

## Relational Contracts and the Value of Relationships<sup>†</sup>

By MARINA HALAC\*

*This article studies optimal relational contracts when the value of the relationship between contracting parties is not commonly known. I consider a principal-agent setting where the principal has persistent private information about her outside option. I show that if the principal has the bargaining power, she wants to understate her outside option to provide strong incentives and then renege on promised payments, while if the uninformed agent has the bargaining power, the principal wants to overstate her outside option to capture more surplus. I characterize how information is revealed, how the relationship evolves, and how this depends on bargaining power. (JEL C78, D82, D83, D86)*

When a principal and an agent engage in a repeated, open-ended relationship, they may be able to complement court-enforced contracts with informal, self-enforced relational contracts. A contract is self-enforcing if the parties prefer to honor the terms of the agreement and continue with the relationship rather than to renege and end the relationship. The payments that each party can credibly promise to make, and thus the scope of incentive provision in a relational contract, are, in turn, limited by the value of the relationship.

But what happens if the contracting parties have incomplete information about the value of the relationship? In many situations, it seems unrealistic to assume that parties can perfectly observe how valuable the relationship is to the other party. Consider, for example, a repeated interaction between a firm and a labor union. Both parties may have bargaining power in determining worker compensation. The workers, however, may not be able to assess the benefit that the firm derives from the relationship with them. How large is the firm's payoff within the relationship relative to its opportunities outside the relationship? Could, for instance, the firm have plans to shut down the factory and start a new business in a remote location and, hence, not care about its reputation?

\* Graduate School of Business, Columbia University, 3022 Broadway, 616 Uris Hall, New York, NY 10027 (e-mail: mhalac@columbia.edu). This paper is a revised chapter of my PhD thesis at UC Berkeley. I am indebted to Ben Hermalin, Steve Tadelis, and Shachar Kariv for their continuous guidance, advice, and support. I thank Patrick Bolton, Bogac Celen, Wouter Dessein, Bob Gibbons, Johannes Hörner, Navin Kartik, Botond Köszegi, Jon Levin, Jin Li, Bentley MacLeod, Kristóf Madarasz, George Mailath, Jim Malcomson, Guillermo Noguera, Alessandro Pavan, Luis Rayo, Paolo Siconolfi, Joel Watson, seminar participants at Berkeley, Caltech, Chicago, Columbia, Duke-UNC, LSE, MIT, Northwestern, NYU, Toronto, UCLA, UCSD, UPenn, USC, WashU, and Yale, and two anonymous referees for helpful comments. The financial support of the Graduate School of Business at Columbia University is gratefully acknowledged.

<sup>†</sup> To view additional materials, visit the article page at <http://dx.doi.org/10.1257/aer.102.2.750>.

Uncertainty about a firm's prospects can render relational contracts ineffective. For example, when workers at Eastman Kodak Co. became concerned about the continuous restructuring and the possibility of a takeover, the company had to introduce a formal plan guaranteeing severance pay and other benefits to "remove the 'uncertainty factor' which demoralizes and demotivates employees..."<sup>1</sup> Firms understand the reputational cost of not complying with their relational contracts and often refrain from renegeing.<sup>2</sup> But there is also evidence that firms sometimes do renege on their promises. Some well-known examples are First Boston and Goldman Sachs, both of which unexpectedly cut discretionary end-of-year bonuses (Stewart 1993; Endlich 1999), and IBM, which abandoned its implicit agreement of no layoffs in the 1990s (Gibbons 2005).

A large literature studies the design of relational contracts but assumes that parties' commitment to the relationship is observable.<sup>3</sup> Thus, in this literature, parties promise the same payments in every period, and they always honor their promises. This article provides the first analysis of relational contracting when the value of the relationship is not commonly known. I show that a firm may want to make its workers believe that the value of the relationship is higher than it truly is, so that it can provide strong incentives and then renege on promised payments and walk away with a high benefit. On the other hand, if the workers have strong bargaining power, the firm may want to make them believe that the value of the relationship is not higher but lower than it is, so that the workers cannot extract the firm's rents.<sup>4</sup> I characterize how parties learn about the extent of their commitment, how information revelation affects the dynamics of the relationship, and how this depends on bargaining power.

To address these issues, I develop an infinite-horizon principal-agent model with nonverifiable outcomes and hidden action in which the principal has persistent private information about her outside option.<sup>5</sup> The principal can be a "low" type with a low outside option (high relationship value) or a "high" type with a high outside option (low relationship value).<sup>6</sup> In every period, one party makes a take-it-or-leave-it offer to the other, specifying an enforceable fixed wage and a discretionary output-contingent payment (a bonus if positive, a fine if negative) for the agent. If the offer is accepted, the agent chooses effort which generates output for the principal. Depending on the realized level of output, the principal or the agent then chooses whether to honor or renege on the promised contingent payment. If the offer is rejected, the parties receive their outside options.

As is well known, the set of equilibria in repeated games with incomplete information can be large. Limiting attention to equilibria that are subgame perfect and (constrained) Pareto optimal is often not sufficient to make predictions. Following a large strand of the literature, I thus focus on a subset of these equilibria by placing restrictions on what parties can do in every contingency. I assume the relationship

<sup>1</sup>The quote is from the firm's announcement of the plan in its employee newspaper. See Hymowitz (1990).

<sup>2</sup>Conlin (2001) provides examples of companies like Southwest Airlines which refrained from renegeing on their no-layoff policies even in times of crisis.

<sup>3</sup>Main references are MacLeod and Malcolmson (1989) and Levin (2003).

<sup>4</sup>Consistent with this idea, empirical evidence shows that firms with substantial implicit claims with workers are more likely to choose income-increasing accounting methods, while firms that face powerful unions try to make profits look smaller. See DeAngelo and DeAngelo (1991) and Bowen, DuCharme, and Shores (1995) among others.

<sup>5</sup>The case in which the agent is the privately informed party is very similar and is briefly discussed in Section IV.

<sup>6</sup>This formulation is equivalent to one in which the principal's outside option is observable but the principal has private information about a fixed component of her return within the relationship.

ends with positive probability if a default occurs (without loss of generality) but remains on the Pareto-optimal frontier otherwise. That is, I distinguish “cheating” from “tough bargaining”: I postulate that a failure to honor a payment may lead to conflict and breakup, while a rejection or an unexpected offer is not a credible excuse to walk away or impose similar punishments.

I start by showing that in equilibrium, the principal can never fully convey the extent of her commitment through her choice of contract; that is, through the contract that she offers or accepts. Intuitively, suppose the low-outside-option type wanted to choose a contract that communicates her strong commitment and thus allows her to provide strong incentives to the agent. Then such a contract would also be attractive, and even more so to the high-outside-option type, who could provide strong incentives and earn a rent by renegeing on the promised payments. Similarly, if the high type wanted to choose a contract that conveys a low commitment and gives her a large share of the surplus when the agent makes the offers, the low type would want to do that as well. As a result, relational contracting is always less efficient than under symmetric information, and the principal’s type can be revealed only if one type either reneges on a payment or rejects an offer.

I concentrate on two settings, one where the informed principal has the bargaining power and another one where the uninformed agent does. Unlike under symmetric information, here the allocation of bargaining power plays a central role, as it essentially determines the source of the inefficiency.

When the informed principal has the bargaining power, the high-outside-option type wants to mimic the low-outside-option type—she wants to pretend to be committed to the relationship to provide high-powered incentives. A Pareto-optimal contract involves either pooling, where both principal types behave as if their type were high and offer weak incentives, or separation, where both types offer strong incentives and at some point the high type reneges. Because the agent must be compensated for the risk of default, separation of types is optimal only if the probability of a low type is sufficiently high and thus this risk is sufficiently low. The analysis describes how fast separation is induced and how the relationship changes as information is revealed. Incentives and effort may increase over time, although a default and breakup also become more likely.

Relationship dynamics are different when the uninformed agent has the bargaining power. Here the low-outside-option type wants to mimic the high-outside-option type—she wants to “play hard to get” to capture a larger share of the surplus. Because the low type can ensure herself a larger share in the future by rejecting the agent’s offer in the present, the Pareto-optimal contract often involves pooling, with the agent being unable to extract the low type’s rents. Separation of types is feasible only if the future is heavily discounted, and in this case it is fully induced via the high type’s rejecting immediately in the first period.

By studying how parties learn about each other and adapt to new information, this paper sheds light on how agency relationships evolve. The results can help us understand why incentives and effort follow different paths in different environments, why some relationships are prone to default and breakup at certain stages, and why others last for a long time. In the employment example, if the bargaining power is on the firm’s side, incentives may build slowly and the relationship become more productive over time as workers grow more convinced that the firm is strongly

committed, until either the firm reneges and the relationship ends, or incentives become stationary. If the bargaining power is instead on the workers' side, the workers' initial offer may force the firm to permanently raise wages or suffer a strike, but the relationship's path remains stationary thereafter.

Section I discusses the related literature. Section II presents the model and the symmetric information benchmark. Section III introduces asymmetric information and characterizes optimal relational contracts. Section IV considers some extensions of the model and discusses the restrictions on strategies. Section V concludes. Proofs not given in the text can be found in the Appendix.

## I. Related Literature

Levin (2003) studies relational contracts under private information about the agent's effort choice and the agent's (time-specific) cost of effort. Levin (2003), MacLeod (2003), and Fuchs (2007) consider asymmetric information about the outcome variable. This article is complementary in that it shows how the form of the relational contract changes when commitment to the relationship is uncertain. As noted, I consider persistent private information about outside payoffs and study contracting under different bargaining protocols. The model then emphasizes elements that are naturally absent from these papers, such as how information is revealed over time, how this affects incentives and surplus, and how this depends on bargaining power.<sup>7</sup>

My analysis of information revelation is related to the literature on reputation building (e.g., Sobel 1985; Ghosh and Ray 1996; Kranton 1996; Watson 1999, 2002).<sup>8</sup> In that literature, when parties cannot observe whether others are the "cooperative" or "noncooperative" type, they optimally increase the stakes of the relationship gradually over time. Similar paths arise in many of my model's equilibria. Unlike this earlier literature, however, here I consider a principal-agent setting where transfers between parties are allowed and incentive provision plays a role. Additionally, in this literature, cooperative types never want to cheat, while noncooperative types always do, so no trade occurs if a player is known to be noncooperative. I take a different view. In my model, *all* types cooperate in some way under symmetric information, and *all* might wish to misbehave in some way under asymmetric information. I believe this is essential to understanding situations with many types, where some of these types may not be very different from each other. In Section IV, I discuss some implications of these different modeling assumptions.

Another related literature considers the ratchet effect (see Freixas, Guesnerie, and Tirole 1985; Laffont and Tirole 1988). This literature shows that if an informed party reveals herself to have high productivity or valuation for a good, the uninformed party will recontract and extract the informed party's rents in subsequent periods.<sup>9</sup> This ratchet effect is present in the model below, but its relevance as a determinant of final outcomes depends on the bargaining power structure.

<sup>7</sup>These elements are not present in Bull (1987) either, which considers a model with "honest" and "dishonest" types but restricts the former to always offer the same contract and the latter to always renege with probability one.

<sup>8</sup>In studying how incentive provision reveals information, the paper also relates to various models where the contract can be a signal, like Aghion and Bolton (1987), Aghion and Hermalin (1990), Spier (1992), and Hermalin (2002).

<sup>9</sup>Also related is the literature on countervailing incentives (Lewis and Sappington 1989), which shows that, as in this article, a party may want to overstate or understate her private information depending on its realization.

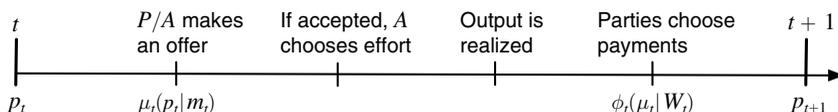


FIGURE 1. TIMING AND BELIEFS

The role played by the bargaining power structure has been studied in various papers. The closest to this work is perhaps Genicot and Ray (2006). In a lender-borrower model with self-enforcing credit contracts, the authors analyze how an (observable) increase in the borrower's outside option affects the contract and the borrower's payoff, and how this depends on the allocation of bargaining power.<sup>10</sup> Their results are consistent with those of this paper if symmetric information is imposed.

