

Relational Contracts and the Value of Relationships*

Marina Halac[†]

University of California, Berkeley

JOB MARKET PAPER

November 1, 2008

Abstract

Relational contracts can be used to provide incentives if the future value of a relationship between contracting parties is sufficiently large. But what happens if the relationship's value is not commonly known? This paper studies optimal relational contract design in a principal-agent setting where the principal's outside option is her private information. I show that incentive provision is always less efficient than under symmetric information. The source of the inefficiency depends on the allocation of bargaining power. If the principal has strong bargaining power, the high-outside-option-type wants to mimic the low-outside-option-type to provide strong incentives and then renege and walk away. If the agent has strong bargaining power, the low type wants to mimic the high type to receive a high transfer when the agent proposes compensation. Both types may want to mimic the other type simultaneously under some bargaining power distributions. I characterize when separation of types is optimal, how it occurs in equilibrium, and how this depends on the parties' bargaining positions. Information may be revealed through default or rejection, which may occur immediately or gradually, and may be delayed.

*I am indebted to Ben Hermalin, Shachar Kariv, and Steve Tadelis for their continuous guidance, advice, and support. I thank Botond Kőszegi, Jon Levin, Kristóf Madarász, and participants at the Microeconomic Theory seminar and the Oliver E. Williamson seminar at UC Berkeley for helpful comments.

[†]Department of Economics, University of California, Berkeley. E-mail: mhalac@econ.berkeley.edu.

1 Introduction

Repeated interaction can improve the efficiency of agency relationships. 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. Informal agreements are self-enforcing if the consequences of non-compliance ensure that neither party wants to renege. For discretionary payments to be self-enforcing, the relationship must then be sufficiently valuable, so that the parties prefer to honor the payments and continue with the relationship rather than to renege and end the relationship. This self-enforcement requirement in turn implies that the scope of incentive provision in a relational contract is 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. For example, consider a repeated interaction between a firm and an association of workers or a labor union. Both the firm and the workers 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. This would require information not only about the firm's payoff within the relationship, but also about the firm's opportunities outside the relationship. How valuable are the firm's outside options? How valuable will they be tomorrow? Does the firm find it optimal to maintain the relationship with the workers? Or does it have plans to shut down the factory and start a new business in a remote location, and hence does not care about its reputation?

This paper investigates a whole array of new issues that arise when the value of the relationship is not commonly known. Continuing with the example, I show that the firm may want to make the workers believe that the value of the relationship is higher than it truly is, so that it can provide high-powered incentives and then renege on the contract and walk away with a high benefit. On the other hand, if the workers are likely to propose compensation, 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. As for the workers, they may not want to exert high effort if they anticipate that the firm may not honor the promised payments. Additionally, when proposing compensation, the workers may want to offer the firm a relatively low wage to ensure that it will accept, or a relatively high one to force a firm with a low relationship value to reveal itself by rejecting the offer.

These issues are present in real-world contracting. Leap (1995, p. 307) notes that "the union generally does not have access to the employer's production, financial, and personnel information."¹ Bowen, DuCharme, and Shores (1995) provide empirical evidence showing that the extent to which a firm has implicit claims with workers and other stakeholders is positively related to the firm's choice of income-increasing accounting methods. Since reported financial data are used to assess the firm's

¹Because of this, Leap explains, the National Labor Relations Board (NLRB) holds that management must provide information requested by the union if such information is relevant to contract negotiations. However, the author notes that "the NLRB has taken a liberal view as to what constitutes relevant information."

reputation for fulfilling these claims, firms have an incentive to inflate these numbers. Uncertainty about the firm’s prospects can clearly render relational contracts ineffective. For example, when workers at Eastman Kodak Co. became concerned about the continuous restructuring and the (remote) possibility of a takeover, the firm had to introduce a formal plan guaranteeing severance pay and other benefits to “remove the ‘uncertainty factor’ which demoralizes and demotivates employees...”^{2,3} And there is also evidence that firms sometimes end up renegeing on their relational contracts. Some well-known stories include those of First Boston and Goldman Sachs, both of which unexpectedly cut discretionary end-of-year bonuses, and IBM, which abandoned its long-standing implicit agreement of “no-layoffs” in the 1990s.⁴

The empirical evidence is also consistent with the idea that firms’ incentives depend on the strength of organized labor. Bowen et al. (1995) find that one of their eight measures of implicit contracting—the use of defined benefit pension plans—tends to favor income-decreasing rather than income-increasing accounting methods.⁵ The reason, they claim, is that this measure is also a proxy for a firm’s dependence on contracts negotiated with unions. For the steel industry during the 1980s, DeAngelo and DeAngelo (1991) document that firms’ reported net income is significantly lower during union negotiations, controlling for cash flows, than in nonnegotiation years. Losses due to restructuring decisions are deliberately reported during union negotiations to help the firms’ case for wage and other concessions from the union.

To formally address these issues, this paper develops an infinite-horizon principal-agent model with nonverifiable outcomes and hidden action in which the principal’s opportunities outside the relationship are known only to her. I consider a setup with two principal types, a “low” type with a low outside option (high relationship value) and a “high” type with a high outside option (low relationship value). In any period, the principal or the agent makes a take-it-or-leave-it offer to the other party. An offer consists of a compensation package for the agent, which combines an enforceable fixed wage with discretionary output-contingent bonus payments, or other forms of performance-based rewards such as raises, promotions, and prizes. If the offer is accepted, the agent chooses effort and this generates an 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 bonus payments. If the offer is rejected, the parties receive their outside options. I assume that the

²The quote is from the firm’s announcement of the plan, as it appeared in its employee newspaper, the *Kodakery*. See Hymowitz (1990).

³A related example concerns the use of seniority wages or so-called “Lazear contracts” (Lazear, 1979). Since these agreements involve informal, delayed wage raises, workers will be reluctant to accept them if they fear that the firm will renege and dismiss the workers as they become senior. According to Hutchens (1989), this can explain why some firms—particularly small firms whose reputation value is uncertain—find it difficult to implement these contracts.

⁴Stewart (1993) describes the case of First Boston. Many bankers quit after a second consecutive year in which bonuses were lower than they expected given their implicit agreement with the bank. Credit Suisse, First Boston’s parent company, argued that lower bonuses were just the result of lower performance, and that bonuses were not supposed to match the market, as employees claimed, but the bank’s financial outcomes. Gibbons (2005, p. 8) comments that “the Swiss may have claimed such a misunderstanding as a way to cover their decision to lower bonuses after the junk-bond market collapsed, which caused bankers specializing in mergers and acquisitions to become less valuable, reducing the present value of future profits to the Swiss from their relationship with the bankers.” The cases of Goldman Sachs and IBM are described in Endlich (1999) and Gibbons respectively.

⁵For a discussion of the implicit nature of pension contracts, see Ippolito and James (1992).

principal's type is fixed over time and consider parameters such that, under symmetric information about the value of the relationship, trade is possible with both types.

As is well known, the set of equilibria in repeated games with incomplete information is very large. Limiting attention to equilibria that are subgame perfect and (constrained) Pareto optimal is generally not sufficient to obtain precise predictions. Following a large strand of the literature, I thus concentrate on a subset of the Pareto-optimal equilibria by introducing restrictions on what parties can do in every contingency. I assume that if a default occurs, the relationship ends with some positive probability (without loss), but, otherwise, given the agent's beliefs and given her type in the case of the principal, the relationship always remains on the Pareto-optimal frontier. Put differently, I distinguish between "cheating" and "tough bargaining." I postulate that a failure to honor a payment may lead to conflict and breakup, while a rejection or an "unexpected" offer does not constitute a credible excuse to walk away or impose similar punishments. Without doubt, these assumptions affect the parties' bargaining powers and may not describe all bargaining situations. I discuss when these assumptions are plausible and how the outcomes change if they are relaxed later in the paper. Below I summarize my main results.

I find that full separation of types can never occur through the contract itself; that is, there exists no contract-separating equilibrium. When the principal has strong bargaining power (namely, she is very likely to make the offers), the high type wants to mimic the low type to provide strong incentives and then renege and walk away with a high payoff. When the agent has strong bargaining power, the low type wants to mimic the high type to signal a high outside payoff and then receive a high payment when the agent makes the offer. I show that one of these deviations is always profitable for some principal type. This result holds for any two contracts and any bargaining protocol. A direct implication is that informational asymmetries about the relationship value always lead to distortions in incentive provision.

Separation of principal types may still occur in equilibrium, however. In addition to partial separation through contracts, separation may occur through default (one type reneges while the other honors the contract) or through rejection (one type rejects while the other accepts the agent's offer). The former is feasible only if the principal has some bargaining power, the latter only if the agent does.

I start by studying Pareto-optimal contracts when the principal has all the bargaining power. In this setting, only payments from the principal to the agent are self-enforcing, and it is the high type who has incentives to misrepresent her type. I show that an equilibrium with revelation exists if and only if the probability that the principal's type is low is sufficiently high. When revelation is not feasible, the unique Pareto-optimal equilibrium implements the symmetric-information contract of the high type—the contract that is optimal when the principal's type is known to be high. When revelation is feasible, any Pareto-optimal equilibrium induces revelation through default. Partial revelation through contracts is never Pareto optimal.

When the probability of a low principal type is not at the extremes, there exist multiple Pareto-optimal equilibria with revelation through default, prescribing different paths of information revelation and leading to full separation at different speeds. I characterize revelation in equilibrium

as being immediate or gradual, and also possibly delayed. Since the high type imposes a negative externality on the low type, it may seem that the low type will want to separate as fast as possible. I find, in contrast, that the opposite is true. The low type’s preferred equilibrium delays information revelation relative to the high type’s preferred equilibrium. Furthermore, during the delay phase, this equilibrium often implements a “probationary contract” that induces minimal separation. More generally, for any given speed of revelation, I show that the probability of separation increases along the path of revelation, and either the bonus payments also increase or the punishment for default falls.

Next I study the case where the agent has all the bargaining power. In this setting, only payments from the agent to the principal are self-enforcing, and it is the low type who has incentives to misrepresent her type. An equilibrium with revelation exists if and only if the probability that the principal’s type is low is sufficiently high and the discount factor is sufficiently low. This latter requirement implies that revelation is unlikely, as relational incentives are enforceable only if the discount factor is relatively high. Similar to the case above, when revelation is not feasible, the unique Pareto-optimal equilibrium implements the symmetric-information contract of the high type, and when it is feasible, any Pareto-optimal equilibrium induces revelation through rejection. But, as one goes from signaling to screening, the results differ from those above in that the Pareto-optimal equilibrium is now always unique. A Coase-conjecture-like argument shows that either immediate revelation can be induced in the first period, or no revelation can be induced in equilibrium. And the fact that the agent can now internalize the externality implies that either immediate revelation is optimal in the first period, or no revelation is optimal in equilibrium.

Finally, I consider an intermediate case where both parties have positive bargaining power. This case is interesting in itself, as it provides insights that cannot be obtained from any of the two extreme cases described above. The main reason is that, for some parameter values and intermediate bargaining power distributions, the two principal types have incentives to misrepresent their types simultaneously. The high type wants to pretend to be the low type while the low type wants to pretend to be the high type. As a result, and in contrast to the cases above, Pareto-optimal equilibria with and without revelation can coexist in this setting.

I characterize the equilibria for the intermediate case by studying how revelation changes as one party receives more bargaining power. There are two effects at play here. One is internalization of the externality; that is, the fact that when the agent proposes compensation, he can capture (at least part of) the rent that the high type obtains when renegeing, and hence internalize both the benefits and costs of inducing revelation. The other is conflicting incentives; that is, the fact that both principal types may want to pretend to be of the other type at the same time. I show that as bargaining power is shifted from the principal to the agent, revelation increases due to internalization of the externality, but falls due to conflicting incentives. The specification of off-the-equilibrium-path beliefs influences, but may not fully determine, the relative importance of these two effects.

Whereas the bargaining power structure is irrelevant in terms of joint surplus when the value of the relationship is observable, this paper shows that it plays an important and not obvious role when

the value is not observable. Essentially, under asymmetric information, the allocation of bargaining power determines the source of the inefficiency. This, in turn, results in new predictions in terms of the parties' individual payoffs. In particular, when the relationship value is not observable, it is no longer the case that a party's payoff is everywhere increasing in his or her bargaining power. The agent may prefer to have only some bargaining power over having all the bargaining power; the principal may ex ante (before learning her type) prefer to have no bargaining power over having only some bargaining power.

