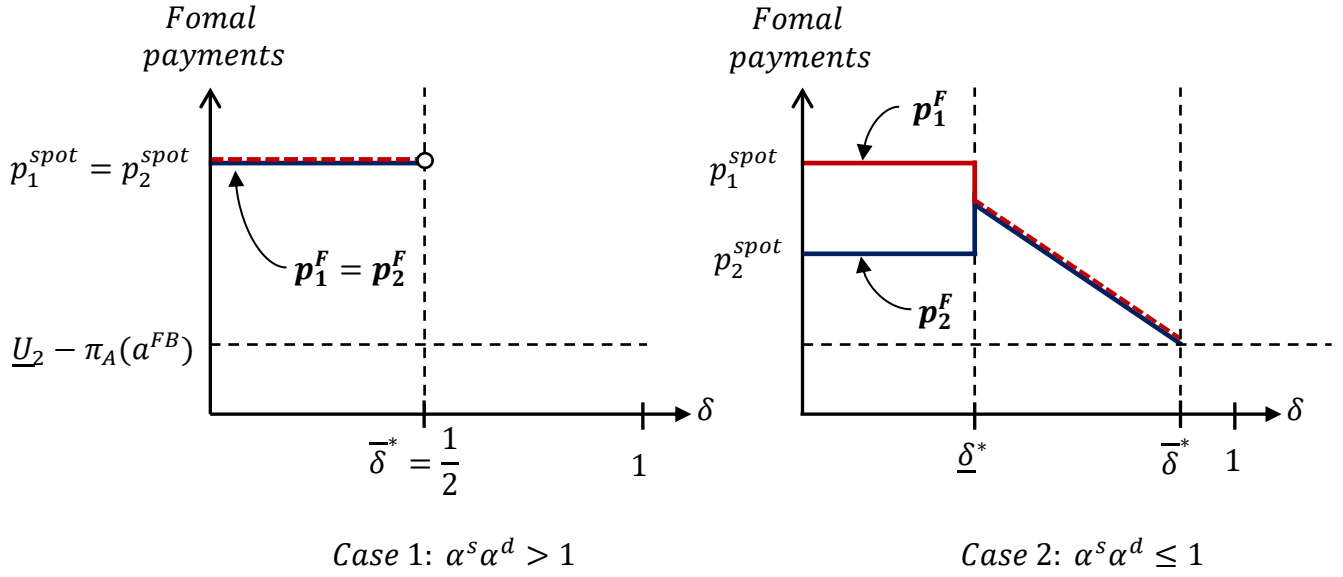


Figure 2: Formal payments for small values of  $\delta$

**Proposition 4.** *Suppose the parties are not highly patient ( $\delta < \bar{\delta}^*$ ). If the agents have high sabotage power ( $\alpha^s \alpha^d > 1$ ), then the optimal relational contract coincides with the optimal spot contract. If the agents have low sabotage power ( $\alpha^s \alpha^d \leq 1$ ), then there exists  $\underline{\delta}^* \in [0, \bar{\delta}^*)$  such that:*

1. *If the parties are impatient ( $\delta \in (0, \underline{\delta}^*)$ ), the (unique) optimal relational contract coincides with the optimal contract under spot transactions.*
2. *If the parties are moderately patient ( $\delta \in (\underline{\delta}^*, \bar{\delta}^*)$ ), the (unique) optimal relational contract has the following characteristics:*
  - (i) *Both formally and informally agreed actions are efficient:  $a_i^F = a_i^I = a^{FB}$  for  $i = 1, 2$ ;*



(ii) *Formally agreed payments are identical:  $p_1^F = p_2^F = \underline{U}_1 - \pi_A(a^{FB}) -$*

$$\frac{\delta}{1-\delta}(\underline{U}_1 - \underline{U}_2)\psi;$$

(iii) *Informally agreed payments extract the high-outside-option agent's rents:  $p_1^I =$*

$$\underline{U}_1 - \pi_A(a^{FB}); \text{ and}$$

(iv) *Informally agreed payments may leave some rents to the low-outside-option*

$$\text{agent: } p_2^I = \underline{U}_1 - \pi_A(a^{FB}) - \delta(\underline{U}_1 - \underline{U}_2)(1 + \psi).$$

3. *If  $\delta = \underline{\delta}^*$  then there are multiple optimal contracts but they all have the following characteristics:*

(i) *Formally and informally agreed actions, as well as the informally agreed*

$$\text{payment to agent 1, are as in part 2: } a_i^F = a_i^I = a^{FB} \text{ for } i = 1, 2, \text{ and } p_1^I = \underline{U}_1 - \pi_A(a^{FB});$$

(ii) *The informally agreed payment to agent 2 and the formally agreed payments are a convex combination of the optimal spot contract and the contract described in part 2:*

$$i. \quad p_2^I = \gamma p_2^{spot} + (1 - \gamma)[\underline{U}_1 - \pi_A(a^{FB}) - \delta(\underline{U}_1 - \underline{U}_2)(1 + \psi)],$$

$$ii. \quad p_1^F = \gamma p_1^{spot} + (1 - \gamma)\left[\underline{U}_1 - \pi_A(a^{FB}) - \frac{\delta}{1-\delta}(\underline{U}_1 - \underline{U}_2)\psi\right],$$

$$iii. \quad p_2^F = \gamma p_2^{spot} + (1 - \gamma)\left[\underline{U}_1 - \pi_A(a^{FB}) - \frac{\delta}{1-\delta}(\underline{U}_1 - \underline{U}_2)\psi\right],$$

$$\text{where } \gamma \in [0, 1].$$

**Proof.** See Appendix.

The above proposition shows that if the members of the organization are highly impatient, then the future value of the relationship is too small for relational contracting to be effective—

that is, no informally agreed payments and actions can be self-enforcing. In that case, the best the principal can do is to offer each agent the same (formal) contract as under spot transactions.

Alternatively, if the members of the organization are moderately patient, the future value of the relationship allows the principal to offer more homogeneous formal contracts—thereby reducing social comparison costs—, but the principal may need to leave some rents to the low-outside-option agent to ensure that the latter is willing to honor the informal agreement. Figure 2 illustrates the behavior of the formally agreed payments to the agents when the discount factor is relatively small.

## **6. Social Comparison Costs and Organization Design**

Nickerson and Zenger (2008) argue that the need to manage social comparison costs importantly distorts an organization’s design—that is, its formal compensation policy, boundaries, and internal architecture. Moreover, they argue that these distortions increase in the scope and size of the organization: “The scale and scope of the firm are primary drivers of social comparison costs. Hence, as scale and scope increase, the need to attenuate social comparison costs through incentive dampening, production efficiency compromises, or boundary restriction increases” (Nickerson and Zenger, 2008, p. 1431).

In this section, we discuss implications of our model for organizational design and performance. We show that in the presence of social comparison costs, pay compression and distortions in firm boundaries and architecture are primarily driven by the depth and strength of informal relationships inside the firm. We also show that the relationship between organizational distortions and the extent to which agents engage in social comparisons is not as clearcut as argued by Nickerson and Zenger (2008).

## 6.1 Formal Compensation Policy

We start by analyzing how social comparison costs shape the organization's formal compensation policy. We define *formal pay compression* as the compensation differential between the two agents if formal contracts were enforced, that is:  $|p_1^F - p_2^F|$ . For the next result, we make the (realistic) assumption that the principal incurs a small cost for drafting a formal contract but once this cost is incurred, he can use the same contract for the two agents at no additional cost. As discussed earlier on, this assumption ensures that when the principal and the agents are patient (large  $\delta$ ), the optimal relational contract entails equal formal payments and actions for the two agents (Proposition 4).

**Proposition 5.** *As transactions within the organization become more relational (that is, as the discount factor  $\delta$  increases), formal pay compression increases.*

**Proof.** See Appendix.

Proposition 5 may seem counterintuitive because it suggests that informal relationships increase contractual rigidity. Yet, the argument behind it is simple, and rests on the fact that unlike in the existing literature, we allow for formally contracted pay levels to differ from the actual, informally agreed ones. As the organization gets more relational, it becomes easier for the principal to credibly commit to honor an informal agreement with the agents. In turn, the possibility of using relationally-enforced informal payments allows the principal to offer more homogeneous formal payments—thereby minimizing social comparison costs—without having to actually overpay any of the agents. Finally, the reduction in social comparison costs limits output losses due to retaliation and sabotage by the two agents—in other words, it reduces organizational conflict.

## 6.2. Firm Boundaries and Organizational Architecture

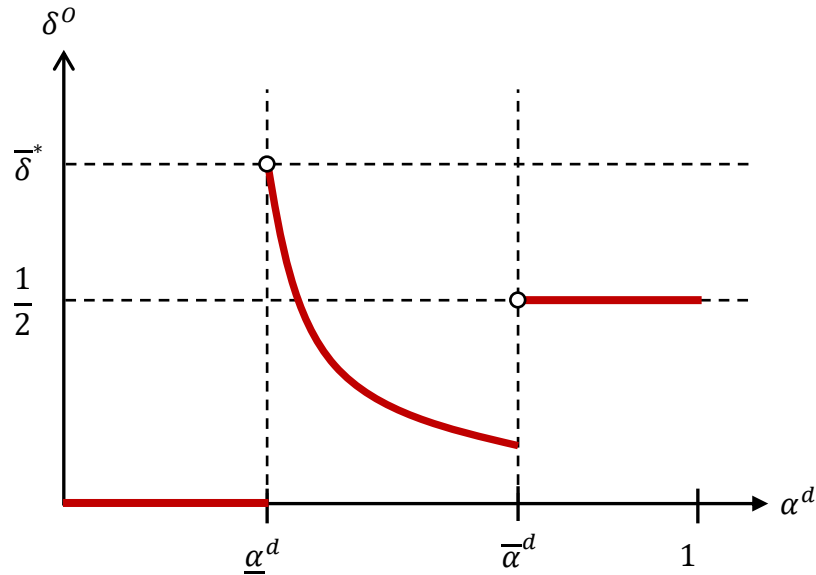


Figure 3: Threshold  $\delta^O$  as a function of  $\alpha^d$

Although social comparisons may occur both between and within organizations, the relative strength of such comparisons is likely to vary with organizational boundaries and internal architecture. In particular, Nickerson and Zenger (2008) argue that while employees within a firm compare each other's status, they are less likely to engage in similar comparisons with employees outside the firm, due to the lower degree of physical proximity and social interaction that links them to those individuals, and to the lack of a common sense of identity. A similar, albeit perhaps smoother reduction in social comparisons may arise when employees belong to different departments or divisions within the same organization. Quite naturally given these observations, Nickerson and Zenger (2008) also suggest that an organization may choose to alter firm boundaries, or its internal architecture, to manage social comparison costs, and that the incentive to enact such alterations is stronger the more propense and sensitive to social comparisons the organization's members.