Finally, the need to think more carefully about the standard assumption that parties' outside options are observable has been stressed in other contexts, including the literature on the boundaries of the firm (see Holmstrom 1999; Williamson 2000; and Schmitz 2006).

## II. The Model

### A. Setup

Consider a risk-neutral principal (she) who can trade with a risk-neutral agent (he) in periods  $t = 0, 1, \dots$ . Both parties have the same discount factor  $\delta \in (0, 1)$ . The principal has private information about her outside option.<sup>11</sup> The principal's per-period outside option (reservation value) is  $r_\theta$ , where  $\theta \in \{\ell, h\}$  is the principal's constant type and  $r_\ell < r_h$ . The agent's outside option is  $r_A$ .

The sequence of events is shown in Figure 1. At the beginning of a period  $t$ , one party makes a take-it-or-leave-it offer to the other. The superscript  $i \in \{A, P\}$  indicates the party that makes the offer in the current period. The other party can then accept or reject this offer. If the offer is accepted, the agent chooses effort  $e_t \in [0, \bar{e}]$  at private cost  $c(e_t)$ , where  $c(0) = 0$ ,  $c'(\cdot) > 0$ ,  $c''(\cdot) > 0$ ,  $c'(\bar{e}) = \infty$ . This effort choice is the agent's private information. The agent's effort generates, in a stochastic manner, output,  $y_t \in \{\underline{y}, \bar{y}\}$ , for the principal. The probability that  $y_t = \bar{y}$  given effort  $e$  is  $f(e) \in (0, 1)$ , where  $f'(\cdot) > 0$ ,  $f''(\cdot) \leq 0$ .<sup>12</sup> Output is observed by both the principal and the agent, but it cannot be verified by a third party.

An offer is a compensation package for the agent, with the payoff to the principal being the difference between the surplus generated by the relationship and the payment to the agent. The agent's compensation consists of a fixed wage  $w_t$  and an

These incentives arise in a standard agency setting when both the agent's marginal utility and her outside option are assumed to depend on the agent's type. In the model below, instead, only outside options are type dependent; countervailing incentives arise endogenously due to the structure of self-enforcing contracts.

<sup>10</sup> See also Baker, Gibbons, and Murphy (2002).

<sup>11</sup> Section IV discusses the case where this private information is on the agent's side.

<sup>12</sup> These conditions imply that the first-order approach for deriving the optimal contract is valid: the agent's incentive compatibility constraint for effort can be replaced with the first-order condition of his optimization problem.

output-contingent bonus  $b_t$ , where, without loss of generality,  $b_t \in \{\underline{b}_t, \bar{b}_t\}$  corresponding, respectively, to  $\underline{y}, \bar{y}$ ;  $\underline{b}_t \leq 0$  and  $\bar{b}_t \geq 0$ . When not confusing, a contract  $\{w, b\}$  is simply denoted by  $b$ . The fixed wage is formally enforced, but, because output is nonverifiable, the bonus is not. If  $b_t > 0$ , the principal has the decision whether to honor or renege on the bonus payment at the end of period  $t$ ; if  $b_t < 0$ , the agent has this decision. (There is no limited liability.) Total compensation is denoted by  $W_t$ , where  $W_t = w_t + b_t$  if contingent payments are honored, and  $W_t = w_t$  if they are not. The expected surplus is  $s(e) \equiv \mathbb{E}_y[y - c | e]$ .

If a party rejects an offer, both the principal and the agent receive their outside options in the current period. If at any point the relationship ends, they receive their outside options in all periods from then on. I assume that for  $\theta \in \{\ell, h\}$ ,  $\max_e s(e) > r_A + r_\theta \geq s(0)$ .

The updating of beliefs, also shown in Figure 1, is as follows. At the beginning of period  $t$ , the agent believes that he is facing a low type with probability  $p_t$ , where  $p_0$  is the prior belief. Within a period  $t$ , there are two instances at which the agent updates his belief. First, depending on which party makes the offer, the agent observes either the offer or the participation decision of the principal. Given this message  $m_t$  and the prior  $p_t$ , the agent updates his belief that the principal is a low type to  $\mu_t(p_t | m_t)$ . Next, if the offer is accepted and output is high, the agent observes whether the principal honors or reneges on the bonus. Given the principal's payment  $W_t$  and the prior  $\mu_t$ , the agent updates his belief to  $\phi(\mu_t | W_t)$ .

I multiply expected lifetime payoffs by  $(1 - \delta)$  to express them as per-period averages. The principal and agent's expected payoffs at time  $t$  are, respectively,

$$\pi_{\theta t} = (1 - \delta) \mathbb{E} \sum_{\tau=t}^{\infty} \delta^{\tau-t} [\mathbb{I}_\tau (y_\tau - W_\tau) + (1 - \mathbb{I}_\tau) r_\theta],$$

$$u_t = (1 - \delta) \mathbb{E} \sum_{\tau=t}^{\infty} \delta^{\tau-t} [\mathbb{I}_\tau (W_\tau - c(e_\tau)) + (1 - \mathbb{I}_\tau) r_A],$$

where  $\mathbb{I}$  is equal to one if parties trade and zero otherwise. The expected surplus is  $s_{\theta t} = \pi_{\theta t} + u_t$ .

A relational contract specifies, for each date  $t$  and every history of play up to date  $t$ , (i) a take-it-or-leave-it offer for the party proposing compensation; (ii) a participation decision for the party receiving the offer; (iii) an effort choice (conditional on participation) for the agent; (iv) output-contingent bonus payment decisions (conditional on participation and given the output realization) for the principal and the agent; and (v) conditional beliefs for the agent.

### B. Equilibrium Concepts

I characterize Pareto-optimal contracts using two equilibrium concepts, perfect public Bayesian equilibrium (PPBE) and weak Markov perfect Bayesian equilibrium (WMPBE). The online Appendix provides formal definitions of these concepts.

A PPBE is a set of public strategies and posterior beliefs such that the strategies form a Bayesian Nash equilibrium in every continuation game given the posterior

beliefs, and the beliefs are updated according to Bayes' rule whenever possible.<sup>13</sup> Following Fudenberg and Tirole (1991), beliefs are updated using Bayes' rule not only on the equilibrium path but also in continuation games reached with zero probability. Because Bayes' rule does not apply when beliefs are degenerate, an assumption such as Assumption 1 is needed (Rubinstein 1985):

*ASSUMPTION 1: If, at any point, the agent's posterior belief assigns probability one to a given type, then his beliefs continue to do so no matter what happens.*

Even under the requirements of PPBE, the set of Pareto-optimal equilibria is large. Inefficient levels of trade, as well as no-trade outcomes, can be supported in equilibrium, and thus be used by the parties as "punishments" for different behaviors. I concentrate on a subset of the Pareto-optimal equilibria by imposing restrictions on strategies and, hence, on these punishments:

*ASSUMPTION 2: If a party reneges on a payment at time  $t$ , the relationship ends with some probability  $1 - \gamma_{t+1} > 0$  at time  $t + 1$ , and continues on the Pareto-optimal frontier with probability  $\gamma_{t+1}$ . If no party reneges, the relationship continues on the Pareto-optimal frontier with probability one.*

In Section IV, I show that the first part of this assumption is without loss of generality and discuss the consequences of relaxing the second part.

Finally, for some of the results, I focus on a weak version of Markov perfect Bayesian equilibrium. Strategies are Markov if they depend only on payoff-relevant past events (including a player's own payoff-relevant private information). Consistent with Assumption 2, I am interested in a simple form of weak Markov strategies. I require that the parties' decisions to continue with the relationship in a given period depend only on payoff-relevant past events and the last period's history, and that, conditional on the relationship's continuing, their moves depend only on payoff-relevant past events. (Also, for the Markov restriction to be effective, I require that beliefs depend only on payoff-relevant past events.) The payoff-relevant information at the beginning of a period is given by the agent's conditional beliefs,  $\mu$ . It is then useful to define  $\pi_\theta^i(\mu, b)$  and  $u^i(\mu, b)$  as the expected payoffs for a type- $\theta$  principal and the agent when party  $i$  makes the offer, the agent's belief is  $\mu$ , and the contract offered is  $b$ .

I henceforth refer to PPBE and WMPBE as equilibria and Markov equilibria, respectively.

### *C. Symmetric Information Benchmark*

As a benchmark, consider the case of symmetric information about the value of the relationship, where the principal's type is observable. This case is equivalent to Levin (2003)'s moral hazard model. The only difference is that, here, either party

<sup>13</sup> PPBE is the natural extension of perfect public equilibrium (Fudenberg, Levine, and Maskin 1994) for dynamic Bayesian games. A strategy is public if it depends only on the public history of play and the player's own payoff-relevant private information. That is, the parties condition their moves on the sequence of past offers, output realizations, and participation and payment decisions, and on her type in the case of the principal. The agent does not condition on his past, unobserved effort decisions, as these do not affect beliefs nor continuation play in any way.

may propose compensation and, given Assumption 2, such party captures all the surplus from the relationship. Clearly, though, under symmetric information, this is only reflected in the fixed wage and does not affect the joint surplus.

It follows from Levin (2003) that the optimal contract is stationary: given  $\theta \in \{\ell, h\}$ , in every period on the equilibrium path,  $e_t = e(\theta)$ ,  $b_t = b(y, \theta)$ , and  $w_t = w^i(\theta)$ . That is, the agent’s effort rule and the output-contingent payments do not change over time, and the fixed wage changes depending only on which party makes the offer. Thus, the principal and agent’s expected payoffs are

$$\pi_\theta = \lambda \mathbb{E}_y[y - W^P(y, \theta) | e(\theta)] + (1 - \lambda) \mathbb{E}_y[y - W^A(y, \theta) | e(\theta)],$$

$$u = \lambda \mathbb{E}_y[W^P(y, \theta) - c | e(\theta)] + (1 - \lambda) \mathbb{E}_y[W^A(y, \theta) - c | e(\theta)],$$

where  $\lambda \in [0, 1]$  is the probability that the principal makes the take-it-or-leave-it offer in any given period, and  $1 - \lambda$  the probability that the agent makes this offer.

For the compensation schedule to be self-enforcing, neither party can wish to renege on a promised payment. Since, here, no party ever reneges in equilibrium, it is without loss to assume that a default leads to termination of the relationship with probability one, which is the worst punishment (Abreu 1988). A self-enforcing contract then satisfies

$$\delta \pi_\theta \geq (1 - \delta) \bar{b}(\theta) + \delta r_\theta,$$

$$\delta u \geq -(1 - \delta) \underline{b}(\theta) + \delta r_A.$$

Depending on the bargaining power distribution, the fixed wage is adjusted and slack transferred from one constraint to the other. The two conditions above can then be combined into a single enforcement constraint, (E). The optimal symmetric-information contract maximizes expected surplus subject to an incentive compatibility constraint for effort and (E). For  $\theta \in \{\ell, h\}$ ,

$$\max_{e(\cdot), \underline{b}(\cdot), \bar{b}(\cdot)} s = \mathbb{E}_y[y - c | e(\theta)]$$

subject to

$$(IC_A) \quad e(\theta) \in \arg \max_e f(e) (\bar{b}(\theta) - \underline{b}(\theta)) - c(e),$$

$$(E) \quad \frac{\delta}{1 - \delta} (s - r_\theta - r_A) \geq \bar{b}(\theta) - \underline{b}(\theta).$$

The solution is denoted by  $e_\theta, b_\theta$ . I also denote  $s(e_\theta) \equiv s_\theta, f(e_\theta) \equiv f_\theta$ . As is well known, under risk neutrality, there exists a simple contract that implements the first-best level of effort. However, when bonus payments are discretionary, the

enforcement constraint must also be satisfied. This constraint is tighter the lower the discount factor and the higher the parties' outside options. For the analysis to be interesting, I assume that for  $\theta \in \{\ell, h\}$ , parameters are such that (E) binds but some bonus scheme is enforceable. Given that the party making the offer captures all the surplus,  $b_\theta$  implements an effort  $e_\theta$  (below first-best level) by specifying

$$\bar{b}_\theta = \frac{\delta}{1 - \delta} \lambda(s_\theta - r_\theta - r_A),$$

$$\underline{b}_\theta = -\frac{\delta}{1 - \delta} (1 - \lambda)(s_\theta - r_\theta - r_A),$$

with fixed payments  $w_\theta^P = r_A - \mathbb{E}_y[b_\theta - c|e_\theta]$  and  $w_\theta^A = \mathbb{E}_y[y - b_\theta|e_\theta] - r_\theta$ . I call contract  $\{w_\theta, b_\theta\}$  the symmetric-information contract of type  $\theta$ .