The next section discusses the related literature. Section 3 presents the model and the case of symmetric information about the value of the relationship. Section 4 introduces asymmetric information and provides general results on information revelation. Section 5 characterizes the Pareto-optimal equilibria under different bargaining power structures. Section 6 considers some extensions of the model. Section 7 discusses the restrictions imposed on strategies and the consequences of relaxing them. Section 8 concludes by describing some empirical implications of the results. Proofs not given in the text can be found in the Appendix.

2 Related literature

Asymmetric information has been introduced to relational contracts models in previous work.⁶ 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. The model that I propose is complementary to this literature in that it shows how the form of the relational contract changes when some variables are not observable. Different from this literature, I consider persistent private information about outside payoffs and study optimal contracting under different bargaining protocols. My model then emphasizes elements that are naturally absent from these papers, such as how information is revealed over time, how this affects incentive provision, the joint surplus, and the parties' individual payoffs, and how these results depend on the allocation of bargaining power.⁷ These elements are not present in the model of Bull (1987) either, which considers a setup with "honest" and "dishonest" principal types but restricts the former to offer the same contract every period and the latter to always renege with probability one.

My analysis of information revelation is related to the literature on reputation building, including Sobel (1985), Ghosh and Ray (1996), and Watson (1999, 2002). In dynamic games where players cannot observe whether the other player is of the "cooperative" or "noncooperative" type, these papers show that it is optimal to increase the stakes of the relationship gradually over time. The optimal path for contracts in many of the equilibria of my model corresponds closely to this

⁶For an analysis of relational contracts under symmetric information, see MacLeod and Malcolmson (1989). Earlier references include Klein and Leffler (1981), Shapiro and Stiglitz (1984), and Bull (1987).

⁷In studying how incentive provision is used to reveal information, the paper is also somewhat related to a variety of models that take the form of the contract to be a signal, such as Aghion and Bolton (1987), Aghion and Hermalin (1990), Spier (1992), Hermalin (2002), and Bénabou and Tirole (2003). In contrast to these articles, however, here I show that the contract itself cannot be used as a fully-revealing signal in equilibrium.

result. My approach differs from theirs, however, in that I consider a principal-agent setup where transfers between parties are allowed and incentive provision plays a role. Additionally, in these papers, cooperative types never have incentives to cheat, while noncooperative types always do, so no trade takes place if a player is known to be noncooperative. I take a different view in this respect. In my model, *all* principal types cooperate in some way under symmetric information, and *all* types may want to misbehave in some way under asymmetric information. In Section 7, I show that these different modeling assumptions lead to contrasting results regarding the inefficiencies caused by asymmetric information when transfers are allowed.

Another related literature is that concerning the ratchet effect, as pioneered by Freixas, Guesnerie, and Tirole (1985) and Laffont and Tirole (1988). In a dynamic contracting setting with asymmetric information, this literature shows that if an informed party reveals himself to have a high productivity or valuation for a good, the uninformed party will recontract and extract the informed party's rents in subsequent periods. This ratchet effect leads to limitations in information revelation; that is, to equilibria with (partial) pooling. This mechanism is also present in the model developed below, but here its relevance as a determinant of final outcomes depends on the bargaining power structure.

The role played by the bargaining power structure has been studied in a variety of 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. Their findings are consistent with the results of this paper if symmetric information is imposed.⁸

The paper is also related to the literature on countervailing incentives, beginning with the work of Lewis and Sappington (1989). This literature shows that, as in this paper, a party may have incentives to overstate or understate her private information depending on the realization of this information. These countervailing incentives are shown to arise in a standard agency setting where not only the agent's marginal utility, but also the agent's outside option, is assumed to depend on the agent's type. The typical example considers a regulated firm whose marginal costs are increasing in its type (à la Baron and Myerson, 1982), but its outside opportunities are decreasing in its type. In the model presented below, in contrast, only the outside option of the informed party is assumed to depend on this party's type; countervailing incentives arise endogenously due to the structure of self-enforcing contracts.

Finally, the need to think more carefully about the standard assumption that the parties' outside options are observable has been stressed in other contexts, including the literature on the boundaries of the firm (see Holmström, 1999, Williamson, 2000, and Schmitz, 2006). More generally, this paper is related to the literature on bargaining with incomplete information; see Rubinstein (1985) and Admati and Perry (1987).

⁸Also related is the model of Baker, Gibbons, and Murphy (2002), which stresses how the decision to vertically integrate affects the parties' outside options and hence the relational contracts they can implement.

3 The model

3.1 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 $\delta \in (0, 1)$.⁹ The agent's per-period outside option (reservation value) is r_A , which is fixed and known, and the principal's is r_θ , where $\theta \in \{\ell, h\}$ is the principal's type (private information) and $r_\ell < r_h$. The principal learns her type at the beginning of the first period and this type remains the same for all periods.

The sequence of events, shown schematically in Figure 1, is as follows. At the beginning of period t , the principal makes a take-it-or-leave-it (TIOLI) offer to the agent with probability $\lambda \in [0, 1]$, and the agent makes a TIOLI offer to the principal with probability $1 - \lambda$. The superscript $i \in \{A, P\}$ is used to indicate the party that makes the offer in the current period. The other party can then accept or reject this offer, where $d^j \in \{0, 1\}$, $j \in \{A, P\}$, denotes this party's decision. 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 stochastic output, $y_t \in \{y, \bar{y}\}$, for the principal. The probability that $y_t = \bar{y}$ given effort e is $f(e) \in (0, 1)$. Output is observed by both the principal and the agent, but it cannot be verified by a third party. As usual, it is assumed that the Mirrlees-Rogerson conditions are satisfied.¹⁰

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 output-contingent bonus $b_t(y_t)$, where, without loss of generality, $b_t \in \{\underline{b}_t, \bar{b}_t\}$ commensurate with y, \bar{y} and $\underline{b}_t \leq 0$, $\bar{b}_t \geq 0$. When not confusing, a contract $\{w, b\}$ is simply denoted by b . I say that $b \geq b'$ if b is weakly "steeper" than b' ; that is, $\bar{b} - \underline{b} \geq \bar{b}' - \underline{b}'$. The fixed wage is enforceable, but, since output is nonverifiable, the bonus payment is not. If $\bar{b}_t > 0$, the principal has the decision whether to honor or renege on the bonus payment at the end of period t ; if $\bar{b}_t < 0$, the agent has this decision. (Note that there is no limited liability.) Total compensation is denoted by W_t , where $W_t = w_t + b_t(y_t)$ if promised payments are honored, and $W_t = w_t$ if they are not. The agent's per-period payoff is then $W_t - c(e_t)$ and the principal's is $y_t - W_t$. The per-period expected surplus is $s(e) \equiv \mathbb{E}_y[y - c|e]$, and the per-period expected transfer to the principal when the agent makes the offer is $r_P(e) \equiv \mathbb{E}_y[y - W^A|e]$.

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

⁹More realistically, this infinite-horizon game can be interpreted as a game that ends at a random date: in any period, the probability that the relationship can continue in the following period is exogenously given by δ .

¹⁰The Mirrlees-Rogerson conditions are the monotone likelihood ratio condition and the convexity of the distribution function condition. These conditions ensure that the first-order approach is valid: the agent's incentive compatibility constraint for effort can be replaced with the first-order condition of his optimization problem. See Rogerson (1985).

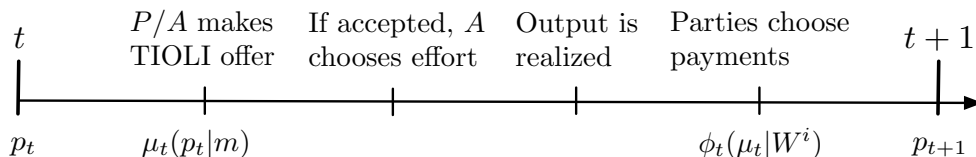


Figure 1: Timing and beliefs

agent assigns a probability p_t to the principal being of the low type, where p_0 is the prior belief. Within period t and before output is realized, the agent updates his belief to $\mu_t(p_t|m)$ if he receives a message m from the principal. This message is the principal's proposed contract if she makes the offer at time t , or the principal's participation decision if the agent makes the offer at time t . More precisely, the agent's belief is $\mu_t(p_t|b^P)$ if the principal offers b^P , $\mu_t(p_t|b^A)$ if the principal accepts b^A , and $\mu_t(p_t|\text{reject } b^A)$ if the principal rejects b^A . Given a posterior $\mu_t(p_t|b^i)$, the probability that the agent assigns to contract b^i being honored by the principal is $\chi_t(\mu_t|b^i)$. The agent then chooses effort and output is realized. If $y_t = \bar{y}$, the agent updates his belief to $\phi_t(\mu_t|W^i)$ upon observing the principal's payment, where $W^i = w^i + \bar{b}^i$ if the principal honors the contract and $W^i = w^i$ if the principal reneges. (Note that if $y_t = \underline{y}$, there is no payment decision to be made by the principal.)

I multiply expected lifetime payoffs by $(1 - \delta)$ to express them as a per-period average; I call these normalized payoffs, or simply payoffs. The principal and agent's expected payoffs at time t are respectively

$$\begin{aligned} \pi_{\theta t} &= (1 - \delta) \mathbb{E} \sum_{\tau=t}^{\infty} \delta^{\tau-t} [d_{\tau}(y_{\tau} - W_{\tau}) + (1 - d_{\tau})r_{\theta}], \\ u_t &= (1 - \delta) \mathbb{E} \sum_{\tau=t}^{\infty} \delta^{\tau-t} [d_{\tau}(W_{\tau} - c(e_{\tau})) + (1 - d_{\tau})r_A], \end{aligned}$$

and 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 TIOLI 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.

3.2 Equilibrium concepts

I characterize Pareto-optimal contracts using two equilibrium concepts, perfect public Bayesian equilibrium (PPBE) and weak Markov perfect Bayesian equilibrium (WMPBE). 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.¹¹ 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 strategies on the sequence of past offers, output realizations, and participation and payment decisions, and on her type in the case of the principal, but the agent does not condition his strategy on his past, unobserved effort decisions. This is intuitive, as the agent's past effort choices do not affect his belief nor payoff, nor the principal's continuation strategy, in any way.

Formally, let $h_t = (b_t^i, d_t^j, y_t, W_t^i)$ denote the public outcome at time t , and $h^t = (h_0, \dots, h_{t-1})$ the public history up to time t . A public strategy for type θ is a triple $\sigma_{\theta t} = (g_{\theta}(h^t, b), a_{\theta}(h^t, b), k_{\theta}(h^t, b^i))$, where g_{θ} (for "gives") is the probability with which θ offers a contract b (when $i = P$), a_{θ} (for "accepts") is the probability with which θ accepts a contract b (when $i = A$), and k_{θ} (for "keeps a promise") is the probability with which θ honors a contract b offered by i (when $y = \bar{y}$). A public strategy for the agent is analogously defined as $\sigma_{At} = (g_A(h^t, b), a_A(h^t, b), e(h^t, b^i), k_A(h^t, b^i))$, where e is the agent's effort choice.

Following Fudenberg and Tirole (1991), I require that beliefs be updated using Bayes' rule not only on the equilibrium path, but also in continuation games that are reached with zero probability. I add a condition on how beliefs are updated when they are zero or one, as Bayes' rule cannot be applied in such a case. (Of course, this is only a problem off the equilibrium path; otherwise, the parties simply play under complete information.) Assumption A1 below says that, once the agent reaches the conclusion that the principal is of a certain type, he cannot revise his belief, even if the principal does not follow that type's strategy (see Rubinstein, 1985). This assumption is not necessary for the qualitative results derived below, but it is convenient to obtain exact and simple expressions for deviation payoffs.

Assumption A1. *If the agent's posterior belief is zero for some history, it remains zero for all subsequent histories, and if it is one for some history, it remains one for all subsequent histories.*

As mentioned in the introduction, even under the requirements of PPBE, the set of Pareto-optimal equilibria is very 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 A2 says that if a default occurs, the relationship ends with some positive probability, and continues on the (constrained-)Pareto-optimal frontier otherwise. This assumption is without loss of generality. Later I show that if a Pareto-optimal equilibrium exists, there exists a Pareto-optimal equilibrium that satisfies A2 and gives the same expected payoffs to all the parties. I introduce A2 upfront for expositional convenience. See Section 7 for further discussion.