To incorporate firm boundaries and organizational architecture into our model in the simplest possible way, suppose that absent social comparison costs, it would be optimal for the principal to employ both agents (integration) and/or allocate the two agents to the same production unit (simple organizational architecture). In the spirit of Nickerson and Zenger (2008), suppose, further, that social comparisons would disappear if the principal outsourced one of the two agents' tasks to a separate firm, or allocated the two agents to different production units (complex organizational architecture). Finally, suppose that deviating from the optimal governance—that is, switching from integration to outsourcing, or from simple to complex organizational architecture—causes a loss  $k > 0$  to the principal.<sup>17</sup>

**Proposition 6.** *Suppose that the cost of separating the two tasks,  $k$ , is not too large.<sup>18</sup> Then, there exists a critical discount rate,  $\delta^o$ , such that the two tasks are undertaken in house (integration) for  $\delta > \delta^o$ , whereas one of the tasks is outsourced for  $\delta < \delta^o$ . Moreover, there exist critical degrees of social comparison propensity,  $\underline{\alpha}^d$  and  $\bar{\alpha}^d$ , such that the outsourcing region, indexed by  $\delta^o$ , varies with  $\alpha^d$  as follows:*

(i)  $\delta^o = 0$  for  $\alpha^d \in [0, \underline{\alpha}^d]$ .

(ii)  $\delta^o > 0$  and strictly decreasing for  $\alpha^d \in (\underline{\alpha}^d, \bar{\alpha}^d]$ , with  $\delta^o < 1/2$  at  $\alpha^d = \bar{\alpha}^d$ .

(iii)  $\delta^o = 1/2$  for  $\alpha^d > \bar{\alpha}^d$ .

---

<sup>17</sup> The literature has pointed out multiple reasons for why the reduced form loss  $k$  may arise. For instance, outsourcing may bias the partner towards profit maximization at the expense of unobservable activities that are valuable to the principal (Holmstrom and Milgrom, 1991, 1994), or it may reduce the principal's incentive to undertake specific investments (Williamson, 1979; Grossman and Hart, 1986; Hart and Moore, 1990). Alternatively, both outsourcing and the complex organizational architecture may reduce the two agents' ability to develop communication routines that help them coordinate production (Cremer et al., 2007).

<sup>18</sup> If the cost of separating the two tasks is *too large*, so that it dominates the benefit in terms of reduction of social comparisons, it would never be optimal to outsource.

**Proof.** See Appendix.

Clearly, the proposition applies identically if one interprets  $\delta^o$  as the degree of distortion in the organization's internal architecture rather than as a distortion in firm boundaries.

The above result highlights two important points. First, long-term informal relationships unambiguously discourage distortions in firm boundaries and organizational architecture. That is because informal relationships permit to reduce social comparison costs within an integrated firm or department through the combination of homogeneous formal compensation and diverse informal compensation. This possibility renders inefficient outsourcing and divisionalization unnecessary.

Second, and most importantly, our model shows that if relational contracts can be used to manage social comparison costs, the link between boundary and architecture distortions on one hand, and the agents' social comparison propensity on the other hand, is non-monotonic. Consistent with Nickerson and Zenger (2008), distortions in firm boundaries and internal organization do not occur till an intermediate level of social comparison propensity,  $\alpha^d$ , is reached. However, once  $\alpha^d$  grows beyond the intermediate threshold, further increases in the agents' social comparison propensity *reduce*, rather than increasing the organizational distortions.

The reason for this counterintuitive result is that in a spot organization, the agents envy each other in equilibrium and react by sabotaging the organization (proposition 2). Thus, high degrees of social comparison propensity make reversion to the spot organization, and hence deviations from the informal compensation agreements, less attractive for the principal. As a result, there is a region where social comparison between the two agents becomes the principal's "friend", in the sense that it enables the principal to homogenize formal contract terms while differentiating

the agents' pay via informal agreements, thereby eliminating the need for distortions in the organization's design.<sup>19</sup>

### 6.3. Job satisfaction and organizational conflict

As discussed in the introduction, there is clear evidence that social comparisons reduce the employees' job satisfaction, and may also trigger retaliation in the form of reduced effort and productivity (e.g., Cohn et al., 2014; Card et al., 2012; Ockenfels et al., 2015). To analyze our model's implications for these phenomena, we define *organizational conflict* as the output destroyed by the two agents' acts of sabotage, plus the disutility, or dissatisfaction, experienced by the agents as a result of negative social comparison:  $\alpha^d(1 + \alpha^s)|\pi_A(a_1^F) + p_1^F - \pi_A(a_2^F) + p_2^F|$ .

**Proposition 7.** *As transactions within the organization become more relational (that is, as the discount factor  $\delta$  increases), organizational conflict decreases.*

**Proof.** See Appendix.

This result establishes that both the employees' perception of unfairness and their acts of retaliation decrease as the organization becomes more relational. This occurs despite the fact that the strength of the informal relationship between the principal and the agents, measured by  $\delta$ , does not affect the agents' sensitivity to social comparisons. The intuition should by now be familiar. Unlike in the spot organization, in a relational organization the principal can use relationally-enforced informal payments to extract surplus from the agents, while offering them homogeneous formal payments to eliminate social comparisons. As a result, the agents do not perceive their treatment as unfair and do not retaliate against the organization.

---

<sup>19</sup> The idea that decreasing the parties' fallback option may facilitate enforcement of their informal agreements has been proposed in different contexts by Baker *et al.* (1994, 2002), Zanarone (2013), and Contreras (2017).

## 7. Conclusion

This paper has developed a formal model to analyze how organizations manage social comparison costs among their members. We have argued that an organization's members are less likely to compare each other's status when status differences are not explicit and formalized. In such a context, an organization engaged in a long-term informal relationship with its members may eliminate social comparison costs by offering them formally homogeneous contract terms, while relying on informal agreements to extract the surplus. This is true even if the members' tasks are verifiable—that is, in the absence of the moral hazard problem normally invoked to motivate informal agreements. We have also shown that in the presence of social comparison costs, the organization may distort its boundaries and internal organization—for instance, by engaging in excess outsourcing and divisionalization to prevent workers from comparing each other's status. However, such distortions are less likely to occur when the organization's informal relationship with its members is strong. Moreover, the members' propensity to engage in social comparisons may reduce organizational distortions by making breach of the informal agreement, and the consequent reversion to arm's-length interaction with its members, less attractive.

The model presented in this paper provides a tractable framework to analyze the role of social comparison costs in organization, which may be both extended and empirically tested. Regarding the extensions, one could use the model to explore how distortions in a firm's production technology—as opposed to distortions in its boundaries and internal architecture—may be driven by social comparisons, as well as the mediating role played by intra-firm informal contracting. Relatedly, one could fully endogenize co-determinants of social comparisons and productivity such as proximity among employees, and analyze the interlinkage between technology, social comparison costs, and organization design. Regarding empirical testing, the model provides a rich



set of predictions that link the expected duration and strength of informal relationships to observable policy outcomes such as pay compression, firm boundaries and the internal organization of the firm. While measuring the expected duration of informal relationships is challenging, recent advances in the empirical literature on relational contracting have made significant progress in this direction,<sup>20</sup> and may enable empirical researchers to systematically investigate the organizational implications of social comparison costs.

## References

Abreu, D., (1988), “On the Theory of Infinitely Repeated Games with Discounting,” *Econometrica*, 56(2), 383-96.

Adams, JS., (1963), “Toward an understanding of inequity,” *Journal of Abnormal and Social Psychology* 67(5): 422–436.

Akerlof, G., and J. Yellen, (1990), “The fair wage-effort hypothesis and unemployment,” *Quarterly Journal of Economics* 105: 255–283.

Baker, G., R. Gibbons, and K. J. Murphy, (1994), “Subjective Performance Measures in Optimal Incentive Contracts,” *Quarterly Journal of Economics*, 109(4), 1125-1156.

Baker, G., R. Gibbons, and K. J. Murphy (2002), “Relational Contracts and The Theory Of The Firm,” *Quarterly Journal of Economics* 117: 39-84.

---

<sup>20</sup> See Gil and Marion (2013), Macchiavello and Morjaria (2015), and Gil et al. (2017). See also Gil and Zanarone (2015) for a detailed discussion of measurement issues.

Baker, G., R. Gibbons, and K. J. Murphy, (2011), “Relational Adaptation,” *Unpublished manuscript, USC Marshall School*.

Bartling, B., and F. von Siemens, (2010), “The Intensity of Incentives in Firms and Markets: Moral Hazard with Envious Agents,” *Labour Economics* 17: 598-607.

Battigalli, P., and G. Maggi, (2008), “Costly contracting in a long-term relationship,” *The RAND Journal of Economics*, 39(2), 352-377.

Bolton, P., and A. Ockenfels, (2000), “A Theory of Equity, Reciprocity and Competition,” *American Economic Review* 90: 166-193.

Bradler, C., R. Dur, S. Neckermann, and A. Non, (2016), “Employee Recognition and Performance: A Field Experiment,” *Management Science* 62: 3085-99.

Breza, E., S. Kaur, and Y. Shamdasani, (2015), “The Morale Effects of Pay Inequality,” Stanford University working paper.