The principal and agent's continuation values under symmetric information are  $\pi_\ell(1, b_\ell) \equiv \pi_\ell(1)$  and  $u(1, b_\ell) \equiv u(1)$  if  $\theta = \ell$ , and  $\pi_h(0, b_h) \equiv \pi_h(0)$  and  $u(0, b_h) \equiv u(0)$  if  $\theta = h$ , where

$$\pi_\ell(1) = \lambda(s_\ell - r_A) + (1 - \lambda)r_\ell, \quad \pi_h(0) = \lambda(s_h - r_A) + (1 - \lambda)r_h,$$

$$u(1) = \lambda r_A + (1 - \lambda)(s_\ell - r_\ell), \quad u(0) = \lambda r_A + (1 - \lambda)(s_h - r_h).$$

As a prelude to the analysis of the next section, note that the principal's continuation value may be decreasing or increasing in her type depending on the bargaining power structure. A lower outside option relaxes the enforcement constraint and allows the principal to provide higher-powered relational incentives. But a lower outside option also relaxes the principal's participation constraint and reduces her expected transfer when the agent makes the offer. As a result,  $\pi_\ell(1) > \pi_h(0)$  if  $\lambda$  is sufficiently high, but  $\pi_\ell(1) < \pi_h(0)$  otherwise. The agent's continuation value, on the other hand, is always nonincreasing in the principal's type:  $u(1) \geq u(0)$  for all  $\lambda \in [0, 1]$ .

### III. Optimal Incentive Contracts

Consider now the case of asymmetric information about the value of the relationship, where the principal's type is observed only by the principal. I start by showing that in equilibrium, low- and high-outside option types will never fully separate through their choice of contract; that is, by offering or accepting different contracts. This result simplifies the analysis of optimal incentive contracts, as it implies that types can separate only if one type either reneges on a payment or rejects an offer. Using this result, I next characterize optimal contracting when the principal has the bargaining power and when the agent does. I show that the allocation of bargaining power plays a central role in determining information revelation and the dynamics of the relationship.

### A. Revelation of Information through the Choice of Contract

Let a *contract-separating equilibrium* be an equilibrium where full information is revealed through the contract that the principal offers or accepts. That is, if the principal makes the offer, the two types offer different contracts with probability one; if the agent makes the offer and he offers more than one contract, the two types accept different contracts with probability one.

**PROPOSITION 1:** *Regardless of the bargaining protocol, a contract-separating equilibrium does not exist.*

**PROOF:**

Suppose by contradiction that there exists an equilibrium where  $\ell$  and  $h$  offer or accept different contracts in some period  $t$ . Let  $b_1$  be  $\ell$ 's contract and  $b_2$   $h$ 's contract. By Assumption 2, after this period  $t$ , the continuation play is the symmetric-information equilibrium, with payoffs  $\pi_\ell(1), u(1)$  if  $\theta = \ell$ , and  $\pi_h(0), u(0)$  if  $\theta = h$ . Let  $e_1, e_2$  be the effort levels that solve the agent's incentive compatibility constraint given  $b_1, b_2$  and the agent's beliefs. Suppose first that if the principal makes the take-it-or-leave-it offer, both  $b_1$  and  $b_2$  are such that the agent accepts, so in equilibrium all parties honor the payments and the relationship ends if a party reneges. Then, omitting time subscripts,

$$\pi_\ell^i(1, b_1) = (1 - \delta)\mathbb{E}_y[y - W_1^i | e_1] + \delta\pi_\ell(1),$$

$$\pi_h^i(0, b_2) = (1 - \delta)\mathbb{E}_y[y - W_2^i | e_2] + \delta\pi_h(0).$$

Now if  $\ell$  deviates to  $b_2$ , she obtains

$$(1) \quad \pi_\ell^i(0, b_2) = (1 - \delta)\mathbb{E}_y[y - W_2^i | e_2] + \delta\pi_\ell(0) = \pi_h^i(0, b_2),$$

where  $\pi_\ell(0)$  is  $\ell$ 's continuation value when the agent's posterior belief is zero, and where the last equality follows from  $\pi_\ell(0) = \pi_h(0)$ . If  $h$  deviates to  $b_1$ , she obtains

$$\begin{aligned} \pi_h^i(1, b_1) &= (1 - \delta)\mathbb{E}_y[y - W_1^i | e_1] \\ &+ \max\{\delta\pi_h(1), f(e_1)[(1 - \delta)\bar{b}_1 + \delta r_h] + (1 - f(e_1))\delta\pi_h(1)\}, \end{aligned}$$

where  $\pi_h(1)$  is  $h$ 's continuation value when the agent's posterior belief is one. Given such belief,  $h$ 's optimal strategy is to offer  $b_\ell$  when the principal makes the offer and

renege on bonus payments  $\bar{b}_\ell$  if output is  $\bar{y}$ , and to reject when the agent makes the offer. Thus,

$$\begin{aligned}\pi_h(1) &= \frac{\lambda(1-\delta)(s_\ell - r_A) + \lambda f_\ell[(1-\delta)\bar{b}_\ell + \delta r_h] + (1-\lambda)(1-\delta)r_h}{1 - \lambda(1-f_\ell)\delta - (1-\lambda)\delta} \\ &= \frac{\lambda(1-\delta)(s_\ell - r_A) + \lambda f_\ell \delta(\pi_\ell(1) - r_\ell + r_h) + (1-\lambda)(1-\delta)r_h}{1 - \lambda(1-f_\ell)\delta - (1-\lambda)\delta}.\end{aligned}$$

Note that  $h$  gains from the option of being able to renege and reject. Hence,

$$\pi_h(1) - \pi_\ell(1) = \left[ 1 - \frac{\lambda(1-\delta)}{1 - \delta(1 - \lambda f_\ell)} \right] (r_h - r_\ell) > 0.$$

As a result,

$$(2) \quad \pi_h^i(1, b_1) > \pi_\ell^i(1, b_1).$$

A contract-separating equilibrium exists if and only if for  $\delta \in (0, 1)$ ,  $\lambda \in [0, 1]$ , we have that  $\pi_\ell^i(1, b_1) \geq \pi_\ell^i(0, b_2)$  and  $\pi_h^i(0, b_2) \geq \pi_h^i(1, b_1)$ . But equations (1) and (2) imply that one of these two incentive compatibility constraints is always violated. Contradiction.

Suppose next that if the principal makes the offer,  $b_1$  or  $b_2$  or both are such that the agent rejects. If  $b_1$  is rejected, the argument above is strengthened, since the difference between  $h$  and  $\ell$ 's expected payoffs under  $b_1$  increases. If  $b_2$  is rejected,  $h$  has incentives to deviate to  $b_h$ , which (by Assumption 2) is always accepted and gives  $h$  a continuation value at least as high as  $b_2$ .

The proof applies to  $\lambda \in [0, 1]$  and any other bargaining protocol such as alternating offers.

Proposition 1 says that the principal's choice of contract cannot be used to fully signal or screen the principal's type. Intuitively, if the low type chooses the high type's contract, her payoff is the same as the high type's, while if the high type chooses the low type's contract, her payoff is strictly larger. The latter is due to the fact that the high type enjoys a positive option to behave opportunistically when she chooses the low type's contract; namely, once the agent is convinced that the principal's type is low, the high type can earn a rent by renegeing on the promised payments when output is high and rejecting the agent's offer when the agent proposes compensation.

This result implies that incentive provision is always inefficient relative to the symmetric information case, and full separation of types must occur either through default—one type reneges while the other honors the promised bonus—or through rejection—one type rejects while the other accepts the agent's offer.

### B. When the Informed Party Has the Bargaining Power

Consider a setting in which the informed principal has the bargaining power; i.e.,  $\lambda = 1$ . To simplify the analysis, and without loss of generality, I make the following assumption:

**ASSUMPTION 3:** *The agent's belief about the principal's type at any point is independent of the fixed wages proposed by the principal up to that point.*

Assumption 3 rules out equilibria where the agent's beliefs are used as a threat to increase the equilibrium fixed wage. I consider such equilibria "unreasonable" in light of Proposition 1, as it is then commonly known that, given a bonus scheme, any type would prefer to offer a wage that gives her a higher surplus.<sup>14</sup>

Allocating the bargaining power to the principal has several direct implications for contracts and incentives. First, because the relationship has no value for the agent, no payments from the agent to the principal can be enforced. Hence,  $\underline{b} = 0$  in any contract. Second, because the agent never proposes compensation, separation of types cannot occur through rejection, and full separation must involve default. Finally, in this setting, the low-outside option type has no incentives to misrepresent her type, whereas the high type wants to pretend to be a low type.

An optimal contract may induce pooling or separation of types. A key point here is that (given Assumption 2) the principal can always choose to act as if her type were high and obtain the high type's symmetric-information payoff  $\pi_h(0)$ . Hence, any pooling equilibrium implements the high type's symmetric-information contract  $b_h$  in every period, as this is the best the principal can do when no information is revealed. Moreover, if a separating equilibrium exists, it gives both types a payoff at least as large as  $\pi_h(0)$ , and thus Pareto-dominates the pooling equilibrium.

To study separation of types, I start by considering contract-pooling Markov equilibria, where both types offer the same contract and thus information revelation occurs only through default. The agent in this case updates his belief only when output is high and the principal decides whether to honor or renege on the bonus payments. I denote the agent's updated belief that the principal is a low type when the principal honors by  $\phi(p|w + \bar{b}) \equiv p_b$ ,  $\phi(p_b|w + \bar{b}) \equiv p_{bb}$ , and so on.

The first step to characterizing these equilibria is to determine the conditions for default. Suppose that, given belief  $p$ , a contract  $b(p)$  is implemented and the principal's type is fully revealed through default. Because the high type has a higher outside opportunity than the low type, it must be the high type who reneges, while the low type honors. This means that if a default occurs and, with some probability  $\gamma(p)$ , the relationship continues, the high type's symmetric-information contract is implemented from then on. If no default occurs, the low type's symmetric-information contract is implemented. Thus, the high type will indeed default only if

$$(D_h) \quad \delta\pi_h(1) \leq (1 - \delta)\bar{b}(p) + \delta[r_h + \gamma(p)(\pi_h(0) - r_h)],$$

<sup>14</sup>Different refinements proposed in the literature would rule out these equilibria. See, for example, Grossman and Perry (1986)'s perfect sequential equilibrium refinement and Assumption B-2 in Rubinstein (1985).

and the low type's enforcement constraint will indeed be satisfied only if

$$(E_\ell) \quad \delta\pi_\ell(1) \geq (1 - \delta)\bar{b}(p) + \delta[r_\ell + \gamma(p)(\pi_h(0) - r_\ell)].$$

Similar conditions apply if a contract induces partial revelation through default. In this case, the high type reneges on the bonus with some probability strictly between zero and one, so she must be indifferent between honoring and renegeing. If a default occurs and the relationship continues, the high type's symmetric-information contract is implemented; if no default occurs, the agent's belief increases from  $p$  to  $p_b$ , and a contract  $b(p_b)$  is implemented in the next period. Of course, at some point, the contract must induce the high type to renege with probability one—otherwise, the high type would not be indifferent between honoring and renegeing in previous periods.