Assumption A2. *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 γ_{t+1} .*

¹¹PPBE is the natural extension of perfect public equilibrium (Fudenberg, Levine, and Maskin, 1994) for dynamic Bayesian games.

Assumption A3 says that if no default occurs, play always remains on the Pareto-optimal frontier (both on and off the equilibrium path). This implies that the parties cannot credibly threaten to end the relationship or switch to any inefficient continuation play following a rejection or an unexpected offer. This assumption is not without loss. On the other hand, any Pareto-optimal equilibrium under A3 is also Pareto optimal when A3 is not imposed, so this assumption only reduces the set of Pareto-optimal equilibria. See Section 7 for further discussion.

Assumption A3. *If no party reneges on a payment at time t , the relationship continues on the Pareto optimal frontier with probability one at time $t + 1$.*

Finally, for some of the results, I focus on a weak version of Markov perfect Bayesian equilibrium (MPBE). An MPBE requires the strategies to be Markov; that is, to depend only on payoff-relevant past events (including a player's own payoff-relevant private information). Consistent with assumptions A2-A3, I am interested in a simple form of weak Markov strategies.¹² I require that the parties' decisions to continue the relationship at time t depend only on payoff-relevant past events and the last period's history h_{t-1} , and that, conditional on the relationship continuing at time t , their moves depend only on payoff-relevant past events. (Trivially, after the relationship ends, continuation strategies also depend on payoff-relevant past events only.) In addition, although the Markov assumption does not technically prevent off-the-equilibrium-path beliefs to change when no new information arrives, I believe that such a change in beliefs is contrary to the spirit of the Markov equilibrium. Indeed, allowing for this change in beliefs would render the Markov assumption ineffective, as strategies depend on off-the-equilibrium-path beliefs. Hence, I also require that beliefs depend only on payoff-relevant past events.

In the model, the only payoff-relevant information at the beginning of period t is given by the agent's conditional beliefs μ_t .¹³ Within period t , payoff-relevant information also includes the party who makes the offer i , the offer b^i , whether it is accepted d^j , and the realized level of output y . If the parties continue the relationship at time t , type θ then chooses $(g_\theta(\mu_t, b), a_\theta(\mu_t, b), k_\theta(\mu_t, b^i))$, and the agent chooses $(g_A(\mu_t, b), a_A(\mu_t, b), e(\mu_t, b^i), k_A(\mu_t, b^i))$.¹⁴ It is also useful to define $\pi_\theta^i(\mu, b^A, b^P)$ and $u^i(\mu, b^A, b^P)$ as the type θ and agent's normalized expected payoffs when i makes the offer in the current period, the agent's conditional belief is μ , and the contract offered by i is b^i . When not confusing, I simply denote these payoffs by $\pi_\theta^i(\mu, b)$ and $u^i(\mu, b)$. Then for $\theta \in \{\ell, h\}$,

$$\begin{aligned}\pi_\theta(\mu, b) &= \lambda \pi_\theta^P(\mu, b) + (1 - \lambda) \pi_\theta^A(\mu, b), \\ u(\mu, b) &= \lambda u^P(\mu, b) + (1 - \lambda) u^A(\mu, b),\end{aligned}$$

¹²See Fudenberg, Levine, and Tirole (1985) and Ausubel and Deneckere (1989).

¹³Different from other models such as that of Mailath and Samuelson (2001), the Markov state is defined to be μ and not just p . The reason is that here, there are off-the-equilibrium-path messages, and thus off-the-equilibrium-path beliefs that affect strategies. Note, however, that μ is a function of payoff-relevant information only.

¹⁴To define the weak Markov strategies formally, let the parties make a decision to continue or end the relationship at the beginning of period t . Denote the probabilities with which type θ and the agent continue the relationship by $\gamma_{\theta t}$ and $\gamma_{A t}$ respectively. Let $\Gamma_t = 1$ if the principal and agent's observed decisions at time t are both to continue, and $\Gamma_t = 0$ otherwise. The weak Markov strategies for type θ and the agent can be defined as $\sigma_{\theta t}^{wm} = (\gamma_\theta(\mu_t, h_{t-1}), g_\theta(\Gamma_t, \mu_t, b), a_\theta(\Gamma_t, \mu_t, b), k_\theta(\Gamma_t, \mu_t, b^i))$ and $\sigma_{A t}^{wm} = (\gamma_A(\mu_t, h_{t-1}), g_A(\Gamma_t, \mu_t, b), a_A(\Gamma_t, \mu_t, b), e(\Gamma_t, \mu_t, b^i), k_A(\Gamma_t, \mu_t, b^i))$.

and the normalized expected surplus is $s_\theta(\mu, b) = \pi_\theta(\mu, b) + u(\mu, b)$.

A formal definition of the solution concepts follows.

Definition 1. Let $g_\theta(h^t, b) \equiv g_\theta$, $a_\theta(h^t, b) \equiv a_\theta$, $k_\theta(h^t, b^i) \equiv k_\theta^i$. A PPBE is a quintuple $(\sigma_\ell, \sigma_h, \sigma_A, \mu, \phi)$ such that assumptions A1-A3 are satisfied and

1. σ_ℓ, σ_h , and σ_A are mutual best responses for all t and h^t ,
2. $\mu(p|b^P) = \frac{p g_\ell}{p g_\ell + (1-p)g_h}$ for all b^P s.t. $g_\theta > 0$ for some θ ,
3. $\mu(p|b^A) = \frac{p a_\ell}{p a_\ell + (1-p)a_h}$ for all b^A s.t. $a_\theta > 0$ for some θ ,
4. $\mu(p|\text{reject } b^A) = \frac{p(1-a_\ell)}{p(1-a_\ell) + (1-p)(1-a_h)}$ for all b^A s.t. $(1-a_\theta) > 0$ for some θ ,
5. $\phi(\mu(p)|w^i + \bar{b}^i) = \frac{\mu(p|b^i)k_\ell^i}{\mu(p|b^i)k_\ell^i + (1-\mu(p|b^i))k_h^i}$ for all b^i s.t. $k_\theta^i > 0$ for some θ ,
6. $\phi(\mu(p)|w^i) = \frac{\mu(p|b^i)(1-k_\ell^i)}{\mu(p|b^i)(1-k_\ell^i) + (1-\mu(p|b^i))(1-k_h^i)}$ for all b^i s.t. $(1-k_\theta^i) > 0$ for some θ .

A WMPBE is a PPBE where strategies are weak Markov and beliefs are Markov as defined above.

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

3.3 Symmetric information benchmark

As a benchmark, I consider the case of symmetric information about the value of the relationship, where the principal's type is observed by both the principal and the agent. (I always maintain the assumption that the agent's effort choice is his private information.) This case is equivalent to Levin (2003)'s relational contracts model with hidden action. The only difference is that, here, the party who proposes compensation changes randomly over time and, given assumption A3, 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 that, in this setup, 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 effort rule of the agent and the output-contingent payments do not change over time, and the fixed payment only changes depending on which party makes the offer. Hence, the respective normalized payoffs for type θ and the agent are

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

For the compensation schedule to be self-enforcing, neither party can wish to renege on a promised payment. Since, here, no party ever reneges on the equilibrium path, 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

$$\begin{aligned}\delta\pi_\theta &\geq (1-\delta)\bar{b}(\theta) + \delta r_\theta, \\ \delta u &\geq -(1-\delta)\underline{b}(\theta) + \delta r_A.\end{aligned}$$

Depending on the bargaining power distribution, the fixed wage is adjusted and slack transferred from one constraint to the other. Thus, the two conditions above can 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 $\theta \in \{\ell, h\}$,

$$\max_{e(\cdot), \underline{b}(\cdot), \bar{b}(\cdot)} s = \mathbb{E}_y[y - c|e(\theta)] \quad (\text{P1})$$

$$\text{subject to } e(\theta) \in \arg \max_e f(e)(\bar{b}(\theta) - \underline{b}(\theta)) - c(e), \quad (\text{IC}_A)$$

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

The solution is denoted by e_θ, b_θ , and the maximized surplus by $s(e_\theta) \equiv s_\theta$. I also denote $f(e_\theta) \equiv f_\theta$, $c(e_\theta) \equiv c_\theta$. As it is well known, under risk neutrality, an optimal contract makes the agent the residual claimant by setting $W(y) = w + y$ and implements the first-best level of effort, e^{FB} . However, in this setup in which output-contingent payments are not formally enforced, the enforcement constraint must also be satisfied. This constraint is tighter the lower the discount factor and the higher the parties' outside options. For things to be interesting, I assume that for $\theta \in \{\ell, h\}$, parameters are such that (E) binds but some bonus scheme is enforceable. Therefore, and given that, as just mentioned, the party making the TIOLI offer captures all the surplus, b_θ implements $e_\theta < e^{FB}$ by specifying

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

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 θ .

The principal and agent's normalized 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 $\theta = h$. More precisely,

$$\begin{aligned}\pi_\ell(1) &= \lambda(s_\ell - r_A) + (1-\lambda)r_\ell, & \pi_h(0) &= \lambda(s_h - r_A) + (1-\lambda)r_h, \\ u(1) &= \lambda r_A + (1-\lambda)(s_\ell - r_\ell), & u(0) &= \lambda r_A + (1-\lambda)(s_h - r_h).\end{aligned}$$

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 λ is sufficiently high, but $\pi_\ell(1) < \pi_h(0)$ otherwise. The agent's continuation value, on the other hand, is always decreasing in the principal's type: $u(1) \geq u(0)$ for all $\lambda \in [0, 1]$.

4 Information revelation

4.1 Contract-pooling equilibria

Consider the case of asymmetric information about the value of the relationship, where the principal's type is observed only by the principal. Let a *contract-pooling equilibrium* be an equilibrium where no information is revealed through the contract itself in finite time. If the principal makes the offer, then the probability that she offers any given contract is the same for both principal types; if the agent makes the offer and he offers more than one contract, then conditional on acceptance, the probability that the principal accepts any given contract is the same for both principal types. That is, in a contract-pooling equilibrium the contract can neither be used to signal nor to screen the principal's type. It turns out that, in pure strategies, only this class of equilibria is feasible.

Proposition 1. *Regardless of the bargaining protocol, if $\theta \in \{\ell, h\}$ is privately observed by the principal, only contract-pooling equilibria can exist in pure strategies.*

Proof. Suppose not. Then there exists an equilibrium where ℓ and h offer or accept different contracts in some period t . Let b_1 be ℓ 's contract and b_2 h 's contract. By A3, 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 $i = P$, 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. (By the definition of contract-pooling equilibrium, if $i = A$, both b_1 and b_2 are such that the principal accepts.) Then, omitting time subscripts,

$$\begin{aligned}\pi_\ell^i(1, b_1) &= (1 - \delta)\mathbb{E}_y[y - W_1^i|e_1] + \delta\pi_\ell(1), \\ \pi_h^i(0, b_2) &= (1 - \delta)\mathbb{E}_y[y - W_2^i|e_2] + \delta\pi_h(0).\end{aligned}$$

Now note that if ℓ deviates to b_2 , she obtains

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

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

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

where $\pi_h(1)$ is h 's continuation value when the agent's posterior is one. It can be shown that $\pi_h(1) = \pi_\ell(1) + \rho$ for $\rho > 0$, where ρ is h 's opportunism option (defined below). Hence,

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

A contract-separating equilibrium exists if and only if for $\delta \in (0, 1), \lambda \in [0, 1]$,

$$\begin{aligned} \pi_\ell^i(1, b_1) &\geq \pi_\ell^i(0, b_2), \\ \pi_h^i(0, b_2) &\geq \pi_h^i(1, b_1). \end{aligned}$$

But equations (1) and (2) imply that one of these two incentive compatibility constraints is always violated. Contradiction.

Suppose next that if $i = P$, either 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 ℓ 's expected payoffs under b_1 increases. If b_2 is rejected, h has incentives to deviate to b_h , which (by A3) is always accepted and gives h a continuation value at least as high as b_2 .