Burkus, D., (2016), “Why Keeping Salaries a Secret May Hurt your Company,” *Harvard Business Review*, March 2016.

Card, D., A. Mas, E. Moretti, and E. Saez, (2012), “Inequality at Work: The Effect of Peer Salaries on Job Satisfaction,” *American Economic Review* 102: 2981-3003.

Cohn, A., E. Fehr, B. Hermann, and F. Schneider, (2014), “Social Comparison and Effort Provision: Evidence from a Field Experiment,” *Journal of the European Economic Association* 12: 877-898.

Contreras, O., (2017), “Relational Contracts, Financing Constraints and Social Welfare,” working paper.

Cremer, J., L. Garicano, and A. Prat, (2007), “Language and the Theory of the Firm,” *Quarterly Journal of Economics* 122: 373-407.

Danziger, L., and E. Katz, (1997), “Wage Secrecy as a Social Convention,” *Economic Inquiry* 35: 59-69.

Edwards, M., (2005), “The Law and Social Norms of Pay Secrecy,” *Journal of Employment and Labor Law* 26: 41-63.

Englmaier, F., and S. Leider, (2012), “Contractual and Organizational Structure with Reciprocal Agents,” *American Economic Journal: Microeconomics*, 4(2), 146-83.

Englmaier, F., and A. Wambach, (2010), “Optimal incentive contracts under inequity aversion,” *Games and Economic Behavior*, 69(2), 312-328.

Akerlof, G. A. and J. L. Yellen., (1990), “The Fair Wage-Effort Hypothesis and Unemployment,” *Quarterly Journal of Economics*, 105(2), 255-283.

Fehr, E., and K. Schmidt, (1999), “A Theory of Fairness, Competition, and Cooperation,” *Quarterly Journal of Economics* 114: 817-868.

Fehr, E., O. Hart, and C. Zender, (2015), “How do Informal Agreements and Revision Shape Contractual Reference Points?” *Journal of the European Economic Association* 13: 1-28.

Futrell, C., (1978), “Effects of Pay Disclosure on Satisfaction for Sales Managers,” *Academy of Management Journal* 21: 140-144.

Gil, R., and J. Marion J, (2013), “Self-enforcing Agreements and Relational Contracts: Evidence from California Highway Procurement,” *Journal of Law, Economics and Organization* 29: 239-277.

Gil, R., and G. Zanarone, (2015), “On the Determinants and Consequences of Informal Contracting,” Johns Hopkins University working paper.

Gil, R., and G. Zanarone, (2017), “Formal and Informal Contracting: Theory and Evidence,” *Annual Review of Law and Social Science* (forthcoming).

Gil, R., M. Kim, and G. Zanarone, (2017), “The Value of Relational Adaptation in Outsourcing: Evidence from the 2008 Shock to the US Airline Industry,” Johns Hopkins University, working paper.

Gillan, S., J. Hartzell, and R. Parrino, (2009), “Explicit Versus Implicit Contracts: Evidence from CEO Employment Agreements,” *Journal of Finance* 64: 1629-55.

Grossman, S., and O. Hart, (1986), “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration,” *Journal of Political Economy* 94: 2691-2719.

Hall, B., (2000), “Compensation and Performance Evaluation at Arrow Electronics,” Harvard Business School Case 9-800-290.

Halonon, M., and O. Hart, (2013), “More is Less: Why Parties May Deliberately Write Incomplete Contracts,” Harvard University working paper.

Hart, O., (2009), “Hold-Up, Asset Ownership, and reference Points,” *Quarterly Journal of Economics* 124: 267-300.

Hart, O., and B. Holmstrom, (2010), “A Theory of Firm Scope,” *Quarterly Journal of Economics* 125: 483-513.

Hart, O., and J. Moore, (1990), “Property Rights and the Nature of the Firm,” *Journal of Political Economy* 98: 1119-58.

Hart, O., and J. Moore, (2008), “Contracts as Reference Points,” *Quarterly Journal of Economics*, 123(1), 1-48.

Hill, A., (2016), “Pay Transparency is the Last Taboo in Business,” *Financial Times*.

Holmstrom, B., and P. Milgrom, (1991), “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design,” *Journal of Law, Economics and Organization* 7: 24-52.

Holmstrom, B., and P. Milgrom, (1994), “The Firm as an Incentive System,” *American Economic Review* 84: 972-91.

Klein, B. (2000), “The Role of Incomplete Contracts in Self-Enforcing Relationships,” *Revue d'Économie Industrielle* 92: 67-80.

Klein, B. and K. M. Murphy (1988), “Vertical Restraints as Contract Enforcement Mechanisms,” *Journal of Law and Economics* 31: 265-97.

Kragl, J., (2015), “Group versus Individual Performance Pay in Relational Employment Contracts when Workers are Envious,” *Journal of Economics and Management Strategy*, 24(1), 131-150.

Kragl, J., and J. Schmid, (2009), “The impact of envy on relational employment contracts,” *Journal of Economic Behavior and Organization*, 72(2), 766-779.

Kvaløy, O., and T. E. Olsen, (2009), “Endogenous Verifiability and Relational Contracting,” *American Economic Review*, 99(5), 2193-2208.

Lemieux, T., B. MacLeod, and D. Parent, (2009), “Performance Pay and Wage Inequality,” *Quarterly Journal of Economics*, 124: 1-49.

Levin, J., (2002), “Multilateral Contracting and the Employment Relationship,” *Quarterly Journal of Economics*, 117(3), 1075-1103.

Levin, J., (2003), “Relational Incentive Contracts,” *American Economic Review*, 93(3), 835-857.

Macchiavello, R., and A. Morjaria, (2015), “The Value of Relationships: Evidence from a Supply Shock to Kenyan Rose Exports,” *American Economic Review* 105: 2911-45.

MacLeod, W. B., (2003), “Optimal Contracting with Subjective Evaluation,” *American Economic Review*, 93(1), 216-240.

MacLeod, W. B., (2007), “Reputations, Relationships, and Contract Enforcement,” *Journal of Economic Literature* 45: 595-628.

MacLeod, W. B., and J. M. Malcomson, (1989), “Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment,” *Econometrica*, 57(2), 447-80.

Malcomson, J. M., (2013), “Relational Incentive Contracts,” In Gibbons, R., and J. Roberts (editors), *The Handbook of Organizational Economics*, Princeton: Princeton University Press.

McAdams, R., (1992), "Relative Preferences," *Yale Law Journal* 9: 1-104.

McAdams, R., (1995), "Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination," *Harvard Law Review* 108: 1008-1083.

Nickerson, J. A., and T. R. Zenger, (2008), "Envy, comparison costs, and the economic theory of the firm," *Strategic Management Journal*, 29(13), 1429-1449.

Ockenfels, A., D. Sliwka, and P. Werner, (2015), "Bonus Payments and Reference Point Violations," *Management Science* 61: 1496-1513.

Ohlendorf, S., (2009), "Expectation Damages, Divisible Contracts, and Bilateral Investment," *The American Economic Review* 99:1608-1618.

Rabin, M., (1993), "Incorporating Fairness into Game Theory and Economics," *American Economic Review* 83: 1281-1302.

Shavell, S., (1980), "Damage Measures for Breach of Contract," *Bell Journal of Economics* 11: 466-490.

Shaw, J., (2014), "Pay Dispersion," *Annual Review of Organizational Psychology and Organizational Behavior* 1: 521-44.

Zanarone, G. (2013), "Contract Adaptation under Legal Constraints," *Journal of Law, Economics and Organization* 29: 799-834.

Zanarone, G., D. Lo, and T. Madsen, (2015), "The Double-Edged Effect of Knowledge Acquisition: How Contracts Safeguard Pre-Existing Resources," *Strategic Management Journal* (forthcoming).

## Appendix

### Proof of Proposition 1.

Part (i). The result follows from the following three observations: (a) since actions and payments are both contractible, the proposed contracts can be (unilaterally) enforced by the principal regardless of the value of  $\delta$ ; (b) since the two agents receive payoffs equal to their outside options, the proposed contracts will be accepted; and (c) since the proposed contracts specify first-best actions (thereby maximizing surplus) and extract all rents from the two agents through the monetary payments, they must be optimal.

Part (ii). Observe that under the contracts described in part (i), the principal receives a per period payoff of  $2S(a^{FB}) - \underline{U}_1 - \underline{U}_2$ , the maximum per period payoff consistent with the agents' participation constraints. The result then follows since all optimal contracts must yield the principal the same payoff. ■

### Proof of Proposition 2.

The result that all parties will adhere to their respective formal contracts was proved in the main text.

We first show that the optimal spot contracts specify first-best actions for the two agents, i.e.,  $a_1^{spot} = a_2^{spot} = a^{FB}$ . Suppose to the contrary that  $a_i^{spot} \neq a^{FB}$  for some  $i \in \{1,2\}$ . Consider now an alternative contract  $(\tilde{a}_i, \tilde{p}_i)$  which sets  $\tilde{a}_i = a^{FB}$  and  $\tilde{p}_i = p_i^{spot} + \pi_A(a_i^{spot}) - \pi_A(a^{FB})$ . It is straightforward to verify that the alternative contract satisfies the agents' participation constraints and increases the principal's payoff by  $\pi_P(a^{FB}) + \pi_A(a^{FB}) - \pi_P(a_i^{spot}) - \pi_A(a_i^{spot}) > 0$ , where



the inequality follows from the definition of  $a^{FB}$  and the assumption that  $a_i^{spot} \neq a^{FB}$ . This, however, contradicts the optimality of the original contract.

We now turn to the optimal spot payments. In what follows, define  $G_i(a_i, p_i) \equiv \pi_A(a_i) + p_i$  (hereafter, agent  $i$ 's *material payoff*). As an intermediate step, we prove that: (I)  $G_1(a_1^{spot}, p_1^{spot}) \geq G_2(a_2^{spot}, p_2^{spot})$ , and (II)  $G_1(a_1^{spot}, p_1^{spot}) = \underline{U}_1$ .