The next step is to determine the conditions for the principal to offer a contract that induces default. As explained above, the principal's payoff must be at least as large as  $\pi_h(0)$ . Now, because the high type has a positive option to renege, her payoff is always larger than the low type's; thus, it is sufficient to find a contract that induces default and gives the low type a payoff larger than  $\pi_h(0)$ . Given belief  $p$ , let  $k_h(p)$  be the probability that the high type honors the bonus payments (keeps a promise) and  $\chi(p) = p + (1 - p)k_h(p)$  the probability that the agent assigns to the principal's honoring the payments. The updated belief when the principal honors is  $p_b = p/\chi(p)$ . Take probabilities  $k_h(p)$  and beliefs  $p = p_0, p_{0b}, p_{0bb}, \dots$ , such that  $k_h(p) \in [0, 1)$  for all  $p$ ,  $k_h(p) = 0$  for some  $p$ .<sup>15</sup> Given  $k_h(p)$ ,  $p_b = p/\chi(p)$ ,  $\pi_\ell(1, b) = \pi_\ell(1)$ , the low type's problem is to find  $b(p), e(p), \gamma(p)$  that maximize her expected payoff subject to an incentive compatibility constraint for the agent's effort,  $(IC_A)$ ; conditions for partial and full revelation through default as just discussed,  $(I_h)$ ,  $(D_h)$ , and  $(E_\ell)$ ; a revelation constraint so that the low type, and thus the high type, have no incentives to deviate to the pooling contract  $b_h$  in the future,  $(R_\ell)$ ; and the agent's participation constraint, which always binds and is incorporated in the low type's expected payoff below:

$$\begin{aligned} \max_{b(p), e(p), \gamma(p)} \quad & \pi_\ell(p_0, b) \\ = \quad & \frac{(1 - \delta)[s(e(p_0)) - r_A - f(e(p_0))(1 - \chi(p_0))\bar{b}(p_0)] + \delta f(e(p_0))\pi_\ell(p_{0b}, b)}{1 - \delta(1 - f(e(p_0)))} \end{aligned}$$

subject to

$$(IC_A) \quad e(p) \in \arg \max_e f(e)\chi(p)\bar{b}(p) - c(e) \quad \text{for } p = p_0, p_{0b}, \dots,$$

$$(I_h) \quad \bar{b}(p) = \frac{\delta}{1 - \delta} [\pi_h(p_b, b) - r_h - \gamma(p)(\pi_h(0) - r_h)] \quad \text{if } k_h(p) > 0,$$

<sup>15</sup>In a contract-pooling Markov equilibrium with revelation,  $k_h(p)$  is always strictly less than one. Otherwise, the agent's belief and thus the principal's strategy would not change in subsequent periods, so revelation would not occur.

$$(D_h) \quad \bar{b}(p) \geq \frac{\delta}{1 - \delta} [\pi_h(1) - r_h - \gamma(p)(\pi_h(0) - r_h)] \quad \text{if } k_h(p) = 0,$$

$$(E_\ell) \quad \bar{b}(p) \leq \frac{\delta}{1 - \delta} [\pi_\ell(p_b, b) - r_\ell - \gamma(p)(\pi_h(0) - r_\ell)] \quad \text{for } p = p_0, p_{0b}, \dots,$$

$$(R_\ell) \quad \pi_\ell(p, b) \geq \pi_h(0) \quad \text{for } p = p_{0b}, p_{0bb}, \dots,$$

where

$$\begin{aligned} \pi_h(p, b) = \pi_\ell(p, b) &+ \frac{f(e(p))\{(1 - \delta)\bar{b}(p) + \delta[r_h + \gamma(p)(\pi_h(0) - r_h) - \pi_\ell(p_b, b)]\}}{1 - \delta(1 - f(e(p)))}. \end{aligned}$$

Let  $b^*(p), e^*(p), \gamma^*(p)$  denote the solution. Then a contract-pooling Markov equilibrium with revelation (through default) exists if and only if there exist  $k_h^*(p) \in [0, 1]$ , such that for  $\chi^*(p) = p + (1 - p)k_h^*(p), p_b = p/\chi^*(p)$ , and  $\pi_\ell(1, b) = \pi_\ell(1), \pi_\ell(p_0, b^*) \geq \pi_h(0)$ , or equivalently,

$$\begin{aligned} (3) \quad (1 - \delta)(s(e^*(p_0)) - s_h) + \delta f(e^*(p_0))(\pi_\ell(p_{0b}, b^*) - \pi_h(0)) \\ \geq (1 - \delta)f(e^*(p_0))(1 - \chi^*(p_0))\bar{b}^*(p_0). \end{aligned}$$

The Appendix shows that this condition is in fact necessary for any separating equilibrium to exist. In particular, I show that the Markov restriction does not bind here, and that the principal’s payoff in a separating equilibrium cannot be increased by having some (partial) information revealed through the principal’s choice of contract.<sup>16</sup>

Condition (3) has a simple interpretation: revelation can be induced in equilibrium if and only if the benefits outweigh the costs for the low type. The benefits come from the ability to provide stronger incentives and generate a higher surplus as revelation occurs. The costs are in the form of a compensation to the agent for the risk of default. If the prior probability that the principal is a low type,  $p_0$ , is sufficiently high, the benefits indeed exceed the costs, and revelation of information is feasible. Otherwise, the low type prefers to offer the pooling contract  $b_h$  and induce no revelation.

**PROPOSITION 2:** *There exists  $\hat{p} \in (0, 1)$  such that if  $p_0 \geq \hat{p}$ , the Pareto-optimal equilibrium induces separation of types through default. If  $p_0 < \hat{p}$ , the Pareto-optimal equilibrium is pooling and implements  $b_h$  in every period.*

<sup>16</sup>The intuition for the latter is straightforward: if a contract-semiseparating equilibrium exists, it involves the high type’s offering some contract  $b'$  on which she reneges and the low type’s mixing between this and another, less steep contract  $b''$  and thus entails weaker incentives than an equilibrium where both types pool in  $b'$ .

I next characterize how revelation of information occurs on the equilibrium path: How fast is it induced? How do the agent's compensation, the agent's effort, and the probability of default change over time as revelation occurs? I focus on Markov equilibria.<sup>17</sup>

To study how fast revelation is induced, I look at how fast the probability that the high type honors becomes zero as a function of the agent's belief. If  $k_h(p_0) = 0$ , full revelation is induced immediately. If  $k_h(p_0) > 0$ , revelation is gradual: the high type initially honors with positive probability and reneges with probability one only after output has been high and the agent's belief updated. Revelation is faster the fewer times the agent updates before full revelation is induced.

It follows directly from condition (3) above that how fast revelation can be induced in equilibrium depends on how likely the relationship is to have a high value. If the agent's prior belief that the principal is a low type,  $p_0$ , is relatively high, the cost of revelation—the compensation to the agent for the risk of default—is low, and this condition can be satisfied with  $\chi(p_0) = p_0$ . That is, full revelation can be induced immediately. If, instead,  $p_0$  is relatively low (but still higher than  $\hat{p}$ ), a sequence of contracts that induce partial revelation and progressively increase the agent's belief must be implemented prior to inducing full revelation, so that the risk of default and hence the cost of revelation are reduced.

The fastest possible speed of revelation is also generally the one that maximizes the high type's payoff.<sup>18</sup> Intuitively, if revelation is delayed, the high type must be indifferent between honoring and reneging, so she is better off if a steeper contract on which she prefers to renege is implemented. The low type, in contrast, may prefer slower revelation. This is at first counterintuitive—since the high type imposes a negative externality on the low type, it seems that the low type would want to separate as fast as possible. However, it is exactly because of the externality that the low type may want to delay revelation. Because she bears the cost of revelation but internalizes only part of the benefits, if  $p_0$  is relatively low, the low type can gain from delaying the cost to the future while using the high type's future rent to sustain strong incentives in the present.

Given a speed of revelation, consider now how the relationship evolves as information is revealed. Along the revelation path, the high type must be indifferent between honoring and reneging on the bonus payments. This requires that the high type's expected payoff be increasing along this path. For a given belief  $p$ , the high type's payoff is

$$\pi_h(p, b) = \frac{(1 - \delta)(s(e(p)) - r_A) + f(e(p))\{(1 - \delta)\chi(p)\bar{b}(p) + \delta[r_h + \gamma(p)(\pi_h(0) - r_h)]\}}{1 - \delta(1 - f(e(p)))}$$

It is immediate that for  $\pi_h(p, b)$  to increase,  $\chi(p)\bar{b}(p)$  or  $\gamma(p)$  must increase. That is, the high type is willing to honor with positive probability today only if tomorrow she can renege on a larger effective bonus or with a lower punishment for

<sup>17</sup> One technical issue is that in a contract-pooling Markov equilibrium with revelation, the probability that the high type honors,  $k_h(p)$ , must be strictly less than one. On the other hand, an optimal contract may require  $k_h(p)$  to be as high as possible, so the set of Pareto-optimal Markov equilibria may not be well defined. To address this problem, I set an upper bound on  $k_h(p)$  and consider equilibria where this bound is arbitrarily close to one.

<sup>18</sup> Note that given a set of parameter values, there may exist different probabilities  $k_h(p)$ , with different associated speeds of revelation, for which condition (3) can be satisfied.

default. Now, in principle, either of these changes may be optimal for the low type. If the effective bonus increases, effort increases, but, because the bonus increases, the agent's loss in the event of default and thus his required compensation also increase. If, instead, the punishment for default falls, the high type reneges on a smaller bonus, so in this case both effort and the default risk compensation fall.<sup>19</sup>

By placing some additional structure on the model, however, the evolution of these variables can be pinned down. In particular, if parameters are such that the symmetric-information contracts are sufficiently far from first-best (that is,  $(\bar{y} - \underline{y}) - \bar{b}_\ell$  is sufficiently large), then the cost of reducing effort is sufficiently large relative to the benefits of reducing the default risk compensation, and thus lowering the bonus is never beneficial. Hence, under this condition, a Pareto-optimal contract always specifies the worst punishment for default ( $\gamma(p) = 0$ ), and an effective bonus that increases as revelation occurs. This in turn implies that the bonus and the agent's effort also increase over time, while the fixed wage goes down.

The relationship's productivity is then increasing along the revelation path under the aforementioned condition. The probability of a default and breakup, though, can also be shown to be optimally increasing. That is, while the probability that the principal is a high type ( $1 - p$ ) falls over time, the probability that the high type reneges ( $1 - k_h(p)$ ) increases in a larger proportion, so the probability that the principal reneges ( $1 - \chi(p) = (1 - p)(1 - k_h(p))$ ) increases. The logic is by now familiar: a contract that involves paying the agent a low compensation for default risk today relative to tomorrow reduces the overall cost of inducing revelation (for both types).

Summarizing,

**PROPOSITION 3:** *Consider  $p_0 \geq \hat{p}$ . Pareto-optimal Markov equilibria have the following features:*

- (i) *Information is revealed immediately for relatively high values of  $p_0$  (i.e., if  $p_0 \geq \bar{p}$  for  $\hat{p} < \bar{p} < 1$ ) and gradually for relatively low values (i.e., if  $p_0 < \underline{p}$  for  $\hat{p} < \underline{p} \leq \bar{p}$ ). The low type prefers slower revelation than the high type.*
- (ii) *There exists  $\Delta \in \mathbb{R}$  such that, if parameters are such that  $(\bar{y} - \underline{y}) - \bar{b}_\ell \geq \Delta$ , the bonus and the agent's effort increase along the revelation path, while the fixed wage decreases. Furthermore, the probability of default increases along this path.*

Consider the example from the introduction, where a firm is privately informed about the prospects of its relationship with workers relative to outside opportunities. The results in this section say that when these prospects are not very likely to be good, the firm acts as if they were indeed relatively bad, as proving otherwise would be too costly. The relationship then follows a stationary path with low bonus incentives and low surplus. Instead, when the firm's prospects are likely to be good, the firm promises the workers high bonuses, and progressively convinces them that the relationship is highly valuable by honoring its promises. Incentives and surplus

<sup>19</sup>Examples showing that either of these paths may arise in a Pareto-optimal equilibrium are available from the author upon request.

thus increase over time, until, at some point, either the firm reneges and the relationship ends, or the contract becomes stationary. Interestingly, the model shows that an improvement in a firm's outlook can then lead to a substantial increase in surplus, but also to an event of default that would not have occurred or would have been delayed otherwise.