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

Let incentive distortions be defined relative to the symmetric-information contract. Then,

Corollary 1. *Regardless of the bargaining protocol, if $\theta \in \{\ell, h\}$ is privately observed by the principal, a separating equilibrium without distortions does not exist.*

Proof. This result follows from Proposition 1. \square

To gain intuition, consider first the result stated in Corollary 1. If the agent believes that each principal type will choose her symmetric-information contract, then either the high type may mimic the low type, or the low type may mimic the high type. Consider the high type's deviation. By distorting incentives upward, the high type can increase the expected surplus and then renege and walk away from the relationship with a high payoff when output is high. On the other hand, if the agent makes the offer in the current period, this deviation entails a cost, as the high type must accept a low expected transfer $r_P = r_\ell$ to signal a low type.¹⁵ Thus, pretending to be the low type is profitable for the high type if the principal makes the offer in the current period, or if the agent does but the principal's bargaining power and the discount factor are sufficiently high.

Consider next the low type's deviation. By distorting incentives downward, the low type can signal that her outside payoff is high and then receive a high expected transfer in periods in which the agent makes the offer. On the other hand, weakened incentives also result in a lower payoff for the low type in periods in which the principal makes the offer. Thus, pretending to be the high type is profitable for the low type if the agent's bargaining power and the discount factor are sufficiently high, or if the agent makes the offer in the current period and the discount factor is sufficiently low.

¹⁵Given A1, once the high type has signaled that her type is low in period t , she can reject the agent's offer in subsequent periods without affecting beliefs.

The proof of Proposition 1 shows that (at least) one of the two deviations described above is always profitable for some principal type. As a result, informational asymmetries about the value of the relationship inevitably lead to distortions in incentive provision.

But Proposition 1 says more than that. It not only shows that the principal types will never immediately separate by choosing their symmetric-information contracts, but also that they will never separate by choosing *any* two different contracts with probability one. This means that it is not possible to design two contracts, involving any sort of transfers and given any bargaining protocol, such that deviations by both types are simultaneously deterred. Hence, any equilibrium in pure strategies is contract-pooling, and incentive distortions relative to the benchmark are arguably not negligible.

To prove that some principal type always has incentives to deviate, it is shown that the low type can obtain the same payoff as the high type by choosing the high type's contract, while the high type can obtain a payoff strictly higher than the low type's by choosing the low type's contract. 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—the high type can earn a rent by renegeing on the bonus payments and by rejecting the agent's offers once the agent is convinced that the principal's type is low. When the agent's posterior belief is equal to one, this opportunism option has a simple form.

Lemma 1. *Let $\pi_h(1) = \pi_\ell(1) + \rho$, where ρ is the opportunism option that the high type enjoys when the agent believes that the principal's type is low with probability one. Then,*¹⁶

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

4.2 Default and rejection

While Proposition 1 shows that full separation of types cannot occur through contracts, full separation may still be possible in equilibrium. In particular, full separation can occur through default—a principal type reneges on the contract while the other honors it—or rejection—a principal type rejects the agent's offer while the other accepts it.

Since the low type has a lower outside opportunity than the high type, an equilibrium in which the principal's type is revealed through default is one in which the low type honors and the high type reneges. But then, such an equilibrium exists only if the low type prefers to honor rather than to renege and signal that her type is high (to then receive high transfers from the agent if the relationship continues), and the high type prefers to renege rather than to honor and signal that her type is low (to then renege on a larger bonus payment in a future period). Hence,¹⁷

¹⁶This expression for ρ is derived using assumption A1. However, while some assumption on beliefs like A1 is necessary to compute ρ , this condition is by no means necessary to show that $\rho > 0$. Indeed, without A1 or any other restriction of this kind, it is straightforward to show that

$$\rho \geq \frac{\delta f_\ell}{1 - \delta(1 - f_\ell)} (r_h - r_\ell) > 0.$$

¹⁷The conditions in Lemma 2 are satisfied if the implemented contract is the symmetric-information contract of

Lemma 2. *Suppose that in equilibrium a contract b' is implemented and the principal's type is fully revealed through default. Let γ' be the probability that the relationship continues when the default occurs. Then,*

$$(1 - \delta)\bar{b}' + \delta[r_\ell + \gamma'(\pi_h(0) - r_\ell)] \leq \delta\pi_\ell(1), \quad (\text{E}_\ell)$$

$$(1 - \delta)\bar{b}' + \delta[r_h + \gamma'(\pi_h(0) - r_h)] \geq \delta\pi_h(1). \quad (\text{D}_h)$$

If the probability that the relationship continues after the principal reneges is relatively high, then the low type wants to renege whenever the high type does. As a result,

Corollary 2. *An equilibrium in which the principal's type is fully revealed through default exists only if the relationship ends with probability $(1 - \gamma) \geq (1 - \bar{\gamma})$ after the default, where*

$$\bar{\gamma} = \frac{\lambda(1 - \delta)}{1 - \delta(1 - \lambda f_\ell)}.$$

Proof. This result follows from Lemmas 1 and 2. □

Since the low type has a lower outside opportunity than the high type, an equilibrium in which the principal's type is revealed through rejection is one in which the low type accepts and the high type rejects. But then, such an equilibrium exists only if the low type prefers to accept rather than to reject and signal that her type is high, and the high type prefers to reject rather than to accept and signal that her type is low. Hence,

Lemma 3. *Suppose that in equilibrium the agent proposes a contract b' with an expected transfer to the principal $r'_P = \mathbb{E}_y[y - W'|e']$, and the principal's type is fully revealed through rejection. Then,*

$$(1 - \delta)(r'_P - r_\ell) \geq \delta(\pi_h(0) - \pi_\ell(1)),$$

$$(1 - \delta)(r_h - r'_P) \geq \max\{\delta\pi_h(1), f(e')[(1 - \delta)\bar{b}' + \delta r_h] + (1 - f(e'))\delta\pi_h(1)\} - \delta\pi_h(0).$$

If the discount factor is relatively high, then the low type wants to reject whenever the high type does. As a result,

Corollary 3. *An equilibrium in which the principal's type is fully revealed through rejection exists only if $\delta \leq \bar{\delta}$, where*

$$\frac{\bar{\delta}}{1 - \bar{\delta}} = \frac{1 - \bar{\delta}(1 - \lambda f_\ell)}{(1 - \bar{\delta})(1 - \lambda) + \bar{\delta}\lambda f_\ell}.$$

Proof. This result follows from Lemmas 1 and 3. □

Thus, while the discount factor must be sufficiently high for relational incentives to be self-enforcing, it must be sufficiently low for revelation through rejection to occur. This tension will

the low type, prescribing termination with probability one after default, as in this case the low type cannot gain from reneging and the high type cannot gain from staying in the relationship. However, this contract may not be implementable in equilibrium. Revelation through default is feasible only if a contract satisfying conditions (E_ℓ) and (D_h) is implementable.

often result in revelation through rejection not being feasible.

5 Characterization of the equilibria

An implication of Proposition 1 is that fixed transfers cannot be used to reveal the principal’s type. In what follows, I then assume that the agent’s belief cannot depend on the fixed wage that the principal proposes (but can depend on the proposed bonus payments). This assumption is without loss of generality.¹⁸ I introduce it to simplify the analysis and to disregard equilibria where the agent’s beliefs are used as a threat to increase the equilibrium wage and transfer the principal’s rents to the agent. I consider such equilibria “unreasonable” in light of Proposition 1, as it is then commonly known that, given a bonus scheme, *any* principal type would always strictly prefer to offer a wage that gives her a higher surplus, provided that the agent’s beliefs are unchanged. In fact, different refinements proposed in the literature would rule out these equilibria.¹⁹

Assumption A4. *The agent’s belief about the principal’s type at any stage of the game is independent of the fixed wages proposed by the principal up to that stage.*

The next subsections characterize optimal relational contracts under assumptions A1-A4 for different allocations of bargaining power.

5.1 When the principal has all the bargaining power

I begin by considering a setting in which the principal has all the bargaining power; that is, $\lambda = 1$. I first analyze what classes of equilibria are feasible and then identify and characterize those that are Pareto optimal. Because all offers are made by the principal, I omit the superscript i throughout this subsection.

Allocating all the bargaining power to the principal has several direct implications for contracts and incentives. First, because the value of the relationship for the agent is zero, 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 pretend to be the high-outside-option type, and it is thus only the high type who wants to misrepresent her type.

The first result concerns equilibria with no revelation. It is straightforward to show that such equilibria can be sustained by the threat that the agent will believe that the principal’s type is high if a contract that induces revelation is offered. On the other hand, by assumptions A3 and A4,

¹⁸The results on how revelation optimally occurs as well as those regarding the joint surplus generated by the relationship are unchanged if this assumption is dropped. The main insights on how the parties’ individual payoffs may be affected by the allocation of bargaining power (Proposition 13) are strengthened if this assumption is dropped.

¹⁹For example, the perfect sequential equilibrium refinement introduced by Grossman and Perry (1986) would eliminate these equilibria. In their terminology, any updating rule that supports a principal’s proposed wage that gives the agent an expected payoff strictly above his outside option is not “credible.” An equivalent refinement is also used by Rubinstein (1985); see assumption B-2 in his paper.

regardless of the agent's beliefs, the principal can always offer the high type's symmetric-information contract b_h and obtain the high type's symmetric-information payoff $\pi_h(0)$. Thus,

Proposition 2. *A pooling equilibrium always exists. Moreover, any such equilibrium implements b_h in all periods.*

I next turn to the analysis of equilibria with revelation. Consider contract-pooling equilibria, where both types offer the same contract b^* and thus revelation occurs only through default. Determining whether such an equilibrium exists can be simplified by two observations. First, given Proposition 2 above, the payoff to the principal on and off the equilibrium path is always at least as large as the high type's symmetric-information payoff $\pi_h(0)$, regardless of the principal's true type. Second, since, in this class of equilibria, the high type has a positive option to renege relative to the low type, one can concentrate on the low type's problem. In sum, existence can be shown by finding a contract that induces default and, in each period, gives the low type an expected discounted payoff at least as large as $\pi_h(0)$.

Program (P2) in Definition 2 below does exactly this. Note that, given assumptions A2-A3 (and the probability of termination following default), the continuation play after a default occurs and the principal's type is revealed is fixed. I then take this as given and consider only play until full revelation is induced. If b^* is the equilibrium pooling contract, the agent's belief is $\mu(p|b^*) = p$. Then let $\phi(p|w+\bar{b}) \equiv p_b$, $\phi(p_b|w+\bar{b}) \equiv p_{bb}$, and so on. Define $\mathbf{k}_h = (k_h(p_0), k_h(p_{0b}), \dots)$ and $\gamma = (\gamma(p_0), \gamma(p_{0b}), \dots)$ to be respectively a sequence of probabilities with which the high type honors the pooling contract and a sequence of probabilities with which the relationship continues after default, given the agent's posterior belief and until full revelation is induced (that is, until the agent's posterior reaches $p = p'$ such that $k_h(p') = 0$ and thus $p'_b = 1$).²⁰ Given some \mathbf{k}_h and γ , (P2) finds corresponding sequences of contracts and effort levels, $\mathbf{b}^* = (b^*(p_0), b^*(p_{0b}), \dots)$, $\mathbf{e}^* = (e^*(p_0), e^*(p_{0b}), \dots)$, that maximize the low type's expected payoff when the strategies are (weak) Markov and revelation is induced through default.²¹ Note that revelation in this case requires that the high type renege with probability one in finite time, and with strictly positive probability in each period. The former follows from the fact that, for the high type to renege with probability strictly between zero and one in a given period, she must be indifferent between honoring and renegeing, so she must be able to renege on a larger payment or with a lower punishment for default in a future period, and eventually renege with probability one. The latter follows from the fact that if the high type reneges with probability zero in a given period, the agent's posterior and hence the principal's Markov strategy do not change in subsequent periods, so revelation cannot occur in a Markov equilibrium. Therefore, (P2) considers sequences \mathbf{k}_h that are finite, with last element $k_h(p) = 0$, and with all elements strictly less than one. Additionally, since in equilibrium the agent's belief that the pooling contract will be honored is correct, I specify $\chi(p) = p + (1-p)k_h(p)$, and, by consistency of beliefs, I specify $p_b = p/\chi(p)$.

²⁰Note that $k_\ell(p)$ is always equal to one in equilibrium.

²¹Alternatively, the problem could be formulated such that \mathbf{k}_h and γ are choice variables. I formulate it in the way it is for consistency with other results presented below, which use \mathbf{k}_h and γ to characterize the equilibria.