(I)  $G_1(a_1^{spot}, p_1^{spot}) \geq G_2(a_2^{spot}, p_2^{spot})$ . Suppose to the contrary that  $G_1(a_1^{spot}, p_1^{spot}) < G_2(a_2^{spot}, p_2^{spot})$ . From agent 1's participation constraint, it follows that  $G_1(a_1^{spot}, p_1^{spot}) > \underline{U}_1$ , which in turn implies that  $U_2(a_1^{spot}, p_1^{spot}, a_2^{spot}, p_2^{spot}) = G_2(a_2^{spot}, p_2^{spot}) > \underline{U}_2$ . Consider now decreasing agent 2's payment by a small  $\varepsilon > 0$ . It is easy to verify that the perturbation induces a new contract that satisfies both agents' participation constraints and increases the principal's utility by  $\varepsilon(1 + \alpha^s \alpha^d)$ , a contradiction.

(II)  $G_1(a_1^{spot}, p_1^{spot}) = \underline{U}_1$ . Suppose to the contrary that  $G_1(a_1^{spot}, p_1^{spot}) > \underline{U}_1$ . There are two cases to consider depending on the values of  $G_1(a_1^{spot}, p_1^{spot})$  and  $G_2(a_2^{spot}, p_2^{spot})$ . Suppose first that  $G_1(a_1^{spot}, p_1^{spot}) = G_2(a_2^{spot}, p_2^{spot})$ . From agent 2's participation constraint, it follows that  $U_2(a_1^{spot}, p_1^{spot}, a_2^{spot}, p_2^{spot}) = G_2(a_2^{spot}, p_2^{spot}) > \underline{U}_2$ . Consider now decreasing both agents' payments by a small  $\varepsilon > 0$ . It can be verified that for small enough  $\varepsilon$ , this perturbation induces a new contract that satisfies both agents' participation constraints and increases the principal's utility by  $2\varepsilon$ , which is a contradiction. Suppose next that  $G_1(a_1^{spot}, p_1^{spot}) > G_2(a_2^{spot}, p_2^{spot})$ . In this case, it is easy to show that decreasing agent's 1 payment by a sufficiently small  $\varepsilon > 0$  induces a new contract that satisfies both agents' participation constraints and increases the principal's utility by  $\varepsilon$ , another contradiction.

We are now ready to characterize the optimal spot payments. The result that  $p_1^{spot} = \underline{U}_1 - \pi_A(a^{FB})$  follows from substituting  $a_1^{spot} = a^{FB}$  into agent 1's participation constraint and using the previous results that  $G_1(a_1^{spot}, p_1^{spot}) \geq G_2(a_2^{spot}, p_2^{spot})$  and  $G_1(a_1^{spot}, p_1^{spot}) = \underline{U}_1$ . To find agent 2's optimal payment, notice that since  $G_1(a_1^{spot}, p_1^{spot}) \geq G_2(a_2^{spot}, p_2^{spot})$  and  $a_1^{spot} = a_2^{spot} = a^{FB}$ , it must be that  $p_2^{spot} \leq p_1^{spot}$ . Moreover, notice that (after substituting the optimal values of  $a_1^{spot}$ ,  $a_2^{spot}$  and  $p_1^{spot}$ ) the derivative of the principal's payoff with respect to  $p_2$  equals  $-1 + \alpha^s \alpha^d$  for  $p_2 < p_1^{spot}$ . Clearly, if  $-1 + \alpha^s \alpha^d > 0$ , then it is optimal to set  $p_2^{spot} = p_1^{spot}$ . Alternatively, if  $-1 + \alpha^s \alpha^d < 0$ , then it is optimal to make  $p_2^{spot}$  as small as possible. This is achieved by setting  $p_2^{spot} = \frac{\underline{U}_2 + \alpha^d \underline{U}_1}{1 + \alpha^d} - \pi_A(a^{FB})$ , which is the minimum value of  $p_2$  that satisfies agent 2's participation constraint. Finally, if  $-1 + \alpha^s \alpha^d = 0$ , then any value between  $\frac{\underline{U}_2 + \alpha^d \underline{U}_1}{1 + \alpha^d} - \pi_A(a^{FB})$  and  $p_1^{spot}$  is optimal. ■

### Proof of Proposition 3.

Define  $\bar{\delta}^* \equiv \max\left\{\frac{1}{2}, \frac{1 + \alpha^d}{1 + \alpha^d + \alpha^d(1 + \alpha^s)}\right\}$  and consider the relational contract  $\{a_i^I, a_i^F, p_i^I, p_i^F\}_{i=1,2}$  which sets  $a_i^I = a_i^F = a^{FB}$ ,  $p_i^I = \underline{U}_i - \pi_A(a^{FB})$  and  $p_i^F = \underline{U}_2 - \pi_A(a^{FB})$  for  $i \in \{1, 2\}$ . Notice that parts (i), (ii), and (iii) of the proposition will follow if we can prove that the proposed contract induces an equilibrium, for in that case: (a) such an equilibrium must be optimal, for it completely eliminates social comparison costs and extracts all rents from the two agents; (b) even if there are other optimal contracts, they must all feature  $a_i^I = a^{FB}$ ,  $p_i^I = \underline{U}_i - \pi_A(a^{FB})$  and  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F$ , for otherwise the principal's welfare would be strictly lower than under the

proposed contract; and (c) all optimal contracts must satisfy  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F \leq \underline{U}_2$ , for otherwise constraint  $(IC_2^{pay})$  would be violated.

To prove that the proposed contract induces an equilibrium, observe that if it does induce an equilibrium, then the fact that  $U_2^R(\cdot) = \underline{U}_2$  implies that we can set  $U_2^{pun} = \underline{U}_2$ . Using this observation, it is straightforward to verify that Eqs.  $(PC_i)$ ,  $(IC_i^{act})$  and  $(IC_i^{pay})$  are all satisfied, whereas Eq.  $(IC_P^{pay})$  will be satisfied provided that  $(\delta/(1-\delta))[2S(a^{FB}) - \underline{U}_1 - \underline{U}_2 - U_P^{spot}] \geq \underline{U}_1 - \underline{U}_2$ , where

$$U_P^{spot} = \begin{cases} 2S(a^{FB}) - 2U_1 & \text{if } a^s a^d \geq 1 \\ 2S(a^{FB}) - U_1 - \frac{U_2 + a^d U_1}{1+a^d} - a^s a^d \left( \underline{U}_1 - \frac{U_2 + a^d \underline{U}_1}{1+a^d} \right) & \text{if } a^s a^d \leq 1 \end{cases}$$

by Proposition 2. Using the fact that  $\bar{\delta}^* = (1 + \alpha^d)/[1 + \alpha^d + \alpha^d(1 + \alpha^s)]$  for  $\alpha^d \alpha^s < 1$  and  $\bar{\delta}^* = 1/2$  otherwise, we can then show that  $(IC_P^{pay})$  would be satisfied as long as  $\delta \geq \bar{\delta}^*$ , which is true by assumption. ■

#### **Proof of Proposition 4.**

To prove the proposition, we first establish a series of lemmas.

**Lemma A.1:** (i) If  $(PC_i)$  and  $(IC_i^{act})$  are both satisfied, then so is  $(IC_i^{pay})$ . (ii) If  $(PC_i)$  is satisfied, then  $(IC_i^{act})$  holds if and only if  $\pi_A(a_i^l) + p_i^l + \frac{\delta}{1-\delta} U_i^R(\cdot) \geq \pi_A(a_i^F) + p_i^F + \frac{\delta}{1-\delta} U_i^{pun}$ .

**Proof.** Part (i). There are two cases to consider depending on the values of  $a_i^F$  and  $a_i^l$ . Consider first the case in which  $a_i^F = a_i^l$ , so that  $p_i^{dev}(a_i^l, a_i^F, p_i^F) = p_i^F$ . We show that  $(IC_i^{pay})$  is

implied by  $(IC_i^{act})$ . Substituting  $a_i^F = a_i^l$  into  $(IC_i^{act})$  and eliminating common terms from both sides of the inequality yields  $p_i^l + \frac{\delta}{1-\delta} U_i^R(\cdot) \geq p_i^F + \frac{\delta}{1-\delta} U_i^{pun}$ , which is identical to  $(IC_i^{pay})$ . The result then follows trivially.

Consider next the case in which  $a_i^F \neq a_i^l$ , so that  $p_i^{dev}(a_i^l, a_i^F, p_i^F) = -D(a_i^l, a_i^F, p_i^F)$ . To show that  $(IC_i^{pay})$  is implied by  $(PC_i)$ , observe that since  $D(a_i^l, a_i^F, p_i^F) \geq 0$  by definition and since  $p_i^l > 0$  by constraint  $(PC_i)$  and Assumption 2 in the text, it must be that  $p_i^l > p_i^{dev}(a_i^l, a_i^F, p_i^F)$ . The result that constraint  $(IC_i^{pay})$  is satisfied then follows from the fact that  $U_i^R(\cdot) \geq U_i^{pun}$  by the definition of  $U_i^{pun}$ .

*Part (ii).* Suppose that  $(PC_i)$  is satisfied. The “only if” part follows directly from the fact that  $\max\{\pi_A(a_i^F) + p_i^F, \pi_A(a_i^{dev}) - D(a_i^{dev}, a_i^F, p_i^F)\} \geq \pi_A(a_i^F) + p_i^F$ . To prove the “if” part, let  $\pi_A(a_i^F) + p_i^F < \pi_A(a_i^{dev}) - D(a_i^{dev}, a_i^F, p_i^F)$  (otherwise the result is trivially true). In this case, Eq.  $(IC_i^{act})$  can be written as  $\pi_A(a_i^l) + p_i^l + \frac{\delta}{1-\delta} U_i^R(\cdot) \geq \pi_A(a_i^{dev}) - D(a_i^{dev}, a_i^F, p_i^F) + \frac{\delta}{1-\delta} U_i^{pun}$ , where  $U_i^R(\cdot) \geq U_i^{pun}$  by the definition of  $U_i^{pun}$ . Since  $\pi_A(a_i^l) + p_i^l \geq \underline{U}_j$  by Eq.  $(PC_i)$  and  $\underline{U}_j > \pi_A(a_i^{dev}) - D(a_i^{dev}, a_i^F, p_i^F)$  by Assumption 2 in the text and the fact that  $D(a_i^{dev}, a_i^F, p_i^F) \geq 0$ , constraint  $(IC_i^{act})$  must hold, proving the result. ■

**Lemma A.2:** Any principal’s optimal relational contract features: (i)  $\pi_A(a_1^F) + p_1^F \geq \pi_A(a_2^F) + p_2^F$ ; (ii)  $p_1^l = \underline{U}_1 - \pi_A(a_1^l)$ ; and (iii)  $a_1^l = a_2^l = a^{FB}$ .