### C. When the Uninformed Party Has the Bargaining Power

Consider now a setting in which the uninformed agent has the bargaining power; i.e.,  $\lambda = 0$ . This allocation has several direct implications for contracts and incentives. First, because the relationship has no value for the principal under symmetric information, no payments from the principal to the agent can be enforced. Hence,  $\bar{b} = 0$  in any contract.<sup>20</sup> Second, as a result, separation of types cannot occur through default, and full separation must involve rejection. Finally, in this setting, the high-outside option type has no incentives to misrepresent her type, whereas the low type wants to pretend to be a high type.

As before, an optimal contract may involve pooling or separation of types. If the agent induces pooling, he offers the symmetric-information contract of the high type, as this is the best the agent can do when no information is revealed. Thus, in every period, the agent proposes a wage  $w_h$  and bonus  $b_h$  such that the expected transfer to the principal is  $\mathbb{E}_y[y - w_h - b_h(y) | e_h] = r_h$ , and both types accept this offer.

A more interesting question is whether the agent can induce separation and when he wants to do so. I proceed as in the previous section. The first step is to determine the conditions for rejection. Suppose that, given belief  $p$ , the agent offers a contract with wage  $w(p)$  and bonus  $b(p)$  such that the principal's expected transfer upon acceptance is  $\mathbb{E}_y[y - w(p) - b(p, y) | e] \equiv r_p(p)$ , and the principal's type is fully revealed through rejection. Because the high type has a higher outside opportunity than the low type, it must be the high type who rejects, while the low type accepts. This means that if the offer is rejected, the high type's symmetric-information contract is implemented from then on. If the offer is accepted, the low type's symmetric-information contract is implemented. Thus, the high type will indeed reject only if

$$(1 - \delta)r_p(p) + \delta\pi_h(1) \leq (1 - \delta)r_h + \delta\pi_h(0),$$

and the low type will indeed accept only if

$$(1 - \delta)r_p(p) + \delta\pi_\ell(1) \geq (1 - \delta)r_\ell + \delta\pi_\ell(0).$$

<sup>20</sup>The argument is as follows. Suppose that a contract with bonus  $\bar{b} > 0$  is implemented. Then the high type reneges (in finite time), as the principal's expected payoff within the relationship is weakly lower than  $r_h$ . So it must be that the low type honors. Thus, if no default occurs, the agent learns that the principal's type is low and offers a contract that gives the principal an expected payoff of  $r_\ell$ . But then the low type also reneges, which gives a contradiction.

For  $\lambda = 0$ ,  $\pi_h(1) = \pi_h(0) = r_h$  and  $\pi_\ell(1) = r_\ell$ . Thus, these two conditions give that full revelation through rejection is feasible only if

$$(4) \quad \delta \leq \frac{1}{2}.$$

This condition is also necessary for partial revelation through rejection. In particular, if (4) does not hold, an equilibrium where the low type mixes between accepting and rejecting and the high type rejects does not exist.<sup>21</sup> A simple Coase conjecture-like argument (Coase 1972) shows why. Under such strategies, after observing a number of rejections, the agent's belief becomes low enough that, if the principal again rejects, the agent optimally offers the symmetric-information contract of the high type, with expected transfer to the principal  $r_h$ , from then on. But then, when that belief is reached, the low type rejects any offer that the high type rejects, so the agent offers a transfer  $r_h$  at that point. Continuing with this reasoning gives that the agent offers a transfer  $r_h$  in all periods.

The next step is to determine the conditions for the agent to offer a contract that induces rejection. Suppose that condition (4) holds. If the agent induces a rejection, he optimally offers a contract with wage  $w^{**}$  and bonus  $b^{**} = b_\ell$  such that the low type's participation constraint binds. That is, for  $\mathbb{E}_y[y - w^{**} - b_\ell(y) | e_\ell] \equiv r_P^{**}$ ,

$$(1 - \delta)r_P^{**} + \delta r_\ell = (1 - \delta)r_\ell + \delta r_h.$$

(Note that under condition (4), this contract is such that the high type rejects.) Hence, the agent is willing to induce full revelation through rejection if and only if  $u_A(p_0, b^{**}) \geq u(0)$ , or equivalently,

$$(5) \quad p_0[(1 - \delta)(s_\ell - r_P^{**}) + \delta(s_\ell - r_\ell) - (s_h - r_h)] \\ \geq (1 - p_0)(1 - \delta)(s_h - r_h - r_A).$$

This condition is also necessary for the agent to induce gradual revelation through rejection. In fact, if separation of types is optimal for the agent, then immediate separation in the first period is optimal. The reason is that, unlike the low type when the principal has the bargaining power, the agent here can internalize the externality. That is, the agent bears both the costs and the benefits of inducing revelation, so his expected payoff falls if the net gain that he obtains is postponed.

Condition (5) parallels condition (3): it says that revelation is induced in equilibrium only if the benefits outweigh the costs for the agent. The benefits come from the ability to implement stronger incentives and a higher wage if the principal is revealed to be a low type. The costs are in the form of forgone surplus due to rejection in the current period if the principal is revealed to be a high type. If the prior probability that the principal is a low type,  $p_0$ , is sufficiently high, the benefits

<sup>21</sup> An equilibrium where the low type accepts and the high type mixes between accepting and rejecting is not feasible either, as the low type would want to reject.

indeed exceed the costs, and revelation of information is optimal. Otherwise, the agent prefers to offer the pooling contract  $b_h$  and induce no revelation.

**PROPOSITION 4:** *There exists  $\hat{p} \in (0, 1)$  such that if  $p_0 \geq \hat{p}$  and  $\delta \leq 1/2$ , the Pareto-optimal equilibrium induces full separation of types through rejection in the first period. If  $p_0 < \hat{p}$  or  $\delta > 1/2$ , the Pareto-optimal equilibrium is pooling and implements  $b_h$  in every period.*

Going back to the employment example, this section shows that relationship dynamics are very different when the bargaining power is on the uninformed workers' side rather than the firm's side. Here, the workers may force the firm to either approve a permanent increase in wages or suffer a strike in the first period, but the path of the relationship remains stationary thereafter. Furthermore, the model suggests that the threat of a strike may often not be effective—in a repeated environment where parties care enough about the future, a firm is not willing to give up future rents to prevent a halt in production today. As a consequence, workers are unable to demand wage increases, and the relationship's path is one of low incentives, low surplus, and relatively low wages.

Although which party makes the contract offers is exogenously determined in the model, we see that “effective” bargaining strengths depend on other factors as well. In this setting, even though the uninformed party makes a take-it-or-leave-it offer in every period, he cannot extract the low type's rents if the discount factor is relatively high. In fact, a Coase-conjecture-like result can be shown to hold; namely, if the length of the time periods is made arbitrarily small, then the uninformed party is always unable to capture the low type's rents, regardless of the discount factor.

## IV. Discussion

### A. Extensions

Below I discuss some extensions and variations of the model. I consider each of these extensions separately, taking the model setup and assumptions as otherwise unchanged.

*Informed Agent.*—Suppose that the principal's outside option is known and equal to  $r_p$ , while the agent's outside option is his private information and equal to  $r_\theta$  for  $\theta \in \{\ell, h\}$ . If the uninformed principal has the bargaining power, the analysis is very similar to the one above where the principal is privately informed and the agent has the bargaining power. If the informed agent has the bargaining power, on the other hand, things are slightly more complicated. The reason is that the agent makes a hidden effort choice, which now may depend on the agent's type. In fact, if revelation is induced through default, a given bonus scheme gives weaker effort incentives to the high type than to the low type, as the high type reneges with positive probability when output is low. Thus, in equilibrium, the realized level of output now provides information about the agent's type.

Still, the main insights of the article remain the same. When the bargaining power is on the informed agent's side, the high type wants to pretend to be a low type to

offer a large contingent payment to the principal, increase the fixed wage, and then renege and walk away with a high payoff when output is low. When the bargaining power is on the uninformed principal's side, the low type wants to pretend to be a high type to signal a high outside payoff and receive a high wage from the principal. The requirements for separation of types are similar to those found above, and separation always involves a default or a rejection.

*Intermediate Allocations of Bargaining Power.*—Suppose that both the principal and the agent have positive bargaining power. In the model, this means that, in any period, either party may make the offer with positive probability; i.e.,  $\lambda \in (0, 1)$ . More realistically, this may also be thought of as a setting where bargaining powers are determined after an initial stage, so that with some probability the principal makes the offer in all future periods, and with some probability the agent does.

This setting is more complex than those above. Payments from the principal to the agent and from the agent to the principal can both be enforced, and information may be revealed through default or rejection. Moreover, both types may want to misrepresent their types and, indeed, they may want to do so simultaneously. A characterization of the Pareto-optimal equilibria is then more difficult, especially because of the many off-the-equilibrium-path beliefs that must be considered.<sup>22</sup>

Yet, this setting may be useful to study how information revelation changes as bargaining power is shifted from one party to the other. For example, it can be shown that if the principal's bargaining power is relatively high, giving more power to the agent can allow for revelation for a larger set of prior beliefs, as the agent captures (part of) the high type's rent and thus internalizes both the costs and benefits of revelation. On the other hand, if the agent's bargaining power is high, increasing it further can make revelation unfeasible. Intuitively, in this case, the low type has incentives to mimic the high type, so both types want to mimic each other at the same time. In turn, a contract that induces revelation may not be sustainable in equilibrium, as beliefs that deter deviations to off-the-equilibrium-path contracts from one type make the deviations attractive for the other type.

*Time-Varying Outside Option.*—The principal's outside opportunities may vary over time. Suppose first that the principal draws her type at the beginning of each period  $t$ , and these draws are uncorrelated over time; i.e.,  $\theta_t \in \{\ell, h\}$ ,  $\text{corr}(\theta_t, \theta_{t+s}) = 0$  for  $s \neq 0$ . Then the expected future value of the relationship, and thus the scope of incentive provision, are independent of the principal's current type. This simplifies the problem significantly. Regardless of bargaining power, optimal incentives solve the symmetric information program for  $r_\theta = p_0 r_\ell + (1 - p_0)r_h$ , and a default never occurs in equilibrium. If the informed principal has the bargaining power, the asymmetry of information is indeed irrelevant. The asymmetry leads to inefficiencies only if the agent has the bargaining power and, given  $p_0$ , he finds it optimal to propose a wage that the high type rejects.

The case with  $\text{corr}(\theta_t, \theta_{t+s}) \in (0, 1)$  is naturally midway between the case with  $\text{corr}(\theta_t, \theta_{t+s}) = 0$  just discussed and the case with  $\text{corr}(\theta_t, \theta_{t+s}) = 1$  studied in the

<sup>22</sup> Because both types may want to pretend to be of the other type, examining extreme off-the-equilibrium-path beliefs is not sufficient to identify the set of feasible equilibria.

article. (This case can be modeled as  $\theta_t$  following a first-order Markov process with  $1 > \Pr(\theta_{t+1} = \ell | \theta_t = \ell) > \Pr(\theta_{t+1} = \ell | \theta_t = h) > 0$ .) Now the expected future value of the relationship does depend on the current type, but only stochastically. Hence, relative to constant types, the agent's learning will be slower and incomplete, and the benefits of inducing information revelation will be lower. More generally, the less serially correlated types are, the less often revelation will be induced in a Pareto-optimal equilibrium.<sup>23</sup>

*Continuum of Types.*—The model can be extended to a continuum of types,  $\theta \in [\underline{\theta}, \bar{\theta}]$ . Suppose that under symmetric information, trade is possible with all these types. Although the analysis becomes more cumbersome, the main insights are unchanged relative to the two-type case. Types cannot be fully separated through their choice of contract, and separation involves a default or a rejection. This means that types are only partially separated in equilibrium (otherwise, a contract with a zero probability of being honored or accepted would be implemented), and the inefficiency persists over time. The density of types that honor a contract or accept an offer can be used to describe how revelation occurs. For example, revelation through default may be induced all at once, or gradually over time, in which case the density of types that honor falls as the agent's belief increases.