The low type's expected payoff is maximized subject to incentive compatibility, revelation, and enforcement constraints. In particular, (IC_A) is the incentive compatibility constraint for the agent's effort, which takes into account the probability $\chi(p)$ that the contract will be honored; (I_h) is an indifference condition for the high type that holds whenever she honors with strictly positive (and strictly less than one) probability, which takes into account the probability $\gamma(p)$ that the relationship will continue after default; (D_h) and (E_ℓ) are respectively a default condition for the high type that holds whenever she honors with zero probability and an enforcement constraint for the low type that ensures that she always honors, both of which take into account the probability $\gamma(p)$, as derived in Lemma 2;²² and (R_ℓ) is a revelation constraint that ensures that the low type will not deviate to the no-revelation contract b_h in future periods. The agent's participation constraint, which always binds, is already incorporated in the expression for the principal's expected payoff.

Definition 2. Given $\mathbf{k}_h = (k_h(p_0), k_h(p_{0b}), \dots, 0)$, $k_h(p) \in [0, 1)$, $\boldsymbol{\gamma} = (\gamma(p_0), \gamma(p_{0b}), \dots)$, $\chi(p) = p + (1-p)k_h(p)$, $p_b = p/\chi(p)$, $\pi_\ell(1, b) = \pi_\ell(1)$, let $\mathbf{b}^* = (b^*(p_0), b^*(p_{0b}), \dots)$, $\mathbf{e}^* = (e^*(p_0), e^*(p_{0b}), \dots)$ be the solution to

$$\max_{e, \mathbf{b}} \pi_\ell(p_0, b) = \frac{(1-\delta)[s(e(p_0)) - (1-\chi(p_0))f(e(p_0))\bar{b}(p_0) - r_A] + \delta f(e(p_0))\pi_\ell(p_{0b}, b)}{1-\delta(1-f(e(p_0)))} \quad (P2)$$

$$\text{subject to } e(p) \in \arg \max_e \chi(p)f(e)\bar{b}(p) - c(e), \quad (IC_A)$$

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

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

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

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

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

Using Definition 2, the following proposition provides a necessary and sufficient condition for contract-pooling equilibria with revelation to exist.

Proposition 3. A contract-pooling Markov equilibrium with revelation (through default) exists if and only if there exist $\mathbf{k}_h^* = (k_h^*(p_0), k_h^*(p_{0b}), \dots, 0)$, $k_h^*(p) \in [0, 1)$, and $\boldsymbol{\gamma}^* = (\gamma^*(p_0), \gamma^*(p_{0b}), \dots)$ 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,

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

A contract-pooling equilibrium with revelation exists if and only if a contract-pooling Markov equilibrium with revelation exists.

²²Note that conditions (I_h) , (D_h) , and (E_ℓ) imply bounds on $\gamma(p)$.

The first part of Proposition 3, concerning Markov equilibria, follows directly from program (P2) and the observations made above. The second part, concerning non-Markov equilibria, is arguably less expected. The main intuition is that revelation is feasible only if the low type prefers inducing to not inducing revelation in every subgame in which revelation is induced. Thus, if there are no Markov strategies that constitute an equilibrium with revelation, then there are no time-dependent strategies that will do so.

Condition (3) has a simple interpretation: revelation can be induced in a contract-pooling 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 value of the relationship is likely to be high (p_0 is high) and the difference between the two types' symmetric-information surpluses is large ($s_\ell - s_h$ is large), the benefits indeed exceed the costs and information revelation is feasible. Otherwise, the low type prefers to offer the pooling contract b_h and induce no revelation.²³

Revelation may also be induced, though only partially, through contracts. Note, however, that a contract-semi-separating equilibrium with no default cannot exist, as both types would want to pool in b_h . And an equilibrium where the low type offers a contract b' and the high type mixes between this and another contract b'' cannot exist either, as the high type would reveal herself by offering b'' (and thus never be indifferent between b' and b'' when the low type prefers b'). Therefore, a contract-semi-separating equilibrium exists only if the high type offers a contract b' on which she reneges, and the low type mixes between this and another contract b'' , where $b'' < b'$. But then such mixed strategy lowers effort incentives relative to a contract-pooling equilibrium—the probability that b' will be honored is lower than if both types pool in b' . This leads to the following results.

Proposition 4. *A contract-semi-separating equilibrium that implements contracts b' and $b'' < b'$ exists only if there exists a contract-pooling equilibrium that implements b' .*

Corollary 4. *A separating equilibrium exists if and only if (3) holds.*

Proof. This result follows from Propositions 3 and 4. □

Since existence of a separating equilibrium requires the low type, and thus the high type, to prefer such equilibrium to the pooling equilibrium, it follows that separation is Pareto optimal whenever it is feasible. Additionally, the reasoning behind Proposition 4 above implies that contract-semi-separation is never Pareto optimal. Hence, the Pareto-optimal equilibria are as follows,

Proposition 5. *If (3) holds, any Pareto-optimal equilibrium is contract-pooling with revelation (through default). If (3) does not hold, the unique Pareto-optimal equilibrium is the pooling one.*

For the set of Pareto-optimal equilibria that induce revelation of information, I next study how revelation occurs on the equilibrium path—how fast revelation is induced, how the form of

²³The comparative static for the discount factor is less straightforward (note that $\bar{b}^*(p_0)$ is a function of δ). In general, a higher discount factor will imply a smaller difference between the two types' symmetric-information surpluses, and thus reduce the benefits from inducing revelation.

the contract evolves over time, and what punishments for default are specified. I characterize the equilibria in terms of the sequence of probabilities with which the high type honors the contract, $\mathbf{k}_h = (k_h(p_0), k_h(p_{0b}), \dots, 0)$, and the sequence of probabilities with which the relationship continues following default, $\boldsymbol{\gamma} = (\gamma(p_0), \gamma(p_{0b}), \dots)$, until full revelation is induced. I restrict attention to Markov equilibria.²⁴

Definition 3. *Restrict $k_h(p) \in [0, \bar{k}_h]$ for all p and some $\bar{k}_h \in (0, 1)$. Let K be the set of sequences $\mathbf{k}_h = (k_h(p_0), k_h(p_{0b}), \dots, 0)$ for which revelation can be induced in a Pareto-optimal Markov equilibrium. Let $n_{\mathbf{k}} \in \mathbb{N}$ be the length of \mathbf{k}_h , with $n^{\min} \equiv \min_{\mathbf{k}_h \in K} n_{\mathbf{k}}$, $n^{\max} \equiv \max_{\mathbf{k}_h \in K} n_{\mathbf{k}}$.*

For a given equilibrium, $n_{\mathbf{k}}$ measures inversely the speed of revelation. For a given prior belief p_0 , n^{\min} and n^{\max} measure respectively how high and how low such speed can be in a Pareto-optimal Markov equilibrium. The larger $n_{\mathbf{k}}$ is, the more slowly revelation occurs on the equilibrium path. Revelation is *immediate* if $n_{\mathbf{k}} = 1$ and *gradual* if $n_{\mathbf{k}} > 1$. Revelation may also be *delayed* if it is induced more slowly than feasible, namely if $n_{\mathbf{k}} > n^{\min}$.

Proposition 6. *n^{\min} and n^{\max} are decreasing in p_0 . There exist $\underline{p} < \bar{p} \in (0, 1)$ such that $K = \emptyset$ for $p_0 < \underline{p}$ (no revelation), $1 < n^{\min} = n^{\max} < \infty$ for $p_0 = \underline{p}$ (gradual revelation), $1 \leq n^{\min} < n^{\max} < \infty$ for $p_0 \in (\underline{p}, \bar{p})$ (gradual or immediate, and delayed revelation), and $n^{\min} = n^{\max} = 1$ for $p_0 \geq \bar{p}$ (immediate revelation).*

Suppose for simplicity that the relationship ends whenever a default occurs, and that (E_ℓ) binds when full revelation is induced. Then to understand the results for n^{\min} , note, first, that if revelation is feasible, the high type always wants to induce revelation as fast as possible. Intuitively, if full revelation is delayed, then, during the delay, the high type must be indifferent between honoring and renegeing. But then it follows that the high type would be better off if a steeper contract were implemented, so that she would strictly prefer to renege. This, in turn, implies that n^{\min} corresponds to the highest revelation speed that is feasible in a Markov equilibrium. That is, n^{\min} is the length of the shortest sequence $\mathbf{k}_h = (k_h(p_0), k_h(p_{0b}), \dots, 0)$ for which the necessary and sufficient condition for revelation, condition (3), can be satisfied.

The results then follow directly. If the prior belief that the principal is a low type is sufficiently high, the cost of revelation for the low type—the compensation to the agent for the risk of default—is low and full revelation can be induced immediately in the first period. On the other extreme, if the prior belief is sufficiently low, revelation is too costly and hence not possible. Lastly, for intermediate values of the prior belief, immediate revelation is not feasible but gradual revelation is. The reason is that, when revelation is gradual, a sequence of contracts that induce partial revelation and progressively increase the agent’s posterior belief is implemented prior to inducing full revelation, so the cost of revelation is reduced.

²⁴In Definition 3, I specify an upper bound for the probability with which the high type honors the contract, $k(p)$. This is necessary for the set of Pareto-optimal Markov equilibria to be well defined, since a Markov equilibrium requires this probability to be strictly less than one, while, as shown below, a Pareto-optimal equilibrium may require this probability to be as high as possible.

Consider next the results for n^{\max} . From the discussion above, it is clear that $n^{\max} > n^{\min}$ only if the low type benefits from delaying revelation. Such a possibility is at first counterintuitive—recall, in this setting, the high type imposes a negative externality on the low type, suggesting that the low type would want to separate from the high type as fast as possible. However, if the prior probability that the principal’s type is low is not high enough, the low type indeed wants to delay revelation. The logic is related to, but not the same as, that for gradual revelation. In this case, full revelation is postponed not to reduce its cost through belief-updating, but rather to just defer this cost until a later time. In fact, during the delay phase, a *probationary* contract that induces minimal separation of types may be implemented.²⁵

Proposition 7. *Restrict $k_h(p) \in [0, \bar{k}_h]$ for all p and some $\bar{k}_h \in (0, 1)$. Take $p_0 \in (\underline{p}, \bar{p})$ such that $n^{\min} < n^{\max}$. Then the high type’s preferred Markov equilibrium induces revelation with $n_{\mathbf{k}} = n^{\min}$. The low type’s preferred Markov equilibrium induces revelation with $n_{\mathbf{k}} = n^{\max}$, and, for many parameter values, initially implements a probationary contract that is honored by the high type with probability \bar{k}_h .*

More generally, regardless of the speed of revelation, it is possible to characterize some common features of an optimal path of revelation. I find that as the agent’s belief that the principal is a low type increases along this path, the probability with which the high type reneges increases, and either the bonus payments or the probability with which the relationship continues after default also increase. The intuition is by now familiar. Given some fixed speed of revelation, a contract that induces a low probability of default today relative to tomorrow reduces the cost of inducing revelation (for both types). And, since the high type must be indifferent between reneging today, and honoring today and reneging tomorrow, it must be that tomorrow either the bonus payments are higher or the punishment for default is lower. Formally,

Proposition 8. *Take $\mathbf{k}_h = (k_h(p_0), k_h(p_{0b}), \dots, 0) \in K$, and corresponding $\gamma = (\gamma(p_0), \gamma(p_{0b}), \dots)$, $\mathbf{b} = (b(p_0), b(p_{0b}), \dots)$. If $p' < p''$, $k_h(p') \geq k_h(p'')$, and either $\bar{b}(p') < \bar{b}(p'')$ or $\gamma(p') < \gamma(p'')$ or both.*

The punishments for default specified in the contract until full revelation is induced, as implied by the sequence $\gamma = (\gamma(p_0), \gamma(p_{0b}), \dots)$, are further described in the next definition and proposition.

Definition 4. *Let $\Gamma_{n_{\mathbf{k}}}$ be the set of sequences $\gamma = (\gamma(p_0), \gamma(p_{0b}), \dots)$ for which revelation can be induced in a Pareto-optimal Markov equilibrium with sequence $\mathbf{k}_h = (k_h(p_0), k_h(p_{0b}), \dots, 0)$ of length $n_{\mathbf{k}}$. Let $\gamma \geq \tilde{\gamma}$ if $\gamma(p) \geq \tilde{\gamma}(p)$ for all p .*

Proposition 9. *Take $\gamma \in \Gamma_{n_{\mathbf{k}}}$. Then $\gamma(p) = 0$ for $p \geq \hat{p}$ or $k_h(p) \geq \hat{k}_h(p)$, where $\hat{p} \in (0, 1)$, $\hat{k}_h(p) \in [0, 1)$. However, depending on parameter values, it may be that $\gamma(p) > 0$ for $p < \hat{p}$, $k_h(p) < \hat{k}_h(p)$. If, given some fixed $n_{\mathbf{k}}$, the low type’s preferred Markov equilibrium specifies $\gamma \in \Gamma_{n_{\mathbf{k}}}$, then the high type’s specifies $\tilde{\gamma} \in \Gamma_{n_{\mathbf{k}}}$ where $\tilde{\gamma} \leq \gamma$.*

²⁵The term “probationary contract” is chosen in reference to contracts that promise employment for a pre-specified period of time, as are commonly used, for example, in academia. This term also has other connotations.