**Proof.** As an intermediate step, we show that if  $\pi_A(a_i^F) + p_i^F > \pi_A(a_j^F) + p_j^F$ , then constraint  $(PC_i)$  must bind—i.e.,  $p_i^l = \underline{U}_j - \pi_A(a_i^l)$ . To this purpose, suppose to the contrary that  $\pi_A(a_i^F) +$

$p_i^F > \pi_A(a_j^F) + p_j^F$  but  $(PC_i)$  is strictly slack. Consider now decreasing  $p_i^F$  by an arbitrarily small  $\varepsilon > 0$  and decreasing  $p_i^l$  by  $(1 - \delta)\varepsilon$ . After some calculations, we can show that the perturbation (strictly) increases the principal's payoff and (weakly) relaxes all the constraints, except for  $(PC_i)$ . But since  $(PC_i)$  is slack by assumption, the perturbation induces a new equilibrium in which the principal is (strictly) better off, contradicting the optimality of the original contract.

Part (i). Suppose to the contrary that  $\pi_A(a_1^F) + p_1^F < \pi_A(a_2^F) + p_2^F$ , which, by the previous paragraph, implies that  $U_2^R(\cdot) = \underline{U}_2$ . We first show that constraint  $(IC_1^{act})$  must hold with strict inequality. Notice that substituting  $U_2^R(\cdot) = \underline{U}_2$  into  $(IC_2^{act})$  and using the fact that  $U_2^{pun} \geq \underline{U}_2$  yields  $\pi_A(a_2^l) + p_2^l \geq \pi_A(a_2^F) + p_2^F$ . Moreover, notice that since  $\underline{U}_1 > \underline{U}_2$  by Assumption 1 in the text, Eq.  $(PC_1)$  implies that  $\pi_A(a_1^l) + p_1^l > \pi_A(a_2^l) + p_2^l$ . Using the two previous observations, together with the original assumption that  $\pi_A(a_1^F) + p_1^F < \pi_A(a_2^F) + p_2^F$ , we can conclude that  $\pi_A(a_1^l) + p_1^l > \pi_A(a_1^F) + p_1^F$ , which, since  $U_1^R(\cdot) \geq U_1^{pun}$  by the definition of  $U_1^{pun}$ , implies that constraint  $(IC_1^{act})$  is strictly slack (as desired).

Consider now increasing  $p_1^F$  by an arbitrarily small  $\varepsilon > 0$ . Since  $\pi_A(a_1^F) + p_1^F < \pi_A(a_2^F) + p_2^F$  by assumption, only equation  $(IC_1^{act})$  is tightened. But we just proved that such a constraint was initially slack, ensuring that the perturbation induces a new equilibrium. Moreover, since the perturbation lowers sabotage, it (strictly) increases the principal's payoff, contradicting the optimality of the original contract.

Part (ii). The result that  $p_1^l = \underline{U}_1 - \pi_A(a_1^l)$  if  $\pi_A(a_1^F) + p_1^F > \pi_A(a_2^F) + p_2^F$  follows directly from the intermediate step which was proved at the beginning. Thus, it remains only to show that  $(PC_1)$  must bind (i.e.,  $p_1^l = \underline{U}_1 - \pi_A(a_1^l)$ ) when  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F$ . Suppose to the contrary that  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F$  but  $(PC_1)$  is strictly slack (i.e.,  $p_1^l > \underline{U}_1 - \pi_A(a_1^l)$ ).

We first prove that constraint  $(IC_1^{act})$  must also be strictly slack. Since  $U_1^R(\cdot) > U_1^{pun}$  by the definition of  $U_1^{pun}$ , it suffices to prove that  $\pi_A(a_1^l) + p_1^l > \pi_A(a_1^F) + p_1^F$ . To this end, observe that  $(PC_2)$  must bind (i.e.,  $U_2^R(\cdot) = \underline{U}_2$ )—otherwise decreasing both  $p_1^F$  and  $p_2^F$  by a small  $\varepsilon > 0$  and decreasing both  $p_1^l$  and  $p_2^l$  by  $(1 - \delta)\varepsilon$  would induce a new equilibrium in which the principal is (strictly) better off, contradicting the optimality of the original contract. But the fact that  $(PC_2)$  binds implies that  $\pi_A(a_1^l) + p_1^l > \pi_A(a_2^l) + p_2^l$  (since  $U_1^R(\cdot) > \underline{U}_1$  by the assumption that  $(PC_1)$  is strictly slack and  $\underline{U}_1 > \underline{U}_2$  by Assumption 1 in the main text) and that  $\pi_A(a_2^l) + p_2^l \geq \pi_A(a_2^F) + p_2^F$  (by equation  $(IC_2^{act})$  and the fact that  $U_2^{pun} \geq \underline{U}_2$ ). The result that  $\pi_A(a_1^l) + p_1^l > \pi_A(a_1^F) + p_1^F$  then follows directly from the assumption that  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F$ .

Finally, using the fact that Eqs.  $(PC_1)$  and  $(IC_1^{act})$  are both strictly slack (the first one by assumption and the second one as shown in the last paragraph), it is easy to verify that decreasing  $p_1^l$  by a small  $\varepsilon > 0$  would induce a new equilibrium in which the principal is (strictly) better off than under the original contract, a contradiction.

*Part (iii).* Suppose to the contrary that  $a_i^l \neq a^{FB}$  for some  $i \in \{1,2\}$ . There are two cases to consider depending on the values of  $a_i^l$  and  $a_i^F$ .

*Case 1:* Suppose first that  $a_i^l = a_i^F$ . Consider an alternative contract for agent  $i$  which sets  $\tilde{a}_i^l = a^{FB}$ ,  $\tilde{a}_i^F = a^{FB}$ ,  $\tilde{p}_i^F = \pi_A(a_i^F) + p_i^F - \pi_A(a^{FB})$ , and  $\tilde{p}_i^l = \pi_A(a_i^l) + p_i^l - \pi_A(a^{FB})$ . Define  $G(a, p) \equiv \pi_A(a) + p$ . Notice that  $G(a_i^l, p_i^l) = G(\tilde{a}_i^l, \tilde{p}_i^l)$  and  $G(a_i^F, p_i^F) = G(\tilde{a}_i^F, \tilde{p}_i^F)$ . Notice also that since  $a_i^l = a_i^F$  and  $\tilde{a}_i^l = \tilde{a}_i^F$ , it must be that  $p_i^{dev} = p_i^F$  and  $\tilde{p}_i^{dev} = \tilde{p}_i^F$ . Using these observations, it is straightforward to show that the alternative contract induces a new equilibrium

and increases the principal's payoff by  $S(a_i^I) - S(a^{FB}) > 0$ , contradicting the optimality of the original agreement.

*Case 2:* Suppose next that  $a_i^I \neq a_i^F$ . We prove that there exists an alternative equilibrium in which the principal is (strictly) better off, contradicting the optimality of the original agreement. To this end, consider the following perturbed contract for agent  $i$ :  $\tilde{a}_i^I = a^{FB}$ ,  $\tilde{a}_i^F = a^{FB}$ ,  $\tilde{p}_i^I = \pi_A(a_i^I) + p_i^I - \pi_A(a^{FB})$ , and  $\tilde{p}_i^F = \pi_A(a_i^F) + p_i^F - \pi_A(a^{FB})$ . It is straightforward to verify that the perturbation increases the principal's welfare by  $S(a^{FB}) - S(a_i^I) > 0$  and affects only Eq. ( $IC_p^{pay}$ ). Moreover, notice that since ( $IC_p^{pay}$ ) was initially satisfied, it will be satisfied under the perturbed contract provided  $\Delta IC_p^{pay} = -[\Delta p_i^I - \Delta p_i^{dev}] + (\delta/(1-\delta))\Delta U_p^R \geq 0$ , where  $\Delta x$  denotes the change in variable  $x$  as we move from the original to the perturbed contract. Notice also that since  $\Delta U_p^R = S(a^{FB}) - S(a_i^I) > 0$ , it suffices to prove that  $\Delta p_i^{dev} - \Delta p_i^I \geq 0$ .

To prove that  $\Delta p_i^{dev} - \Delta p_i^I \geq 0$ , observe that  $\Delta p_i^I = \pi_A(a_i^I) - \pi_A(a^{FB})$  and  $\Delta p_i^{dev} = \pi_A(a_i^F) + p_i^F - \pi_A(a^{FB}) + D(a_i^I, a_i^F, p_i^F)$ , where the last expression uses the facts that  $p_i^{dev}(a_i^I, a_i^F, p_i^F) = -D(a_i^I, a_i^F, p_i^F)$  and  $\tilde{p}_i^{dev}(a_i^I, a_i^F, p_i^F) = \tilde{p}_i^F$ . Hence,  $\Delta p_i^{dev} - \Delta p_i^I = \pi_A(a_i^F) + p_i^F - \pi_A(a_i^I) + D(a_i^I, a_i^F, p_i^F)$ . Consider now the following two cases depending on the values of  $a_i^I$  and  $a_i^F$ . Suppose first that  $S(a_i^F) \geq S(a_i^I)$ . Since  $D(a_i^I, a_i^F, p_i^F) \geq \pi_P(a_i^F) - p_i^F - \pi_P(a_i^I)$ , we have that  $\Delta p_i^{dev} - \Delta p_i^I \geq S(a_i^F) - S(a_i^I) > 0$ , proving the result.