*Continuum of Output Levels.*—The model can also be easily extended to a continuum of output levels,  $y_t \in [\underline{y}, \bar{y}]$ . Let  $y$  have distribution  $F(y|e)$  and density  $f(y|e)$ . For the case of symmetric information about the value of the relationship, Levin (2003) shows that optimal contracts are “one-step”: bonus payments are  $\bar{b} \geq 0$  for  $y \geq \hat{y}(\theta)$  and  $\underline{b} \leq 0$  for  $y < \hat{y}(\theta)$ , where  $\hat{y}(\theta)$  is the point at which the likelihood ratio  $f_e/f(y|e(\theta))$  switches from positive to negative as a function of  $y$ . That is, to give the strongest incentives, a relational contract specifies the maximum and minimum bonus payments that satisfy the enforcement constraint. For the case of asymmetric information, one can show that optimal contracts are also one-step. If no information is revealed, or if it is revealed only through rejection, this follows immediately, as bonuses are used exclusively to provide incentives. If information is revealed through default, bonuses are also used to induce the high type to renege. However, all that matters is the total expected probability with which the high type honors or reneges, which is independent of how many steps are defined in the contract.

### B. Consequences of Default, Rejection, and Unexpected Offers

Below I discuss the restrictions on strategies and how the outcomes of the model change if they are relaxed. The online Appendix provides formal statements and proofs of these results.

Throughout the paper, I assume that following a default, the parties end the relationship with positive probability, and continue on the Pareto-optimal frontier otherwise (first part of Assumption 2). This is without loss of generality: if

<sup>23</sup> Athey and Bagwell (2008) study how the persistence of private information affects equilibrium outcomes in a dynamic Bertrand game.

a Pareto-optimal equilibrium exists, there exists a Pareto-optimal equilibrium that satisfies this assumption and gives the same expected payoffs to all the parties. This is straightforward if no default occurs in equilibrium, as the worst punishment for default is then optimal. Now suppose that a default occurs. The continuation play after default must involve no trade with positive probability; otherwise, the two types would have the same incentives to renege. Assuming termination with some probability is therefore without loss. Further, consider an equilibrium where, following default, the relationship ends with probability  $1 - \gamma$  and continues on an inefficient path of play with probability  $\gamma$ . Then for some  $\gamma'$ , there exists an equilibrium where, following default, the relationship ends with probability  $1 - \gamma'$  and continues on an efficient path with probability  $\gamma'$ , and where the parties' expected payoffs are weakly higher.

The other main restriction introduced in the paper is that, if no default occurs, the relationship always remains on the Pareto-optimal frontier (second part of Assumption 2). This means that the parties cannot end the relationship or switch to an inefficient level of trade following a rejection or an unexpected offer.<sup>24</sup> This is not without loss; as described below, there exist Pareto-optimal equilibria with expected payoffs to the parties that cannot be replicated under this assumption. On the other hand, the Pareto-optimal equilibria under this assumption are also Pareto optimal without it; that is, this assumption does not select equilibria that would not be Pareto optimal otherwise.<sup>25</sup>

The motivation for assuming Pareto-optimal play provided no default has occurred is to study situations where parties' ability to credibly punish behavior depends on the nature of such behavior. In particular, I view renegeing on a promised payment as different in essence from other deviations such as rejecting an offer or making an offer that is not expected. The former is cheating, a breach of contract that, if verifiable, would be punished by law. The latter are part of the strategic bargaining; neither a rejection nor an unexpected proposal would lead to legal action. How effectively parties can punish these deviations may also be related to market conditions. For example, in a lender-borrower setting, borrowers may not be able to get a loan after having defaulted, but market competition may ensure that those who reveal themselves as "bad" types by rejecting strict loan terms are still served.

Yet this restriction on strategies does not describe all situations. To understand how the results change when this assumption is relaxed, consider first allowing the parties to impose inefficient punishments after a rejection. For the same reasons as in the article, a contract-separating equilibrium does not exist. However, the agent can now threaten to walk away if the principal rejects his offer, so revelation through rejection is always feasible, regardless of the discount factor.

<sup>24</sup>This type of restriction on the punishments that parties can impose appears in models where parties can renegotiate after a disagreement. For example, Ramey and Watson (2002) and Miller and Watson (2010) make assumptions about the bargaining stage and the consequences of disagreement at that stage that are similar to those made here. Because all transfers are made at that stage, though, inefficient termination never occurs in their settings.

<sup>25</sup>In particular, note that while relaxing this assumption could allow for equilibria that give a higher payoff to the agent, those equilibria would necessarily give a lower payoff to the low type of principal.

Suppose next that the parties can impose inefficient punishments after an unexpected offer. Then a contract-separating equilibrium exists, as the agent can threaten to walk away if the principal does not offer a contract that gives him the expected value of the relationship with the low type (conditional on payments being honored). The low type in this case is indifferent between this and any other contract, while the high type prefers termination. Indeed, this equilibrium exists for any probability that the principal makes the offers  $\lambda > 0$ . But then, one may argue, this prediction is not very reasonable: the agent captures all the surplus even if the principal has more of, or even all, the bargaining power. Realistically, one may think that the principal would “refuse to play this equilibrium”: no matter her type, the principal is better off in a pooling equilibrium. More generally, when parties can punish each other for not offering the contracts that they expected, the concept of bargaining power loses its meaning, and almost any outcome can be obtained.

Finally, it is worth noting that regardless of the restrictions on strategies, informational asymmetries about the value of the relationship always lead to inefficiencies. This follows from the fact that while trade is profitable with both types, separation of types requires that trade cease with positive probability (or for some period of time) in equilibrium, be it after a default, a rejection, or an unexpected offer. In contrast, inefficiencies may not necessarily arise in a setting where trade with (observable) high types is not profitable. Under this assumption, which is common in the literature on reputation building (see Section I), if revelation through rejection or through the principal’s contract choice is feasible, then the symmetric-information expected surplus can be achieved.

## V. Concluding Remarks

In informal contractual relationships, parties may not perfectly know how much the other party values the relationship. This paper has studied how this asymmetry of information affects optimal contract design in a general agency setting. I showed that a party may want to overstate her value of the relationship to persuade her partner that she will deliver on a promise, or to understate this value to capture a larger share of the surplus. I characterized how parties induce and react to information revelation in different environments, and how in turn their relationships evolve. In particular, I have provided a theory for why incentives and effort may increase over time, why valuable relationships may break down, and why relationship dynamics will critically depend on the parties’ relative bargaining powers. The analysis presented here is broadly applicable and may have relevant implications for employment contracts, interfirm agreements, supply-chain relationships, informal credit contracts, and other settings where contracting tends to be informal and information is typically incomplete.

## APPENDIX

### PROOF OF PROPOSITION 2:

I proceed by proving five claims.

CLAIM 1: *A pooling PPBE always exists, and any such PPBE implements  $b_h$  in all periods.*

For existence, consider this PPBE: the agent's beliefs are  $\mu(p|b_h) = p$ ,  $\mu(p|b' \neq b_h) = 0$ ,  $\phi(\mu|w_h + \bar{b}_h) = \phi(\mu|w_h) = p$ ; the principal offers  $b_h$  (with fixed wage  $w_h$ ) in every period; the agent accepts and chooses  $e_h$ ; both types always honor; if the principal reneges, the relationship ends. Clearly, beliefs are consistent, the agent's participation and effort decisions are optimal, and the principal's payment decision is optimal. To see that offering  $b_h$  is optimal for the principal, note that for  $\theta \in \{\ell, h\}$ ,  $\pi_\theta(p, b_h) = \pi_h(0)$ ;  $\pi_\theta(0, b') < \pi_h(0)$  for any  $b' < b_h$  because the agent then chooses  $e' < e_h$ ; and  $\pi_\theta(0, b'') < \pi_h(0)$  for any  $b'' > b_h$  because the agent rejects, or accepts and chooses  $e'' = 0$ . Finally, since both types follow the same strategy, no information is revealed.

To show that any pooling PPBE implements  $b_h$ , note that such PPBE cannot implement  $b' < b_h$ ; regardless of beliefs (given Assumption 2), both types can increase their expected payoffs by deviating to  $b_h$ . Further, any PPBE that implements  $b'' > b_h$  induces separation. Otherwise, either both types always renege or both always honor. But if both types renege, the agent exerts no effort, so  $b''$  cannot be optimal. And since  $b'' > b_h$ ,  $h$  optimally reneges.

**CLAIM 2:** *A contract-pooling WMPBE with revelation exists if and only if (3) holds.*

For sufficiency, consider this WMPBE: the agent's beliefs are  $\mu(p|b^*) = p$ ,  $\mu(p|b' \neq b^*) = 0$ ,  $\phi(\mu|w^* + \bar{b}^*) \equiv p_b = p/\chi^*(p)$ ,  $\phi(\mu|w^*) = 0$ ; the principal offers  $b^*(p)$  (with  $w^*(p) = r_A + c(e^*(p)) - \chi^*(p)f(e^*(p))\bar{b}^*(p)$ ) while information has not been fully revealed; the agent accepts and chooses  $e^*(p)$ ; if  $y = \bar{y}$ ,  $h$  honors with probability  $k_h^*(p)$  and  $\ell$  with probability one; if  $y = \bar{y}$  and the principal honors,  $b^*(p_b)$  is implemented; if  $y = \bar{y}$  and the principal reneges, the relationship ends with probability  $1 - \gamma^*(p)$ , and continues with contract  $b_h$  with probability  $\gamma^*(p)$ ; if  $y = y$ ,  $b^*(p)$  is again implemented. Clearly, beliefs are consistent, the agent's participation and effort decisions are optimal, and the principal's payment decision is optimal since  $b^*(p)$  satisfies  $(I_h)$ ,  $(D_h)$ , and  $(E_\ell)$ . To see that  $b^*(p)$  is optimal, note that for  $\theta \in \{\ell, h\}$ , the most profitable deviation is  $b_h$ , with expected payoff  $\pi_h(0)$ , and  $\pi_h(p, b^*) > \pi_\ell(p, b^*)$ , so (3) gives that no type wants to deviate. Finally, since (3) holds for  $k_h^*(p) < 1$ , revelation occurs through default.

For necessity, note that if (3) does not hold,  $b^*(p)$  cannot be implemented, as  $\ell$  would want to deviate to  $b_h$  regardless of beliefs (given Assumption 2). And by the definition of  $b^*(p)$ , if (3) does not hold for any  $k_h(p)$  with  $k_h(p) \in [0, 1)$ , then there is no contract-pooling WMPBE that induces revelation. (See the text for a discussion of the program that defines  $b^*(p)$ .)

**CLAIM 3:** *If a contract-pooling WMPBE with revelation does not exist, a contract-pooling PPBE with revelation does not exist.*

I prove the contrapositive:

**Step 1:** Suppose there exists a PPBE with revelation with  $\tilde{\mu}_{\tilde{i}}(p) \neq \tilde{\mu}_{\tilde{i}+1}(p)$  for some  $\tilde{i}$ . (Of course, the beliefs  $\tilde{\mu}_{\tilde{i}}$ ,  $\tilde{\mu}_{\tilde{i}+1}$  would differ only for off-the-equilibrium-path contracts. More precisely, consider a PPBE with  $\tilde{\mu}_{\tilde{i}}(p|\tilde{b}) \neq \tilde{\mu}_{\tilde{i}+1}(p|\tilde{b})$  where either  $g_\ell(\tilde{\mu}_{\tilde{i}}, \tilde{b}) = g_h(\tilde{\mu}_{\tilde{i}}, \tilde{b}) = 0$ , or  $g_\ell(\tilde{\mu}_{\tilde{i}+1}, \tilde{b}) = g_h(\tilde{\mu}_{\tilde{i}+1}, \tilde{b}) = 0$ , or both.) But the

subgame starting at  $\tilde{t}$  is otherwise the same as that starting at  $\tilde{t} + 1$ . Then there exists a PPBE with revelation with  $\tilde{\mu}_t(p) = \tilde{\mu}_{\tilde{t}+1}(p)$  for  $t = \tilde{t}, \tilde{t} + 1$ .