The idea is similar to that above. Note that if the punishment for default falls (that is, the probability that the relationship continues after default increases), the bonus payments on which the high type reneges fall, as the principal’s continuation value after default increases. As a result, if the punishment is reduced, effort incentives and thus the probability that output is high and revelation occurs fall, but the compensation to the agent for the risk of default also falls. Hence, both the expected benefits and costs of inducing revelation go down, and the optimal punishment depends on the relative importance of these two effects.

It is then straightforward that the worst punishment for default is optimal when the probability that the principal is a low type, or the probability that the high type honors, is sufficiently high, as the cost of inducing revelation is then negligible. Furthermore, given a fixed speed of revelation n_k , the punishment for default is relatively lower in the contract preferred by the low type. This again seems contrary to intuition, as it is the high type who reneges and is “punished” in equilibrium. But note that the low type’s continuation value when revelation occurs is relatively lower, so it is this type who gains more from reducing the cost of inducing revelation and, therefore, the punishment for default.²⁶

The finding that an optimal contract may prescribe termination with a probability strictly less than one when “cheating” occurs in equilibrium is reminiscent of Green and Porter (1984)’s results and, in turn, of the punishments used in relational contracts under subjective performance (see Section 2). However, the intuition is rather different. Note that, unlike in those models, there is no uncertainty on whether cheating has occurred or not—the agent knows with certainty that the principal has reneged when he does not receive the promised payment. Here, the punishment for cheating may then be reduced not to lower the inefficiency it imposes ex post, but rather to optimally balance the costs and benefits of inducing this cheating ex ante.

Figures 2 to 4 illustrate the results for two numerical examples.²⁷ I let $k(p) \in [0, \bar{k}]$ for $\bar{k} \rightarrow 1$. Figure 2 shows how the speed of revelation changes with the prior distribution of types. The solid line with square data points depicts n^{\min} , and the dashed line with triangular data points n^{\max} . In Example A, for instance, I find that the Pareto-optimal equilibrium induces revelation if and only if $p_0 \geq 0.627 \equiv \underline{p}$. If $p_0 \geq 0.685$, revelation can be immediate ($n^{\min} = 1$); otherwise, it is always gradual ($n^{\min} > 1$). If $p_0 \geq 0.95 \equiv \bar{p}$, immediate revelation is the unique Pareto-optimal outcome, whereas if $p_0 \in (0.627, 0.95)$, revelation may be delayed ($n_k > n^{\min}$), so there exist multiple Pareto-optimal equilibria. For example, for $p_0 = 0.85$, equilibria with n_k equal to one, two, and three,

²⁶When full revelation is induced, there is an additional reason why reducing the punishment for default is not beneficial for the high type. Note that in a Pareto-optimal equilibrium, when p is such that $k_h(p) = 0$, (E_ℓ) binds and (D_h) is slack. Otherwise, one could increase $\gamma(p)$ and increase the high type’s expected payoff without affecting the low type’s. But then, if (E_ℓ) binds and (D_h) is slack, an increase in $\gamma(p)$ by $d\gamma$ reduces incentives by $(\pi_h(0) - r_\ell)d\gamma$, while increasing the high type’s continuation payoff after default by only $(\pi_h(0) - r_h)d\gamma$.

²⁷Example A considers $\delta = \frac{4}{5}$, $e \in [0, 1]$, $y \in \{0, 1\}$, $f(e) = \frac{e}{2}$, $c(e) = \frac{e^2}{4}$, $r_\ell = \frac{1}{25}$, $r_h = \frac{1}{16}$, and $r_A = 0$. The first-best solution is $e^{FB} = 1$, $s^{FB} = \frac{1}{4}$. Since $\frac{\delta}{1-\delta}[s^{FB} - r_\theta - r_A] < 1$, the enforcement constraint binds for both types. The symmetric-information contracts are $b_\ell = \frac{4}{5}$ and $b_h = \frac{1}{2}$. Example B considers $\delta = \frac{4}{5}$, $e \in [0, 1]$, $y \in \{0, 1\}$, $f(e) = \frac{1}{2} + \frac{e}{4}$, $c(e) = \frac{e^2}{4}$, $r_\ell = \frac{21}{64}$, $r_h = \frac{27}{64}$, and $r_A = 0$. The first-best solution is $e^{FB} = \frac{1}{2}$, $s^{FB} = \frac{9}{16}$. Again, since $\frac{\delta}{1-\delta}[s^{FB} - r_\theta - r_A] < 1$, the enforcement constraint binds for both types. The symmetric-information contracts are $b_\ell = \frac{-2+\sqrt{15}}{2}$ and $b_h = \frac{1}{2}$.

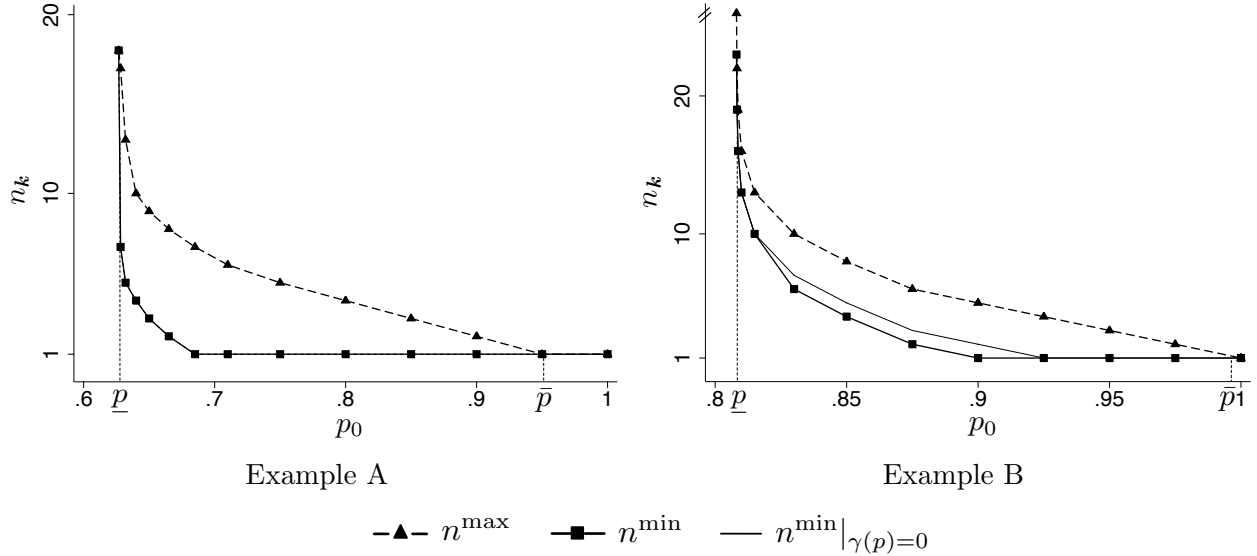


Figure 2: Speed of revelation

are all Pareto optimal. The figure also shows that revelation may take a long time: for the lowest prior belief for which revelation can occur ($p_0 = 0.627$), $n_k = 18$; that is, full separation of types is induced once and only after output has been high and the agent updated his belief 17 times.

Figure 3 describes the path of information revelation for two particular prior distributions, $p_0 = 0.8$ (top panel) and $p_0 = 0.64$ (bottom panel), in Example A. The figure depicts the probability with which the high type honors the bonus payments $k_h(p)$ (left graphs) and the bonus payments implemented $\bar{b}(p)$ (right graphs) on the equilibrium path until full revelation is induced. The x -axis is an index that starts at one and increases by one when the agent's posterior belief p changes (regardless of the magnitude of such change). In each graph, the dashed lines correspond to the low type's preferred Markov equilibrium, and the solid lines to the high type's. The figure shows that revelation is always faster in the high type's preferred equilibrium. For $p_0 = 0.8$, full revelation is induced immediately in the high type's preferred equilibrium, while it is induced only after the agent has updated his belief three times in the low type's. For $p_0 = 0.64$, revelation is always gradual, but it also takes significantly longer in the low type's preferred equilibrium. Further, when revelation is delayed, the low type's preferred equilibria show that a probationary contract is implemented: $k_h(p) = \bar{k}_h$ during the delay phase. Finally, the figure shows that the probability of separation and the bonus payments increase monotonically over time as the agent updates his posterior belief.

The numerical examples also illustrate the punishments for default. In Example A, I find that prescribing termination of the relationship with probability one whenever a party reneges is always optimal. In Example B, this is true for equilibria with $n_k = n^{\max}$, but not necessarily otherwise. To see this, consider Figure 2 again. The thin solid line depicts the highest speed of revelation that is Pareto optimal under the constraint that the relationship end with probability one following a default, $n^{\min}|_{\gamma(p)=0}$. This line is above that for n^{\min} in the graph for Example B, showing that

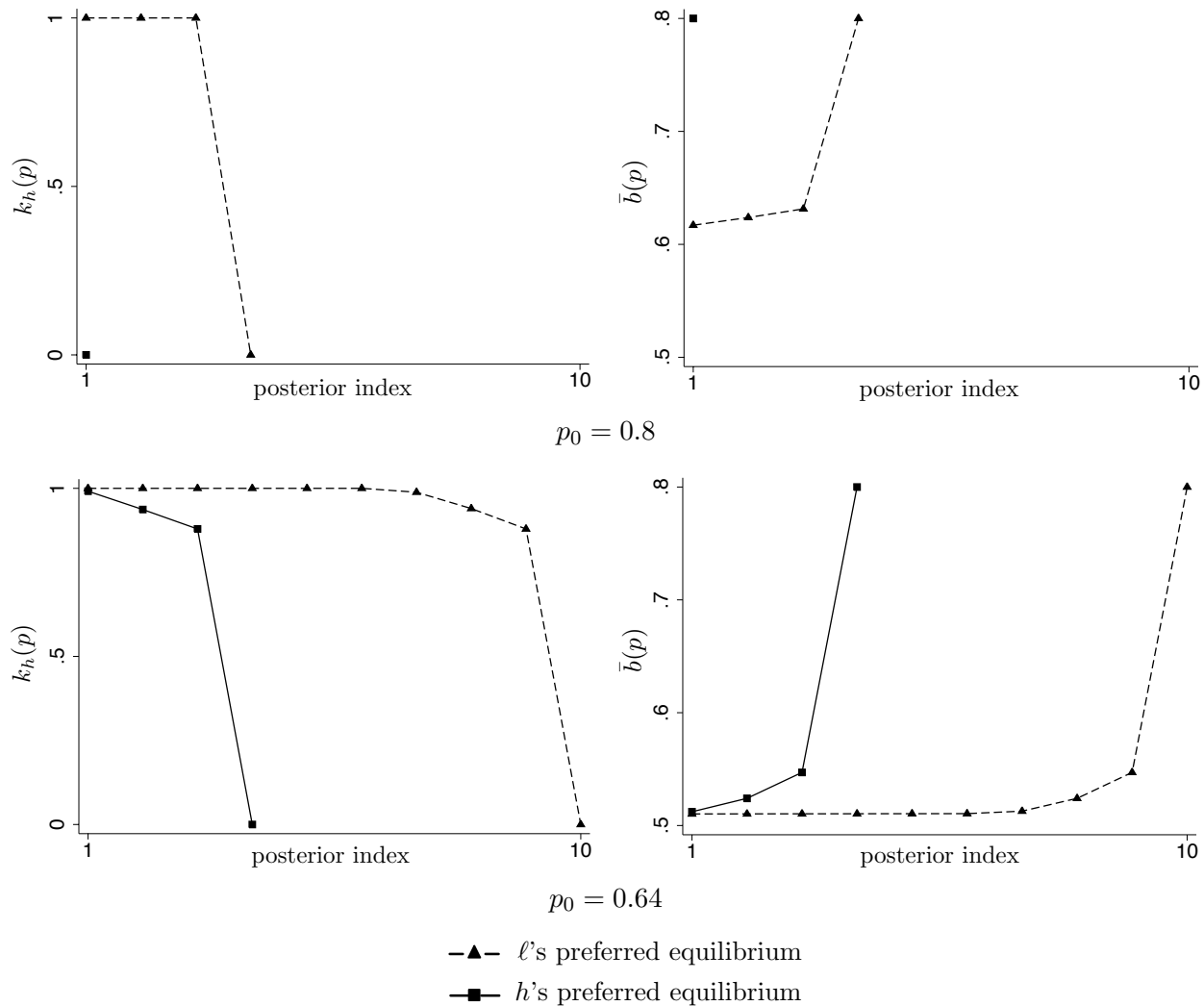


Figure 3: Information revelation paths (Example A)

the restriction on $\gamma(p)$ reduces the set of Pareto-optimal equilibria. In particular, in this example, lowering the punishment for default allows the parties to implement contracts that induce faster revelation, and that are then preferred by the high type.