Suppose next that  $S(a_i^F) < S(a_i^I)$ . We first show that there is no loss of generality in assuming that  $\pi_A(a_i^F) + p_i^F \geq \underline{U}_2$ . To see this, notice that if  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F$ , we can increase both  $p_1^F$  and  $p_2^F$  until  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F = \underline{U}_2$  without violating any constraint nor changing the welfare of any of the parties. Alternatively, if  $\pi_A(a_1^F) + p_1^F > \pi_A(a_2^F) + p_2^F$  (notice

that  $\pi_A(a_1^F) + p_1^F < \pi_A(a_2^F) + p_2^F$  is ruled out by part (i) of the lemma), we can increase  $p_2^F$  until  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F$  (in which case we are back to the previous case), or until  $\pi_A(a_2^F) + p_2^F = \underline{U}_2$  (in which case we can easily prove that the perturbation does violate any constraint nor changes the welfare of any of the parties). Either case, the result follows.

Finally, since  $\underline{U}_2 > \pi_A(a_i^I)$  by Assumption 2 in text, the result that  $\pi_A(a_i^F) + p_i^F \geq \underline{U}_2$  implies that  $\pi_A(a_i^F) + p_i^F > \pi_A(a_i^I)$ . Furthermore, because  $S(a_i^F) < S(a_i^I)$ , the fact that  $\pi_A(a_i^F) + p_i^F > \pi_A(a_i^I)$  implies that  $\pi_P(a_i^F) - p_i^F > \pi_P(a_i^I)$ , which in turn requires  $D(a_i^I, a_i^F, i) = 0$ . Accordingly, it follows that  $\Delta p_i^{dev} - \Delta p_i^I = \pi_A(a_i^F) + p_i^F - \pi_A(a_i^I) > 0$ , as desired. ■

**Lemma A.3:** If  $\delta < \bar{\delta}^*$ , then any principal's optimal contract satisfies: (i)  $\pi_A(a_1^F) + p_1^F > \pi_A(a_2^F) + p_2^F$  or  $U_2^R(\cdot) > \underline{U}_2$  (or both); (ii)  $(IC_2^{act})$  and  $(IC_P^{pay})$  are both binding; and (iii)  $a_1^F = a_2^F = a^{FB}$ .

**Proof. Part (i).** Let  $\delta < \bar{\delta}^*$  and suppose to the contrary that there exists an optimal contract in which both  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F$  (notice that  $\pi_A(a_1^F) + p_1^F < \pi_A(a_2^F) + p_2^F$  is ruled out by Lemma A.2) and  $U_2^R(\cdot) = \underline{U}_2$ . Observe that all optimal contracts must satisfy: (a)  $a_i^I = a^{FB}$ ,  $p_i^I = \underline{U}_i - \pi_A(a^{FB})$  and  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F$ , for otherwise the principal's welfare would be strictly lower than under the original optimal contract; and (b)  $\pi_A(a_1^F) + p_1^F = \pi_A(a_2^F) + p_2^F \leq \underline{U}_2$ , for otherwise constraint  $(IC_2^{pay})$  would be violated. Observe also that using an argument similar to that of case 2 of the proof of Lemma A.2(iii), we can show that there must exist an optimal contract in which  $a_i^I = a_i^F$  for  $i \in \{1, 2\}$ . Combining the two previous observations, it follows that there must exist an optimal contract which sets  $a_i^I = a_i^F = a^{FB}$ ,  $p_i^I = \underline{U}_i - \pi_A(a^{FB})$



and  $p_1^F = p_2^F \leq \underline{U}_2 - \pi_A(a^{FB})$ . But after substituting such values into  $(IC_2^{pay})$  and doing some calculations, we can show that  $(IC_2^{pay})$  would be violated for  $\delta < \bar{\delta}^*$ , a contradiction.

Part (ii). By part (i), we know that  $\pi_A(a_1^F) + p_1^F > \pi_A(a_2^F) + p_2^F$  or  $U_2^R(\cdot) > \underline{U}_2$  (or both). Suppose first that  $\pi_A(a_1^F) + p_1^F > \pi_A(a_2^F) + p_2^F$ . If  $(IC_2^{act})$  is slack, we can show that increasing  $p_2^F$  by a small amount would induce a new equilibrium in which the principal is (strictly) better off, contradicting the optimality of the original contract. Similarly, if  $(IC_p^{pay})$  is slack, we can show that decreasing  $p_1^F$  by a small amount would induce a new equilibrium in which the principal is (strictly) better off, another contradiction.

Suppose next that  $U_2^R(\cdot) > \underline{U}_2$ . If  $(IC_2^{act})$  is slack, we can show that decreasing  $p_2^F$  by a small amount would induce a new equilibrium in which the principal is (strictly) better off, a contradiction. Similarly, if  $(IC_p^{pay})$  is slack, we can show that decreasing both  $p_1^F$  and  $p_2^F$  by a small amount does not violate any constraints and, in fact, relaxes equation  $(IC_2^{act})$  without affecting the principal's payoff. We can then perform the same perturbation as in the previous case, thereby finding another contradiction.

Part (iii). Suppose to the contrary that  $a_i^F \neq a^{FB}$  for some  $i \in \{1,2\}$ . Recall that  $a_i^I = a^{FB}$  by Lemma A.2. Using the same perturbation as in case 2 of the proof of Lemma A.2(iii), we can construct a perturbed contract for agent  $i$  that yields the same payoffs to all the parties and relaxes constraint  $(IC_2^{act})$ . We can then use an argument analogous to that of part (i) to find a perturbation that satisfies all the constraints and strictly increases the principal's welfare, contradicting the optimality of the original contract. ■

**Lemma A.4:** The following is true in any principal's optimal contract: (i) If  $\alpha^s \alpha^d > 1$ , then  $\pi_A(a_1^F) + p_1^F = \pi_A(a_1^F) + p_2^F$ . (ii) If  $\alpha^s \alpha^d \leq 1$ , then there exists a  $\underline{\delta}_* \in [0, \bar{\delta}^*)$  such that  $(PC_2)$  binds for  $\delta < \underline{\delta}_*$  and  $p_1^F = p_2^F$  for  $\delta > \underline{\delta}_*$ . Moreover, when  $\delta = \underline{\delta}_*$ , any equilibrium in which  $a_i^F = a_i^I = a^{FB}$  for  $i \in \{1, 2\}$ ,  $p_1^I = \underline{U}_1 - \pi_A(a^{FB})$ , and in which the values of  $\{p_1^F, p_2^I, p_1^F\}$  are set in such a way that Eqs.  $(IC_2^{act})$  and  $(IC_P^{pay})$  are both binding and both  $(PC_2)$  and  $p_1^F \geq p_2^F$  are satisfied is optimal.

**Proof.** *Part (i).* Suppose to the contrary that  $p_1^F \neq p_2^F$ , which, by Lemmas A.2 and A.3 requires  $\pi_A(a_1^F) + p_1^F > \pi_A(a_2^F) + p_2^F$ . Consider now increasing both  $p_2^F$  and  $p_2^I$  by a small amount. It is straightforward to show that such a perturbation would induce a new equilibrium in which the principal is (strictly) better off, contradicting the optimality of the original contract.

*Part (ii).* Define  $\underline{\delta}_* \equiv \frac{1 - \alpha^s \alpha^d}{1 + \alpha^d + \alpha^d(1 + \alpha^s)}$ . Since  $\alpha^s \alpha^d \in [0, 1]$  and  $\alpha^d > 0$ , we know that  $\underline{\delta}_* \in [0, \bar{\delta}^*)$ , where  $\bar{\delta}^* = \min\left\{\frac{1}{2}, \frac{1 + \alpha^d}{1 + \alpha^d + \alpha^d(1 + \alpha^s)}\right\}$  by Proposition 3. In what follows, let  $\Delta_{F,1} = \frac{-\delta \alpha^d(1 + \alpha^s)}{1 - \delta - \delta \alpha^d(1 + \alpha^s)}$ ,  $\Delta_{I,2} = (1 - \delta) \left[ \frac{1 - \delta - \delta \alpha^d(2 + \alpha^s)}{1 - \delta - \delta \alpha^d(1 + \alpha^s)} \right]$ , and  $\Delta_P = (1 - \delta) \left[ \frac{\delta \alpha^d + \alpha^d \alpha^s}{1 - \delta - \delta \alpha^d(1 + \alpha^s)} - 1 \right]$ . There are four cases to consider depending on the value of  $\delta$ .

*Case 1:* Let  $\delta \in (0, \underline{\delta}_*)$  and suppose to the contrary that  $(PC_2)$  does not bind. Fix an arbitrarily small  $\varepsilon > 0$  and consider a perturbed contract which sets  $\tilde{p}_2^F = p_2^F - \varepsilon$ ,  $p_1^F = p_1^F - \varepsilon \Delta_{F,1}$  and  $p_2^I = p_2^I - \varepsilon \Delta_{I,2}$  while keeping the rest of the original contract unchanged. Observe that  $1 - \delta - \delta \alpha^d(1 + \alpha^s) > 0$  (since  $\delta < \underline{\delta}_*$  by assumption) and thus that  $\varepsilon \Delta_{F,1} < 0$ . After some calculations, it is straightforward to show that Eqs.  $(PC^1)$ ,  $(IC_P^{pay})$  and  $(IC_2^{act})$  remain unchanged and the principal's welfare changes by  $-\Delta_P$ . Moreover, using the fact that  $(IC_P^{pay})$  and  $(IC_2^{act})$  were both

initially binding (see Lemma A.3) and remain unchanged after the perturbation, we can show that  $(IC_1^{act})$  continues to hold. Since Eq.  $(PC_2)$  is slack by assumption, it then follows that the perturbation induces an equilibrium. Finally, using the definition of  $\underline{\delta}_*$ , together with the assumption that  $\delta \in (0, \underline{\delta}_*)$ , we can show that  $-\Delta_P > 0$ , contradicting the optimality of the original contract.