**Step 2:** Suppose there exists a PPBE with revelation with  $\tilde{k}_{h\tilde{t}} = 1$  for some  $\tilde{t}$ ,  $\tilde{k}_{ht} < 1$  for  $t \neq \tilde{t}$ . But by the same reasoning as above, then there exists a PPBE with revelation with  $\tilde{k}_{ht} = \tilde{k}_{ht}$  for  $t < \tilde{t}$ ,  $\tilde{k}_{ht} = \tilde{k}_{ht+1}$  for  $t \geq \tilde{t}$ .

**Step 3:** Suppose there exists a PPBE with revelation with  $\tilde{k}_{h\tilde{t}}(\mu(p), b) \neq \tilde{k}_{h\tilde{t}+1}(\mu(p), b)$  for some  $\tilde{t}$ . But by the same reasoning as above, then there exists a PPBE with revelation with  $\tilde{k}_{ht}(\mu(p), b) = \tilde{k}_{h\tilde{t}+1}(\mu(p), b)$  for  $t = \tilde{t}, \tilde{t} + 1$  and, thus, given steps 1 and 2, a WMPBE with revelation exists.

**CLAIM 4:** *A contract-semiseparating PPBE that implements contracts  $b'$  and  $b'' < b'$  exists only if there exists a contract-pooling PPBE that implements  $b'$ .*

**Step 1:** Suppose there exists a contract-semiseparating PPBE where contracts  $b', b''$  are implemented and no type reneges in equilibrium. Then it must be that  $b' = b'' = b_h$ ; otherwise, regardless of beliefs, some type would want to deviate. But then no separation occurs. Contradiction.

**Step 2:** Suppose there exists a contract-semiseparating PPBE where  $\ell$  offers some contract  $b'$ ,  $h$  mixes between  $b'$  and some  $b'' > b'$ ,  $\ell$  honors  $b'$  with probability one, and  $h$  reneges on  $b''$  with positive probability. Then in equilibrium  $\pi_\ell(\mu(p|b'), b') \geq \pi_h(0)$  (otherwise,  $\ell$  would have a profitable deviation) and  $\mu(p|b'') = 0$ , and, hence,  $\pi_h(\mu(p|b''), b'') < \pi_h(0) \leq \pi_h(\mu(p|b'), b')$ . But then  $h$  is not willing to mix between  $b'$  and  $b''$ . Contradiction.

**Step 3:** By Steps 1 and 2, in any contract-semiseparating PPBE,  $h$  offers some contract  $b'$ ,  $\ell$  mixes between  $b'$  and some  $b'' < b'$ ,  $h$  reneges on  $b'$  with positive probability, and  $\ell$  honors  $b'$  and  $b''$  with probability one. Now note that if  $\ell$  offers  $b'$  with probability  $g'_\ell \in (0, 1)$ ,  $\mu(p|b') = pg'_\ell / [pg'_\ell + (1 - p)] < p$ , and thus given  $k_h$ ,  $\chi(\mu(p|b'), b') < \chi(p, b')$  and  $e(\mu(p|b'), b') < e(p, b')$ . But then  $\pi_\theta(\mu(p|b'), b') < \pi_\theta(p, b')$  for  $\theta = \ell, h$ . Hence, if there exists a PPBE where both types offer  $b'$  given belief  $\mu(p|b')$ , there exists a PPBE where both types offer  $b'$  given belief  $p$ .

**CLAIM 5:** *If (3) holds, the Pareto-optimal PPBE induces separation of types through default. If (3) does not hold, the Pareto-optimal PPBE is the pooling PPBE.*

By Proposition 1 and Claim 1, and the fact that  $\lambda = 1$ , we need to consider only contract-pooling PPBE with revelation through default, contract-semiseparating PPBE, and the pooling PPBE implementing  $b_h$ . Suppose first that (3) holds. Then by Claim 2, a contract-pooling PPBE with revelation exists, which also implies that both types prefer this PPBE to the pooling PPBE. Furthermore, suppose that a contract-semiseparating PPBE also exists. Then using the proof of Claim 4, it is straightforward to see that there exists a contract-pooling PPBE with revelation that Pareto-dominates the contract-semiseparating PPBE. Suppose next that (3) does not hold. Then by Claims 2, 3, and 4, the pooling PPBE is unique, and thus Pareto optimal.

Finally, it follows directly from inspection of condition (3) that there exists  $\hat{p} \in (0, 1)$  such that this condition holds if and only if  $p_0 \geq \hat{p}$ .

**PROOF OF PROPOSITION 3:**

Restrict  $k_h(p) \in [0, \bar{k}_h]$  for all  $p$  and some  $\bar{k}_h \in (0, 1)$ , with  $\bar{k}_h$  arbitrarily close to one. Consider  $p_0 \geq \hat{p}$ . I proceed by proving five claims.

*CLAIM 1: There exists  $p' \in (\hat{p}, 1)$  such that a WMPBE with immediate revelation exists if and only if  $p_0 \geq p'$ .*

It follows directly from condition (3) that there exists  $p' \in [\hat{p}, 1)$  such that  $\pi_\ell(p_0, b) \geq \pi_h(0)$  for some  $b$  satisfying  $(D_h)$  and  $(E_\ell)$  and  $k_h(p_0) = 0$  if and only if  $p_0 \geq p'$ . To show that  $p' > \hat{p}$ , take  $\tilde{p} = p' - \varepsilon$  for some small  $\varepsilon > 0$ . By construction,  $\pi_\ell(\tilde{p}, b) < \pi_h(0)$  if  $k_h(\tilde{p}) = 0$ , but  $\pi_\ell(p', b') \geq \pi_h(0)$  for some  $b'$  and  $k_h(p') = 0$ . Then let  $\tilde{b}$  satisfy  $(I_h)$  for  $\gamma(\tilde{p}) = 0$  and  $p_{\tilde{b}} \equiv p'$ :  $\tilde{b} = [\delta/(1 - \delta)] \times [\pi_h(p', b') - r_h]$ . Note that  $\pi_h(p', b') > \pi_h(0)$ , and, hence,  $\tilde{b} > \bar{b}_h$ . Let  $\tilde{k}_h \in (0, 1)$  be such that for  $k_h(\tilde{p}) = \tilde{k}_h, p_{\tilde{b}} = p'$ . As  $\varepsilon \rightarrow 0, \tilde{k}_h \rightarrow 1$ . But then for  $\varepsilon$  sufficiently small and  $k_h(\tilde{p}) = \tilde{k}_h, \pi_\ell(\tilde{p}, \tilde{b}) \geq \pi_h(0)$ .

*CLAIM 2: There exists  $\bar{p} \in [p', 1)$ , for  $p'$  as defined in Claim 1, such that immediate revelation is Pareto optimal if  $p_0 \geq \bar{p}$ .*

This follows immediately from the low type's maximization program in Section IIIB and the two types' expected payoffs.

*CLAIM 3: Revelation is slower in  $\ell$ 's preferred WMPBE than in  $h$ 's preferred WMPBE.*

Observe that  $h$  receives the same (flow) payoff as  $\ell$  when she honors the payments, while  $h$  receives strictly more when she reneges. Hence,  $h$ 's preferred WMPBE cannot induce slower revelation than  $\ell$ 's. To show that there exist parameters for which  $\ell$  prefers slower revelation, take  $p_0 = p'$ , where  $p'$  is as defined in Claim 1, and assume that  $\gamma^*(p) = 0$  for all  $p$  (see Claim 4 below). Then  $\pi_\ell(p_0, b_\ell | k_h(p_0) = 0) = \pi_h(0), \pi_h(p_0, b_\ell | k_h(p_0) = 0) > \pi_h(0)$ . Moreover,  $h$ 's preferred WMPBE implements  $b_\ell$  and induces immediate revelation. Now consider another WMPBE implementing  $\tilde{b}$  and inducing gradual revelation. Let  $\tilde{k}_h(p_{0\tilde{b}}) = 0, \tilde{b}(p_{0\tilde{b}}) = \bar{b}_\ell, \tilde{b}(p_0) = [\delta/(1 - \delta)](\pi_h(p_{0\tilde{b}}, b_\ell | k_h(p_{0\tilde{b}}) = 0) - r_h)$ , where  $p_{\tilde{b}} = p_0/\tilde{\chi}(p_0), \tilde{\chi}(p_0) = p_0 + (1 - p_0)\tilde{k}_h(p_0)$ . Note that  $\tilde{b}(p_0) > \bar{b}_h$ . Thus, for  $\tilde{k}_h(p_0)$  sufficiently close to one,  $\ell$  prefers this WMPBE with gradual revelation to the WMPBE with immediate revelation.

*CLAIM 4: There exists  $\Delta \in \mathbb{R}$  such that if  $(\bar{y} - \underline{y}) - \bar{b}_\ell \geq \Delta, \gamma^*(p) = 0$  for all  $p = p_0, p_{0b}, \dots$ , in any Pareto-optimal WMPBE.*

In any Pareto-optimal WMPBE, when  $k_h(p) = 0, (E_\ell)$  binds and  $(D_h)$  is slack. Otherwise, one could increase  $\gamma(p)$  and increase  $\pi_h(p, b)$  without affecting  $\pi_\ell(p, b)$ . It is then sufficient to show that if  $(\bar{y} - \underline{y}) - \bar{b}_\ell$  is sufficiently large, the derivative

of  $\pi_\ell(p, b)$  with respect to  $\bar{b}(p)$  given  $k_h(p) = 0$  and  $\gamma(p) = 0$  is positive for all  $p \geq p'$ , for  $p'$  as defined in Claim 1. For  $k_h(p) = 0$ ,

$$\pi_\ell(p, b) = \frac{(1 - \delta)[\underline{y} + f(e)(\bar{y} - \underline{y}) - c(e) - r_A - (1 - p)f(e)\bar{b}] + \delta f(e)\pi_\ell(1)}{1 - \delta(1 - f(e))}.$$

The derivative of  $\pi_\ell(p, b)$  with respect to  $e$  is

$$\begin{aligned} \frac{\partial \pi_\ell(p, b)}{\partial e} &= \frac{1}{[1 - \delta(1 - f(e))]^2} \left\{ (1 - \delta)[f'(e)(\bar{y} - \underline{y}) - c'(e) - (1 - p)f'(e)\bar{b}] \right. \\ &\quad \left. + \delta f'(e)\pi_\ell(1) \right\} [1 - \delta(1 - f(e))] \\ &\quad - \delta f'(e)\pi_\ell(p, b)[1 - \delta(1 - f(e))]. \end{aligned}$$

By constraint (IC<sub>A</sub>),  $f'(e)p\bar{b} - c'(e) = 0$ . Substituting and rearranging terms,

$$\frac{\partial \pi_\ell(p, b)}{\partial e} = \frac{f'(e)[(1 - \delta)(\bar{y} - \underline{y} - \bar{b}) + \delta(\pi_\ell(1) - \pi_\ell(p, b))]}{1 - \delta(1 - f(e))}.$$

The derivative of  $\pi_\ell(p, b)$  with respect to  $\bar{b}$  is then

$$\begin{aligned} \text{(A1)} \quad &\frac{f'(e)[(1 - \delta)(\bar{y} - \underline{y} - \bar{b}) + \delta(\pi_\ell(1) - \pi_\ell(p, b))]}{1 - \delta(1 - f(e))} e'(\bar{b}) \\ &- \frac{(1 - \delta)(1 - p)f(e)}{1 - \delta(1 - f(e))}. \end{aligned}$$

Substituting with  $e'(\bar{b}) = pe'(p\bar{b})$  and rearranging, (A1) is greater or equal to zero if and only if

$$\text{(A2)} \quad p \left\{ f(e) + f'(e)e'(p\bar{b}) \left[ \bar{y} - \underline{y} - \bar{b} + \frac{\delta}{1 - \delta} (\pi_\ell(1) - \pi_\ell(p, b)) \right] \right\} \geq f(e).$$

If  $\bar{y} - \underline{y} - \bar{b}_\ell$  is sufficiently large, then  $\bar{y} - \underline{y} - [\delta/(1 - \delta)](\pi_\ell(p, b_\ell) - r_\ell)$  is sufficiently large, and (A2) holds for  $\bar{b} = \bar{b}_\ell$  and all  $p$  such that  $\pi_\ell(p, b_\ell) \geq \pi_h(0)$ .