Figure 4 looks at the optimal punishments for default in equilibria inducing immediate revelation ($n_k = 1$) in Example B. For the range of priors p_0 for which such equilibria exist, the dashed and solid lines depict the probability that the relationship continues after a default, $\gamma(p_0)$, in the low and high types' preferred Markov equilibria. The figure shows that this probability is zero when the agent's belief is sufficiently high, and is always higher in the low type's preferred equilibrium.

5.2 When the agent has all the bargaining power

I next consider a setting in which the agent has all the bargaining power; that is, $\lambda = 0$. Because all offers are made by the agent, I omit the superscript i throughout this subsection.

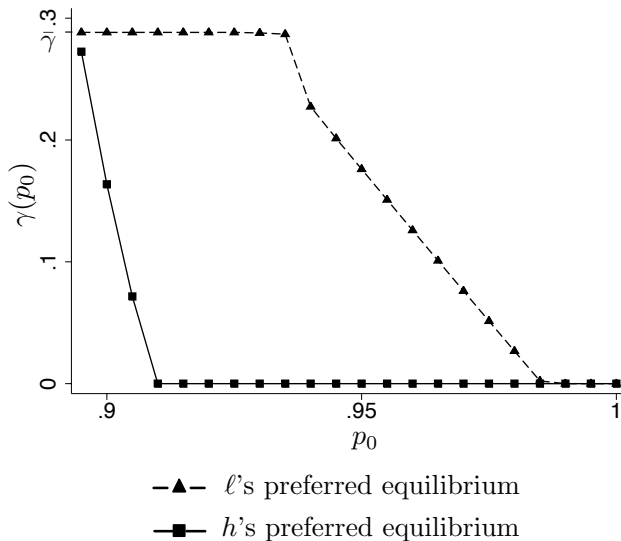


Figure 4: Punishment for default (Example B)

Allocating all the bargaining power to the agent has several direct implications for contracts and incentives. First, because, under symmetric information, the value of the relationship for the principal is zero, no payments from the principal to the agent can be enforced. Hence, $\bar{b} = 0$ in any contract.²⁸ Second, as a result of this, 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 pretend to be the low-outside-option type, and it is thus only the low type who wants to misrepresent her type.

The results for this case are similar to those described above in that whenever revelation is feasible, a contract that induces revelation is Pareto optimal, and whenever revelation is not feasible, the symmetric-information contract of the high type is Pareto optimal. However, as expected, I find that a shift in the bargaining power from the principal—the informed party—to the agent—the uninformed party—reduces the solution to a unique Pareto-optimal equilibrium.

Proposition 10. *The Pareto-optimal equilibrium is unique. If*

$$\delta \leq \frac{1}{2}, \text{ and} \tag{4}$$

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

²⁸The argument is as follows. Suppose that a contract with bonus payments $\bar{b} > 0$ is implemented. Then the high type always reneges on \bar{b} (in finite time), as the principal's expected payoff within the relationship is always weakly lower than r_h in this setting. Then it must be that the low type honors \bar{b} . But then, if no default is observed, the agent learns that the principal's type is low, and thus offers the principal a transfer $r_P = r_\ell$ from then on. But then the low type also reneges on \bar{b} , which gives a contradiction. Note that for this argument to hold, the agent cannot be able to commit to a strategy that gives the principal part of the surplus once her type is revealed to be low. Here, this is ensured by assumption A3. Without this assumption, on the other hand, such a strategy could be self-enforced by the threat that the relationship will be terminated if the agent deviates. However, in Section 7, I show that the agent would never want to do this.

for $r_P^{**} \equiv r_\ell + \frac{\delta}{1-\delta}(r_h - r_\ell)$, this equilibrium implements bonus payments b_ℓ with an expected transfer to the principal r_P^{**} in the first period following acceptance, and induces immediate full revelation (through rejection). Otherwise, this equilibrium is a pooling equilibrium that implements bonus payments b_h with an expected transfer to the principal r_h in all periods.

Condition (4) is required for revelation to be feasible; it is the condition derived in Corollary 3 for the case of $\lambda = 0$. This condition says that revelation is feasible only if the future is discounted enough. Otherwise, the low type would not reveal herself by accepting an offer that the high type rejects. Behind this condition is the fact that only immediate revelation is relevant here: if separation of types cannot be induced immediately in the first period, then it cannot be induced in equilibrium.²⁹ To see this, suppose that immediate revelation is not feasible. Then an equilibrium where the low type accepts with probability one and the high type mixes between accepting and rejecting cannot exist, as the low type would have incentives to reject. So consider an equilibrium where the low type mixes between accepting and rejecting and the high type rejects with probability one. I use a Coase-conjecture-like argument (Coase, 1972) to show that such an equilibrium cannot exist either. If the principal types follow the proposed strategies, then there exists a posterior belief sufficiently low that, if the agent observes rejection, he optimally offers the symmetric-information contract of the high type, with an expected transfer $r_P = r_h$ to the principal, from then on. But then, when the agent's belief reaches that level, the low type rejects with probability one any offer that the high type rejects and, consequently, the agent optimally offers $r_P = r_h$ in that period. Continuing with this reasoning gives that the agent offers $r_P = r_h$ in all periods.

Condition (5) is required for revelation to be optimal; it parallels condition (3) for the case of $\lambda = 0$. This condition says that revelation is optimal if and only if the benefits outweigh the costs for the agent. The benefits come from the ability to pay a lower expected transfer to the principal and provide stronger incentives if the principal is revealed to be of the 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 of the high type. If the value of the relationship is likely to be high (p_0 is high) and the differences between the two types' outside options and the two types' symmetric-information surpluses are large ($r_h - r_\ell$ and $s_\ell - s_h$ are large), the benefits indeed exceed the costs and information revelation is optimal. Otherwise, the agent prefers to offer the pooling contract b_h and induce no revelation.

Behind condition (5) is the fact that only immediate revelation is relevant here too: if separation of types is optimal, then immediate separation in the first period is optimal. The reason that the agent cannot gain from postponing revelation of information is that, unlike the low type in the case of $\lambda = 1$, the agent now internalizes the externality. That is, the agent bears both the costs and benefits of inducing revelation, so his expected payoff falls if the net gain that he obtains from inducing revelation is postponed. The agent's ability to internalize the externality and the effects of this internalization on information revelation are further studied in the next subsection.

It is worth noticing that revelation may be unlikely in this setting. As mentioned in Corollary

²⁹Also behind this condition is the straightforward fact that partial revelation through contracts is not feasible in this setting.

3, there exists a tension between self-enforcement of relational contracts requiring a high discount factor and revelation through rejection requiring a low discount factor. The latter requirement is especially strong when the agent has all the bargaining power, as shown by condition (4).

5.3 When both parties have positive bargaining power

The last setting that I consider is one in which both the principal and the agent have positive bargaining power; that is, $\lambda \in (0, 1)$.

This setting is richer than those studied in the previous subsections. First, the relationship is valuable for both parties; hence, payments from the principal to the agent and from the agent to the principal can both be enforced. Second, in principle, full revelation can occur through default or through rejection, as the principal may have a decision to honor or renege as well as to accept or reject. Finally, in this setting, both types may want to misrepresent their types and indeed, as discussed below, they may want to do so simultaneously.

A full characterization of the Pareto-optimal equilibria is difficult because of the many off-the-equilibrium-path beliefs that must be considered. Given that both default and rejection may be used to reveal information, different belief specifications may result in one or the other being the equilibrium revelation channel. Moreover, both principal types may want to pretend to be of the other type, so the worst off-the-equilibrium-path punishment will in general not be the same for both types. As a result, considering extreme off-the-equilibrium-path beliefs will not be sufficient to identify the set of feasible equilibria, and a given belief specification will entail different deviation payoffs for the two types. Given that there is a continuum of off-the-equilibrium-path beliefs, this complication is not minor. Furthermore, the fact that the two types' incentives to misrepresent their types conflict with each other implies that focusing on Pareto-optimal equilibria will not simplify the problem much.

To make progress, I concentrate on the analysis of how information revelation changes when some bargaining power is shifted from one party to the other. This exercise is useful to understand what forces make revelation increase or decrease in a Pareto-optimal equilibrium, and how these forces are related to the distribution of bargaining power. For the cases of $\lambda = 1$ and $\lambda = 0$, the previous subsections have shown already that the optimal path of information revelation may be very different depending on the parties' relative bargaining positions. The analysis below shows that this is also true for the intermediate case of $\lambda \in (0, 1)$, and sheds light on the mechanisms behind this result.

To simplify the analysis and because, as discussed above, revelation through rejection may in general not be feasible, throughout this subsection I restrict attention to revelation through default. The following assumption implies that revelation cannot be induced through rejection in equilibrium.³⁰

³⁰This assumption is satisfied for both of the numerical examples given in Section 5.1. In Example A, $f_\ell = \frac{e_\ell}{2} = \frac{2}{5}$, so $\bar{\delta} \in [0.5, 0.68]$. In Example B, $f_\ell = \frac{e_\ell}{4} + \frac{1}{2} = \frac{-2+\sqrt{15}}{16}$, so $\bar{\delta} \in [0.5, 0.65]$. Thus, since $\delta = \frac{4}{5}$ in both cases, $\delta > \bar{\delta}$ for all $\lambda \in [0, 1]$ in both examples.

Assumption A5. Let $\bar{\delta}$ be defined as in Corollary 3. Then $\delta > \bar{\delta}$ for all $\lambda \in (0, 1)$.

Given this assumption, the bargaining power structure affects the amount of information that is revealed in equilibrium via two main effects: *Internalization of the Externality* (IE) and *Conflicting Incentives* (CI). IE is the fact that when the agent makes the offer, he can internalize (at least part of) the externality that the high type imposes on the low type. CI is the fact that for some parameter values, the high type wants to pretend to be the low type and the low type wants to pretend to be the high type, so their incentives conflict with each other.

IE and CI have opposing implications for how information revelation changes with the allocation of bargaining power. Through IE, revelation increases as λ falls. The idea is that the higher the agent's bargaining power, the more he can internalize the externality. To see this, consider the principal's participation (individual-rationality) constraints in a contract-pooling Markov equilibrium,

$$(1 - \delta)(r_P(p) - r_\ell) \geq \delta[\pi_\ell(\mu(p|\text{reject } b^A), b) - f(e)\pi_\ell(p_{b^A}, b) - (1 - f(e))\pi_\ell(p, b)], \quad (\text{IR}_\ell)$$

$$(1 - \delta)(r_P(p) - r_h) \geq \delta[\pi_h(\mu(p|\text{reject } b^A), b) - f(e)\pi_h(p_{b^A}, b) - (1 - f(e))\pi_h(p, b)], \quad (\text{IR}_h)$$

where $p_{b^A} \equiv \phi(\mu(p)|w^A + \bar{b}^A)$. In equilibrium, whenever the agent makes the offer, one of these participation constraints binds; otherwise, the agent could increase the fixed wage and increase his expected payoff. Also, it can be shown that if the agent induces revelation (through default) for some off-the-equilibrium-path belief $\mu(p|\text{reject } b^A)$, then he does so when his belief does not change upon observing rejection; namely, $\mu(p|\text{reject } b^A) = p$. Now, given the belief $\mu(p|\text{reject } b^A) = p$, it can further be shown that the high type's participation constraint always binds and the low type's becomes slack as λ falls. This implies that as the agent receives more bargaining power, he can capture the rent that the high type obtains when reneging, and hence internalize both the benefits and costs of inducing revelation. As a result, as λ falls, IE leads to more revelation: the agent induces revelation for a larger range of prior beliefs p_0 and at a faster rate. Furthermore, this also relaxes the condition for the principal to induce revelation herself.