For the next cases, define  $\tilde{\delta} \equiv 1/[1 + \alpha^d(1 + \alpha^s)]$  and notice that  $\tilde{\delta} \in (\underline{\delta}_*, \bar{\delta}^*)$ .

*Case 2:* Let  $\delta \in (\underline{\delta}_*, \tilde{\delta})$  and suppose to the contrary that  $\pi_A(a_1^F) + p_1^F \neq \pi_A(a_2^F) + p_2^F$ , which, as argued before, requires  $\pi_A(a_1^F) + p_1^F > \pi_A(a_2^F) + p_2^F$ . Fix an arbitrarily small  $\varepsilon > 0$  and consider a perturbed contract which sets  $\tilde{p}_2^F = p_2^F + \varepsilon$ ,  $\tilde{p}_1^F = p_1^F + \varepsilon \Delta_{F,1}$  and  $\tilde{p}_2^I = p_2^I + \varepsilon \Delta_{I,2}$  while keeping the rest of the original contract unchanged. Notice that  $1 - \delta - \delta\alpha^d(1 + \alpha^s) > 0$  (since  $\delta < \tilde{\delta}$  by assumption) and thus that  $\varepsilon \Delta_{F,1} < 0$ . After some calculations, it is straightforward to show that Eqs.  $(PC_1)$ ,  $(IC_p^{pay})$  and  $(IC_2^{act})$  remain unchanged, Eqs.  $(PC_2)$  and  $(IC_1^{act})$  are both relaxed, and the principal's welfare changes by  $\Delta_P$ . Moreover, using the definitions of  $\underline{\delta}_*$  and  $\tilde{\delta}$ , together with the assumption that  $\delta \in (\underline{\delta}_*, \tilde{\delta})$ , we can show that  $\Delta_P > 0$ , contradicting the optimality of the original contract.

*Case 3:* Let  $\delta \in [\tilde{\delta}, \bar{\delta}^*)$  and suppose to the contrary that  $\pi_A(a_1^F) + p_1^F > \pi_A(a_2^F) + p_2^F$ . Consider now decreasing  $p_1^F$  by an arbitrarily small  $\varepsilon > 0$  and decreasing  $p_2^I$  by  $\delta\alpha^d\varepsilon$ . After some calculations, it is straightforward to show that Eqs.  $(PC_1)$  and  $(IC_2^{act})$  remain unchanged, Eqs.  $(PC_2)$  and  $(IC_1^{act})$  are relaxed, and Eq.  $(IC_p^{pay})$  is satisfied provided  $1 - \delta - \delta\alpha^d(1 + \alpha^s)$  or  $\delta \geq \tilde{\delta}$ , which is true by assumption. Moreover, since the perturbation increases the principal's welfare by  $\delta\alpha^d(1 + \alpha^s)\varepsilon > 0$ , the original contract cannot be optimal, proving the result.

*Case 4:* Finally, let  $\delta = \underline{\delta}_*$ . Fix a contract, hereafter called contract A, satisfying  $a_i^F = a_i^l = a^{FB}$  for  $i \in \{1,2\}$  and  $p_1^l = \underline{U}_1 - \pi_A(a^{FB})$ , which, in addition, sets  $\{p_1^F, p_2^l, p_1^F\}$  in such a way that Eqs.  $(IC_2^{act})$  and  $(IC_p^{pay})$  are both binding and both  $(PC_2)$  and  $p_1^F \geq p_2^F$  are satisfied. We show that contract A both is an equilibrium and maximizes the principal's welfare.

To prove that contract A induces an equilibrium, notice that  $(IC_2^{act})$ ,  $(IC_p^{pay})$  and  $(PC_2)$  hold by assumption,  $(IC_1^{act})$  is implied by the fact that  $(IC_p^{pay})$  and  $(IC_2^{act})$  are both binding, and  $(PC_1)$  is implied by the facts that: (a)  $\pi_A(a_1^F) + p_1^F \geq \pi_A(a_1^F) + p_2^F$ , and (b)  $p_1^l = \underline{U}_1 - \pi_A(a^{FB})$ . Since  $(IC_1^{pay})$  and  $(IC_2^{pay})$  can be ignored by Lemma A.1, we can conclude that contract A induces an equilibrium, as desired.

To prove that contract A is optimal, let  $\{\tilde{a}_i^l, \tilde{a}_i^F, \tilde{p}_i^l, \tilde{p}_i^F\}_{i=1,2}$  be an arbitrary optimal contract. By Lemmas A.2 and A.3, we know that  $\tilde{a}_i^F = \tilde{a}_i^l = a^{FB}$  for  $i \in \{1,2\}$  and  $\tilde{p}_1^l = \underline{U}_1 - \pi_A(a^{FB})$ , the same as in contract A. Accordingly, there are only three unknown variables in the optimal contract, namely  $\tilde{p}_2^l, \tilde{p}_1^F$ , and  $\tilde{p}_2^F$ . Moreover, since  $(IC_2^{act})$  and  $(IC_p^{pay})$  bind under both contract A and the optimal contract (see Lemma A.3), it is easy to verify that the two will be identical when  $\tilde{p}_1^F - \tilde{p}_2^F = p_1^F - p_2^F$ . When  $\tilde{p}_1^F - \tilde{p}_2^F \neq p_1^F - p_2^F$ , on the other hand, there are two relevant cases to consider. Imagine first that  $\tilde{p}_1^F - \tilde{p}_2^F < p_1^F - p_2^F$ . In this case, we can easily show that performing the same perturbation as in case 1 does not affect the principal's welfare, leaves Eqs.  $(IC_2^{act})$  and  $(IC_p^{pay})$  unchanged, and increases  $\tilde{p}_1^F - \tilde{p}_2^F$ . Accordingly, by setting  $\varepsilon$  in such a way that the new level of  $\tilde{p}_1^F - \tilde{p}_2^F$  equals  $p_1^F - p_2^F$ , we can generate a new optimal contract identical to contract A, as desired. Alternatively, if  $\tilde{p}_1^F - \tilde{p}_2^F > p_1^F - p_2^F$ , we can show that performing the same perturbation as in case 2 does not affect the principal's welfare, leaves Eqs.  $(IC_2^{act})$  and  $(IC_p^{pay})$

unchanged, and decreases  $\tilde{p}_1^F - \tilde{p}_2^F$ . Thus, by setting  $\varepsilon$  in such a way that the new level of  $\tilde{p}_1^F - \tilde{p}_2^F$  equals  $p_1^F - p_2^F$ , we can generate a new optimal contract identical to contract A, as desired. ■

After establishing Lemmas A.1, A.2 and A.3, we are ready to prove Proposition 4.

By Lemmas A.2 and A.3, we know that all optimal relational contracts satisfy  $p_1^I = \underline{U}_1 - \pi_A(a^{FB})$  and  $a_i^I = a_i^F = a^{FB}$  for  $i \in \{1,2\}$ . To find  $\{p_1^F, p_2^I, p_2^F\}$ , recall that: (a) both  $(IC_2^{act})$  and  $(IC_p^{pay})$  are binding by Lemma A.3; and (b)  $p_1^F \geq p_2^F$  by Lemma A.2 and the previous result that  $a_1^F = a_2^F = a^{FB}$ . Recall also that

$$U_p^{spot} = \begin{cases} 2S(a^{FB}) - 2\underline{U}_1 & \text{if } \alpha^s \alpha^d \geq 1, \\ 2S(a^{FB}) - \underline{U}_1 - \frac{U_2 + \alpha^d \underline{U}_1}{1 + \alpha^d} - \alpha^s \alpha^d \left( \underline{U}_1 - \frac{U_2 + \alpha^d \underline{U}_1}{1 + \alpha^d} \right) & \text{if } \alpha^s \alpha^d \leq 1, \end{cases}$$

by Proposition 2.

Consider first the case in which  $\alpha^s \alpha^d > 1$ . Observe that  $p_1^F = p_2^F$  by Lemma A.4. Observe also that  $U_2^{pun} = U_2^R(\cdot) = \pi_A(a^{FB}) + p_2^I$ , as in this case the principal's optimal relational contract is also the equilibrium that minimizes agent 2's welfare. Combining  $(IC_2^{act})$  with  $(IC_p^{pay})$ , and using the two previous observations, we obtain  $p_1^F = p_2^I = p_2^F = \underline{U}_1 - \pi_A(a^{FB})$ , which, as desired, coincides with the outcome under spot transactions described in Proposition 2.

Consider now the case in which  $\alpha^s \alpha^d \leq 1$ . In what follows, let  $\underline{\delta}_* \equiv \frac{1 - \alpha^s \alpha^d}{1 + \alpha^d + \alpha^d(1 + \alpha^s)}$ .

Part (1). Suppose that  $\alpha^s \alpha^d \leq 1$  and  $\delta \in (0, \underline{\delta}_*)$ . From Lemma A.4, we know that  $(PC_1)$  must bind. After Combining  $(PC_1)$ ,  $(IC_2^{act})$  and  $(IC_p^{pay})$ , and doing some calculations, we obtain

$p_1^F = \underline{U}_1 - \pi_A(a^{FB})$  and  $p_2^I = p_2^F = \left(\frac{U_2 + \alpha^d U_1}{1 + \alpha^d}\right) - \pi_A(a^{FB})$ , which coincides with the outcome under spot transactions described in Proposition 2.

Part (2). Suppose next that  $\alpha^s \alpha^d \leq 1$  and  $\delta \in (\underline{\delta}_*, \bar{\delta}^*)$ . Observe that  $p_1^F = p_2^F$  by Lemma A.4. Observe also that  $U_2^{pun} = \underline{U}_2$ , for the optimal spot contract, which is also an equilibrium of the repeated game, yields agent 2 his lowest possible equilibrium payoff of  $\underline{U}_2$  (see Proposition 2). Combining  $(IC_2^{act})$  with  $(IC_P^{pay})$ , and using the two previous observations, we obtain the values of  $\{p_1^F, p_2^I, p_1^F\}$  described in (2).