**CLAIM 5:** *There exists  $\Delta \in \mathbb{R}$  such that if  $(\bar{y} - \underline{y}) - \bar{b}_\ell \geq \Delta$ , in any Pareto-optimal WMPBE,  $\bar{b}(p)$  and  $e(p)$  are increasing along the revelation path, and  $w(p)$  and  $\chi(p)$  are decreasing.*

By  $(I_h)$ , either  $\chi(p)\bar{b}(p)$  or  $\gamma(p)$  increase along the revelation path. By Claim 4,  $\gamma^*(p) = 0$  for all  $p = p_0, p_{0b}, \dots$ , so  $\chi(p)\bar{b}(p)$  must be increasing.  $(IC_A)$  then gives that  $e(p)$  is increasing and the agent's participation constraint gives that  $w(p)$  is decreasing. Because  $\pi_h(p, b)$  is increasing and  $\gamma(p)$  is constant over time,  $(I_h)$  and  $(D_h)$  imply that  $\bar{b}(p)$  is increasing. Finally, note that  $\pi_\ell(p, b)$  and  $\pi_h(p, b)$  are increasing in  $\chi(p)$  and, by the proof of Claim 4, they are also increasing in  $\bar{b}(p)$  for all  $p$  for which revelation is feasible. So if  $\chi(p)$  and  $\bar{b}(p)$  are increasing along the revelation path,  $\pi_\ell(p, b)$  is also increasing. But then (since future payoffs are discounted), if there exists a Pareto-optimal WMPBE where  $\chi(p)$  is increasing, there exists another WMPBE where  $\chi(p)$  is decreasing and  $\pi_\ell(p_0, b)$  and  $\pi_h(p_0, b)$  are strictly larger. Thus,  $\chi(p)$  optimally decreases over time.

**PROOF OF PROPOSITION 4:**

Let an offer be denoted by the proposed bonus and expected transfer to the principal  $(b, r_p)$ , where  $r_p \equiv \mathbb{E}_y[y - w - b(y) | e]$ . I proceed by proving three claims.

*CLAIM 1: If (4) and (5) hold, the Pareto-optimal PPBE implements  $(b_\ell, r_p^{**})$  and induces revelation through rejection in the first period.*

Suppose that (4) and (5) hold. If the agent offers  $(b_\ell, r_p^{**})$ , (4) implies that  $\ell$  accepts and  $h$  rejects. Condition (5) implies  $u(p_0, b_\ell | r_p^{**}) \geq u(0)$ , and, because by definition  $b_\ell$  maximizes enforceable incentives and  $r_p^{**}$  minimizes the principal's transfer,  $u(p_0, b_\ell | r_p^{**}) \geq u(p_0, b' | r')$  for any contract  $(b', r')$  that induces immediate revelation. Finally, condition (5) implies  $u(p_0, b_\ell | r_p^{**}) \geq u(p_0, b' | r')$  for any contract  $(b', r')$  that induces gradual revelation. To see this, note that if delaying revelation one period could increase the agent's expected payoff, then it would be by having  $h$  accept with a relatively high probability  $a'_h$  in the current period. But then, if  $u(p_0, b_\ell | r_p^{**}) = u(0)$ ,  $\underline{b}' = -[\delta/(1 - \delta)](u(p_0, b_\ell | r_p^{**}) - r_A) = \underline{b}_h$ , and thus  $u(p_0, b') < u(0)$  for  $a'_h < 1$ ,  $u(p_0, b') \rightarrow u(0)$  as  $a'_h \rightarrow 1$ . And if  $u(p_0, b_\ell | r_p^{**}) > u(0)$ ,  $\underline{b}' < \underline{b}_h$ , but these higher incentives must result in a less than proportional increase in the expected surplus, as otherwise  $b_h$  would not be optimal when  $\theta = h$  is observable. Thus,  $u(p_0, b') < u(p_0, b_\ell | r_p^{**})$ .

*CLAIM 2: If (4) does not hold, the Pareto-optimal PPBE implements  $(b_h, r_h)$  in every period.*

Suppose (4) does not hold. Then immediate revelation is not feasible, since  $\ell$  rejects whenever  $h$  rejects. Gradual revelation is not feasible either. First note that a PPBE where  $h$  mixes between accepting and rejecting and  $\ell$  accepts with probability one cannot exist, as again  $\ell$  wants to reject. Consider then a PPBE where  $\ell$  mixes between accepting and rejecting and  $h$  rejects with probability one. Given a prior  $p$  at time  $t$ , the agent offers  $(b(p), r_p(p))$ . If the principal accepts, the agent learns that  $\theta = \ell$  and offers  $(b_\ell, r_\ell)$  from then on. If the principal rejects, the agent updates his belief. Then there exists a posterior  $p'$  low enough that the agent optimally offers  $(b_h, r_h)$  from then on. But then, when  $p = p''$  such that  $\mu(p'' | \text{reject}) = p'$ ,  $\ell$  rejects any offer that  $h$  rejects. In turn, when  $p = p''$ , the agent optimally offers  $(b_h, r_h)$ . But then, when  $p = p'''$  such that  $\mu(p''' | \text{reject}) = p''$ ,  $\ell$  rejects any offer that  $h$  rejects.

Continuing with this reasoning gives that revelation is not feasible if immediate revelation is not feasible.

CLAIM 3: *If (5) does not hold, the Pareto-optimal PPBE implements  $(b_h, r_h)$  in every period.*

Suppose that (5) does not hold. Then the agent is better off by offering  $(b_h, r_h)$  than  $(b_\ell, r^{**})$  or any other contract that induces immediate revelation. And by the discussion in Claim 1, he is also better off by offering  $(b_h, r_h)$  than any contract  $(b', r')$  that induces gradual revelation.

Finally, it follows directly from inspection of condition (5) that there exists  $\hat{p} \in (0, 1)$  such that this condition holds if and only if  $p_0 \geq \hat{p}$ .

## REFERENCES

- Abreu, Dilip. 1988. "On the Theory of Infinitely Repeated Games with Discounting." *Econometrica* 56(2): 383–96.
- Aghion, Philippe, and Patrick Bolton. 1987. "Contracts as a Barrier to Entry." *American Economic Review* 77(3): 388–401.
- Aghion, Philippe, and Benjamin Hermalin. 1990. "Legal Restrictions on Private Contracts Can Enhance Efficiency." *Journal of Law, Economics, and Organization* 6(2): 381–409.
- Athey, Susan, and Kyle Bagwell. 2008. "Collusion with Persistent Cost Shocks." *Econometrica* 76(3): 493–540.
- Baker, George, Robert Gibbons, and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *Quarterly Journal of Economics* 117(1): 39–84.
- Bowen, Robert M., Larry DuCharme, and D. Shores. 1995. "Stakeholders' Implicit Claims and Accounting Method Choice." *Journal of Accounting and Economics* 20(3): 255–95.
- Bull, Clive. 1987. "The Existence of Self-Enforcing Implicit Contracts." *Quarterly Journal of Economics* 102(1): 147–59.
- Coase, Ronald H. 1972. "Durability and Monopoly." *Journal of Law and Economics* 15(1): 143–49.
- Conlin, Michelle. 2001. Where Layoffs Are a Last Resort. *Business Week*, October 8.
- DeAngelo, Harry, and Linda DeAngelo. 1991. "Union Negotiations and Corporate Policy: A Study of Labor Concessions in the Domestic Steel Industry during the 1980s." *Journal of Financial Economics* 30(1): 3–43.
- Endlich, Lisa. 1999. *Goldman Sachs: The Culture of Success*. New York: Alfred A. Knopf.
- Freixas, Xavier, Roger Guesnerie, and Jean Tirole. 1985. "Planning under Incomplete Information and the Ratchet Effect." *Review of Economic Studies* 52(2): 173–91.
- Fuchs, William. 2007. "Contracting with Repeated Moral Hazard and Private Evaluations." *American Economic Review* 97(4): 1432–48.
- Fudenberg, Drew, David I. Levine, and Eric Maskin. 1994. "The Folk Theorem with Imperfect Public Information." *Econometrica* 62(5): 997–1039.
- Fudenberg, Drew, and Jean Tirole. 1991. "Perfect Bayesian Equilibrium and Sequential Equilibrium." *Journal of Economic Theory* 53(2): 236–60.
- Genicot, Garance, and Debraj Ray. 2006. "Bargaining Power and Enforcement in Credit Markets." *Journal of Development Economics* 79(2): 398–412.
- Ghosh, Parikshit, and Debraj Ray. 1996. "Cooperation in Community Interaction without Information Flows." *Review of Economic Studies* 63(3): 491–519.
- Gibbons, Robert. 2005. "Incentives Between Firms (and Within)." *Management Science* 51(1): 2–17.
- Grossman, Sanford J., and Motty Perry. 1986. "Perfect Sequential Equilibrium." *Journal of Economic Theory* 39(1): 97–119.
- Hermalin, Benjamin E. 2002. "Adverse Selection, Short-Term Contracting, and the Underprovision of On-the-Job Training." *Contributions to Economic Analysis and Policy* 1(1): Article 5.
- Holmstrom, Bengt. 1999. "The Firm as a Subeconomy." *Journal of Law, Economics, and Organization* 15(1): 74–102.
- Hymowitz, Carol. 1990. "Kodak Gives Parachutes to Workers." *Wall Street Journal*, January 12, B1.

- Kranton, Rachel E.** 1996. "The Formation of Cooperative Relationships." *Journal of Law, Economics, and Organization* 12(1): 214–33.
- Laffont, Jean-Jacques, and Jean Tirole.** 1988. "The Dynamics of Incentive Contracts." *Econometrica* 56(5): 1153–75.
- Levin, Jonathan.** 2003. "Relational Incentive Contracts." *American Economic Review* 93(3): 835–57.
- Lewis, Tracy R., and David E. M. Sappington.** 1989. "Countervailing Incentives in Agency Problems." *Journal of Economic Theory* 49(2): 294–313.
- MacLeod, W. Bentley.** 2003. "Optimal Contracting with Subjective Evaluation." *American Economic Review* 93(1): 216–40.
- MacLeod, W. Bentley, and James M. Malcomson.** 1989. "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment." *Econometrica* 57(2): 447–80.
- Miller, David, and Joel Watson.** 2010. "A Theory of Disagreement in Repeated Games with Renegotiation." <http://dss.ucsd.edu/~d9miller/papers/MillerWatson-Disagreement.pdf> (accessed April 29, 2011).
- Ramey, Garey, and Joel Watson.** 2002. "Contractual Intermediaries." *Journal of Law, Economics, and Organization* 18(2): 362–84.
- Rubinstein, Ariel.** 1985. "A Bargaining Model with Incomplete Information about Time Preferences." *Econometrica* 53(5): 1151–72.
- Schmitz, Patrick W.** 2006. "Information Gathering, Transaction Costs, and the Property Rights Approach." *American Economic Review* 96(1): 422–34.
- Sobel, Joel.** 1985. "A Theory of Credibility." *Review of Economic Studies* 52(4): 557–73.
- Spier, Kathryn E.** 1992. "Incomplete Contracts and Signalling." *RAND Journal of Economics* 23(3): 432–43.
- Stewart, James.** 1993. "Taking the Dare." *The New Yorker*, July 26: 34–9.
- Watson, Joel.** 1999. "Starting Small and Renegotiation." *Journal of Economic Theory* 85(1): 52–90.
- Watson, Joel.** 2002. "Starting Small and Commitment." *Games and Economic Behavior* 38(1): 176–99.
- Williamson, Oliver E.** 2000. "The New Institutional Economics: Taking Stock, Looking Ahead." *Journal of Economic Literature* 38(3): 595–613.