Part (3). Finally, suppose that  $\alpha^s \alpha^d \leq 1$  and  $\delta = \underline{\delta}_*$ . Observe that the following is true under both the optimal spot contract and the relational contract described in (2): (a) formal payments satisfy  $p_2^F \geq p_2^I$ ; and (b) not only are Eqs.  $(PC_2)$ ,  $(IC_2^{act})$  and  $(IC_P^{pay})$  satisfied, but  $(IC_2^{act})$  and  $(IC_P^{pay})$  are both binding. Observe also that provided  $p_2^F \geq p_2^I$ , Eqs.  $(PC_2)$ ,  $(IC_2^{act})$  and  $(IC_P^{pay})$  are all linear in  $p_1^I$ ,  $p_1^F$ ,  $p_2^I$  and  $p_1^F$ . From the previous observations, we can conclude that after setting  $a_i^I = a_i^F = a^{FB}$ ,  $i \in \{1, 2\}$ , any convex combination between the optimal spot contract and the relational contract described in (2) will feature  $p_1^I = \underline{U}_1 - \pi_A(a^{FB})$ , together with values of  $\{p_1^F, p_2^I, p_1^F\}$  for which Eqs.  $(IC_2^{act})$  and  $(IC_P^{pay})$  are both binding and both  $(PC_2)$  and  $p_1^F \geq p_2^I$  are satisfied. The result then follows from the last part of Lemma A.4, which states that any contract with such features is optimal. ■

**Proof of Proposition 5.**

Let  $p_1^F(\delta)$  and  $p_2^F(\delta)$  be the formally agreed-upon payments under an optimal relation contract when the discount factor is  $\delta$ . From Propositions 3 and 4 and the corollary of Proposition 3, we know that if there is a small cost of drafting a formal contract but the same contract can be used for the two agents at no additional costs, then: (a)  $|p_1^F(\delta) - p_2^F(\delta)|$  is (weakly) decreasing on the interval  $[0, \bar{\delta}^*]$ ; and (b)  $|p_1^F(\delta) - p_2^F(\delta)| = 0$  on the interval  $[\bar{\delta}^*, 1]$ . The result then follows directly from our definition of *formal pay compression*:  $|p_1^F - p_2^F|$ . ■

**Proof of Proposition 6.**

Let  $U_p^O \equiv 2S(\alpha^{FB}) - \underline{U}_1 - \underline{U}_2 - k$  denote the principal's payoff from outsourcing one of the tasks, and let  $U_p^R(\delta, \alpha^d)$  denote the principal's payoff without outsourcing and given an optimal relational contract when the discount factor is  $\delta$  and the degree of social comparisons is  $\alpha^d$ . As mentioned in the main text, suppose that  $k$  is not “too large”—otherwise it would never be optimal to outsource. More precisely, suppose that  $U_p^O > U_p^{spot}$  for  $\alpha^d = 1$ . Suppose also that if the principal is indifferent between producing in house or outsourcing, he would produce in house.

By Proposition 2, we know that  $U_p^{spot}$  (strictly) decreases with  $\alpha^d$  on the interval  $[0, \min\{1, \frac{1}{\alpha^s}\}]$  and remains constant on the interval  $[\min\{1, \frac{1}{\alpha^s}\}, 1]$ . Thus, since  $U_p^O > U_p^{spot}$  when  $\alpha^d = 1$  by assumption, and  $U_p^O < U_p^{spot}$  when  $\alpha^d = 0$  by proposition 1, it follows that there exists a  $\bar{\alpha}^d \in (0, \min\{1, \frac{1}{\alpha^s}\})$  such that  $U_p^{spot}(\alpha^d) = U_p^O$  for  $\alpha^d = \bar{\alpha}^d$ ,  $U_p^{spot}(\alpha^d) > U_p^O$  for  $\alpha^d < \bar{\alpha}^d$ , and  $U_p^{spot}(\alpha^d) < U_p^O$  for  $\alpha^d > \bar{\alpha}^d$ .

Consider first the case in which  $\alpha^d \in [0, \bar{\alpha}^d)$ , so that  $U_p^{spot}(\alpha^d) > U_p^O$ . Notice that  $U_p^R(\delta, \alpha^d) \geq U_p^{spot}(\alpha^d)$  for any  $\delta$  because optimality of the relational contract implies that the principal prefers it to the spot contract. Then, it follows that  $U_p^R(\delta, \alpha^d) > U_p^O$  for any  $\delta \in [0, 1]$ , and therefore that  $\delta^O = 0$ .

Consider next the case in which  $\alpha^d \in [\bar{\alpha}^d, \min\{1, \frac{1}{\alpha^s}\}]$ , which, since  $\bar{\alpha}^d \alpha^s < 1$  by definition, is never an empty interval. By Propositions 3 and 4, we know that

$$(1) \quad U_p^R(\delta) = \begin{cases} U_p^{spot} & \text{for } \delta \leq \underline{\delta}^*, \\ 2S(a^{FB}) - 2\underline{U}_1 + \delta(\underline{U}_1 - \underline{U}_2) \left[ 1 + \frac{\alpha^d(1 + \alpha^s)}{1 + \alpha^d} \right] & \text{for } \delta \in [\underline{\delta}^*, \bar{\delta}^*], \\ 2S(a^{FB}) - \underline{U}_1 - \underline{U}_2 & \text{for } \delta \in [\bar{\delta}^*, 1], \end{cases}$$

where  $\underline{\delta}^* \equiv \frac{1 - \alpha^s \alpha^d}{1 + \alpha^d + \alpha^d(1 + \alpha^s)}$  and  $\bar{\delta}^* \equiv \frac{1 + \alpha^d}{1 + \alpha^d + \alpha^d(1 + \alpha^s)}$ . There are two relevant subcases. Suppose first that  $\alpha^d = \bar{\alpha}^d$ . Since  $U_p^{spot} = U_p^O$  by the definition of  $\bar{\alpha}^d$ , it follows that  $U_p^R(\delta) = U_p^O$  for  $\delta \leq \underline{\delta}^*$  and  $U_p^R(\delta) > U_p^O$  for  $\delta > \underline{\delta}^*$ . Because the principal would produce in house if he is indifferent between producing in house and outsourcing, it follows that  $\delta^O = 0$ . Suppose next that  $\alpha^d \in (\bar{\alpha}^d, \min\{1, \frac{1}{\alpha^s}\}]$ . Notice that  $U_p^{spot} < U_p^O$  by the definition of  $\bar{\alpha}^d$  and the assumption that  $\alpha^d > \bar{\alpha}^d$ . Using (1), it is straightforward to check that  $U_p^R(\delta)$  is continuous in  $\delta$  and thus that  $U_p^R(\delta) = U_p^O$  for some  $\delta \in (\underline{\delta}^*, \bar{\delta}^*)$ . Let  $\delta^O$  be such a value, which, after some calculations, is given by  $\delta^O = \left[ 1 - \frac{k}{\underline{U}_1 - \underline{U}_2} \right] \left[ \frac{1 + \alpha^d}{1 + \alpha^d + \alpha^d(1 + \alpha^s)} \right]$ . Since  $U_p^R(\delta)$  is always (weakly) increasing and strictly increasing on the interval  $[\underline{\delta}^*, \bar{\delta}^*]$ , it follows that  $U_p^R(\delta) < U_p^O$  for  $\delta < \delta^O$  and  $U_p^R(\delta, \alpha^d) > U_p^O$  for  $\delta > \delta^O$ , as desired.



Finally, consider the case in which  $\alpha^d \in \left(\min\left\{1, \frac{1}{\alpha^s}\right\}, 1\right]$ . By Propositions 3 and 4, we know that

$$U_P^R(\delta) = \begin{cases} U_P^{spot} & \text{for } \delta < \frac{1}{2}, \\ 2S(a^{FB}) - \underline{U}_1 - \underline{U}_2 & \text{for } \delta \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

Moreover, since  $\alpha^d > \bar{\alpha}^d$ , we know that  $U_P^{spot} < U_P^O$ . Thus,  $\delta^O = \frac{1}{2}$ .

To summarize, we have shown that

$$\delta^O = \begin{cases} 0 & \text{for } \alpha^d \in [0, \bar{\alpha}^d], \\ \left[1 - \frac{k}{\underline{U}_1 - \underline{U}_2}\right] \left[\frac{1 + \alpha^d}{1 + \alpha^d + \alpha^d(1 + \alpha^s)}\right] & \text{for } \alpha^d \in \left(\bar{\alpha}^d, \min\left\{1, \frac{1}{\alpha^s}\right\}\right], \\ \frac{1}{2} & \text{for } \alpha^d \in \left(\min\left\{1, \frac{1}{\alpha^s}\right\}, 1\right]. \end{cases}$$

To complete the proof, let  $\varphi(\alpha^d) = \left[1 - \frac{k}{\underline{U}_1 - \underline{U}_2}\right] \left[\frac{1 + \alpha^d}{1 + \alpha^d + \alpha^d(1 + \alpha^s)}\right]$ . After some calculations, we can show that  $\varphi'(\alpha^d) < 0$  and  $\varphi(\bar{\alpha}^d) = \frac{1}{2} \left[1 - \frac{k}{\underline{U}_1 - \underline{U}_2}\right] < \frac{1}{2}$ , as desired. ■

### Proof of Proposition 7.

Let  $a_i^F(\delta)$  and  $p_i^F(\delta)$  be the formally agreed-upon action and payment offered to agent  $i$  under an optimal relational contract when the discount factor is  $\delta$ . From Propositions 3 and 4, we know that  $|\pi_A[a_1^F(\delta)] + p_1^F(\delta) - \pi_A[a_1^F(\delta)] + p_2^F(\delta)|$  is (weakly) decreasing on the interval  $[0, \bar{\delta}^*]$  and equal to zero on the interval  $[\bar{\delta}^*, 1]$ . The result then follows directly from our measure of *organizational conflict*:  $\alpha^d(1 + \alpha^s)|\pi_A(a_1^F) + p_1^F - \pi_A(a_2^F) + p_2^F|$ . ■