Through CI, on the other hand, revelation falls as λ falls. The idea is that the higher the agent's bargaining power, the stronger the conflict of incentives. Indeed, since the low type has incentives to pretend to be the high type only when λ is relatively low (while the high type has incentives to pretend to be the low type for all $\lambda \in (0, 1)$), CI is relevant only when the agent has sufficiently high bargaining power. Now this implies that as λ falls, it becomes harder to sustain an equilibrium with revelation, be it an equilibrium where both principal types pool at a revelation contract or where both types pool at a no-revelation contract but the agent induces revelation when he proposes compensation. Intuitively, when revelation is induced and there is a conflict of incentives, beliefs that make deviations to off-the-equilibrium-path contracts unprofitable for some principal type inevitably make such deviations profitable for the other type, so it is not possible to deter both deviations simultaneously. As a result, as λ falls, CI leads to less revelation: the parties induce revelation for a smaller range of prior beliefs p_0 and at a slower rate.

Summarizing,

Proposition 11. *As λ falls, IE leads to more and faster revelation, while CI leads to less and slower revelation. There exists $\bar{\lambda} \in (0, 1)$ such that for $\lambda \geq \bar{\lambda}$, CI is non-binding and thus revelation is decreasing in λ .*

Of course, the specification of off-the-equilibrium-path beliefs influences the relative strength of IE and CI. In fact, depending on parameter values, there may exist beliefs such that one or the other effect can be completely “shut down.” For instance, if the discount factor is sufficiently high, there exist off-the-equilibrium-path beliefs such that the low type’s participation constraint always binds, and hence IE plays no role. These beliefs yield the tightest condition for revelation: an equilibrium with revelation exists if and only if the sum of the low type and the agent’s payoffs is higher with revelation than with no revelation. In contrast, if the discount factor is sufficiently low, there exist off-the-equilibrium-path beliefs such that deviations by both types can always be deterred in an equilibrium with revelation, and hence CI plays no role. These beliefs yield the most relaxed condition for revelation: for λ sufficiently low, an equilibrium with revelation exists if and only if the sum of the high type and the agent’s payoffs is higher with revelation than with no revelation. Formally,³¹

Lemma 4. *Let $e(p, b_\ell)$ solve the agent’s incentive compatibility constraint given bonus payments b_ℓ, \bar{b}_ℓ and $\chi(p) = p$. If*

$$(1 - \delta) \leq \delta f(e(p, b_\ell)), \quad (6)$$

then there exist beliefs such that IE is never binding. Given these beliefs, revelation is induced in equilibrium with $\mathbf{b}^ = \{\mathbf{b}^{P^*}, \mathbf{b}^{A^*}\}$ if and only if $s_\ell(p, \mathbf{b}^*) + (1 - \delta)(r_h - r_\ell) \geq s_h$ for $p = p_0, p_{0b^*}, \dots$*

Lemma 5. *If*

$$(1 - \delta)(s_h - r_\ell - r_A) \geq \delta(r_h - r_\ell), \quad (7)$$

then there exist beliefs such that CI is never binding. Given these beliefs and for λ sufficiently low, revelation is induced in equilibrium with $\mathbf{b}^ = \{\mathbf{b}^{P^*}, \mathbf{b}^{A^*}\}$ if and only if $s_h(p, \mathbf{b}^*) \geq s_h$ for $p = p_0, p_{0b^*}, \dots$*

Because of the conflict of incentives between the low and high types, the path of revelation may vary significantly across different Pareto-optimal equilibria. Indeed, unlike the settings where either the principal or the agent has all the bargaining power, in this setting, an equilibrium with no revelation may be Pareto optimal even if an equilibrium with revelation exists. That is,

Proposition 12. *If $\lambda \in (0, 1)$, Pareto-optimal equilibria with and without revelation may coexist.*

To give a simple example, suppose that the conditions given in Lemmas 4 and 5 both hold and λ is relatively low. Further, suppose that if revelation is induced, the sum of the low type

³¹Note that the conditions given in the lemmas are sufficient but not necessary. The condition given in Lemma 5 is satisfied for the numerical example A given in Section 5.1. Thus, in that case, CI is never binding.

and the agent's payoffs is lower than if no revelation is induced, but the sum of the high type and the agent's payoffs is higher. Then clearly, there exist belief specifications that sustain equilibria with and without revelation. Moreover, both classes of equilibria are Pareto optimal: the former maximizes the high type's expected payoff, the latter maximizes the low type's expected payoff.

Another interesting result that can be obtained from the analysis of this and the previous subsections concerns the parties' individual payoffs. It is obvious that varying the distribution of bargaining power will affect the distribution of surplus and thus the parties' expected payoffs. But more surprisingly, I find that the relationship between a party's expected payoff and his or her bargaining power is not necessarily always positive.

Proposition 13. *A party's expected payoff can sometimes be decreasing in his or her bargaining power. The agent may prefer having some bargaining power to having all the bargaining power; the principal may ex ante (before learning her type) prefer having no bargaining power to having some bargaining power.*

To illustrate, consider Example A from Section 5.1. In this example, $\delta = 4/5$, so revelation is not feasible for $\lambda = 0$ (see condition (4)). On the other hand, condition (7) is satisfied, and revelation is feasible for a range of prior distributions for all $\lambda \in (0, 1]$. But then, the agent is better off with λ close to zero than with $\lambda = 0$: giving some bargaining power to the principal allows the agent to induce revelation and thus increase his expected payoff. And, if the prior probability that the principal's type is low is relatively high, the principal is ex ante better off with $\lambda = 0$ than with λ close to zero: giving all the bargaining power to the agent allows the principal to ensure that no revelation is induced and thus increase her ex ante expected payoff.

This example is shown in Figure 5. The prior probability that the principal's type is low is assumed to be $p_0 = 0.8$. The figure depicts the parties' expected payoffs under different bargaining power distributions. The square and the triangle indicate the principal and agent's payoffs when the agent has all the bargaining power ($\lambda = 0$) and thus no revelation is induced. The solid and dashed lines indicate the principal and agent's minimum and maximum Pareto-optimal payoffs when the principal's bargaining power is positive ($\lambda > 0$) and as shown on the x -axis. In this case, the Pareto-optimal equilibrium always involves revelation of information.³² The figure then shows that the agent can always increase his expected payoff by reducing his bargaining power from $1 - \lambda = 1$ to $1 - \lambda \in [0.82, 1)$, and sometimes also by reducing it from $1 - \lambda = 1$ to $1 - \lambda \in [0.77, 1)$. Similarly, the principal can always increase her ex ante expected payoff by reducing her bargaining power from $\lambda \in (0, 0.03]$ to $\lambda = 0$, and sometimes also by reducing it from $\lambda \in (0, 0.08]$ to $\lambda = 0$.

³²The principal's minimum payoff and the agent's maximum payoff correspond to a Pareto-optimal equilibrium with $\mu(p|\text{reject } b^A) = p$, where the agent induces immediate revelation and the principal no revelation. The agent's minimum payoff and the principal's maximum payoff correspond to a Pareto-optimal equilibrium with $\mu(p|\text{reject } b^A) = 0$, where both parties induce delayed revelation, for a range of λ , and to a Pareto-optimal equilibrium with $\mu(p|\text{reject } b^A) = 1$, where the agent induces immediate revelation and the principal delayed revelation, for another range of λ .

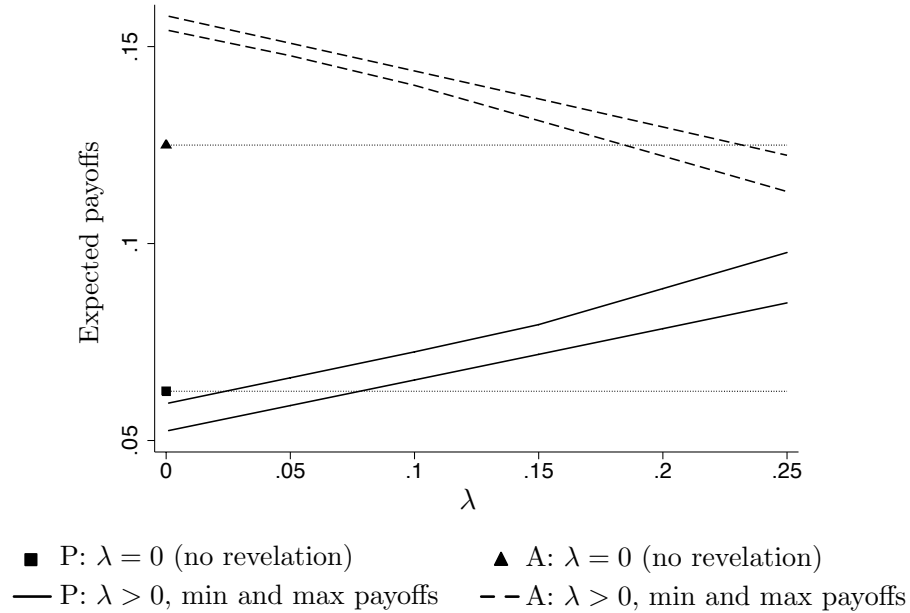


Figure 5: Individual payoffs (Example A)

6 Extensions

This section considers some extensions and variations of the model. I discuss how the results change if the agent's, rather than the principal's, outside option is not commonly known, if the principal's type varies over time, and if there is a continuum of principal types or output levels. I consider each of these issues separately, taking the model setup and assumptions as otherwise unchanged.

Asymmetric information about the agent's outside option. Suppose that the principal's outside option is fixed and known, and equal to r_P , while the agent's outside option is her private information, and equal to r_θ for $\theta \in \{\ell, h\}$. The main difference with the model studied above comes from the fact 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, then a given bonus scheme will provide weaker effort incentives to the high type than to the low type, as the high type reneges with positive probability when output is low. As a result, in equilibrium, the principal will update her belief about the agent's type not only upon observing the agent's payment decision, but also upon observing the realized level of output.

Still, although the analysis becomes more complicated, the main insights of the paper stay the same. If the agent has strong bargaining power, the high type wants to pretend to be the low type to increase the fixed wage (by promising a large bonus payment $-\underline{b}$ to the principal) and then renege and walk away with a high payoff when output is low. If the principal has strong bargaining power, the low type wants to pretend to be the high type to signal a high outside payoff and then receive a high payment when the principal proposes compensation. The requirements for revelation of information to be optimal are similar to those found above, and revelation always involves default

or rejection.

Time-varying outside option. The principal’s opportunities outside the relationship may vary over time. Consider first the extreme case where the principal draws her type at the beginning of each period t , and these draws are uncorrelated over time; that is, $\theta_t \in \{\ell, h\}$, $\text{corr}(\theta_t, \theta_{t+s}) = 0$ for $s \neq 0$. Then the expected future value of the relationship is the same regardless of the principal’s current type, and thus the scope of relational incentive provision under symmetric information is independent of θ_t . This simplifies the problem significantly. Given any allocation of bargaining power, optimal incentives solve program (P1) for $r_\theta = p_0 r_\ell + (1 - p_0) r_h$, and default never occurs on the equilibrium path. If $\lambda = 1$, the asymmetry of information about the principal’s type is indeed irrelevant. The asymmetry of information can lead to inefficiencies, however, when the agent proposes compensation. In such a case, if (and only if) p_0 is sufficiently high, the agent optimally proposes a high fixed wage that induces the high type to reject, so with probability $(1 - p_0)$ trade does not occur.

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 paper. (One could model this intermediate case, for example, by assuming that the principal’s type follows a first-order Markov process, with $1 > \text{Prob}(\theta_{t+1} = \ell | \theta_t = \ell) > \text{Prob}(\theta_{t+1} = \ell | \theta_t = h) > 0$.) Now the expected future value of the relationship does depend on the principal’s current type, but only stochastically. Hence, relative to the case of constant types, the agent’s learning will be slower and always incomplete, and the benefits of inducing revelation of information will be lower. More generally, the less serially correlated the principal types are, the less often revelation will be induced in a Pareto-optimal equilibrium.³³

Continuum of types. The model can be extended to a continuum of types, $\theta \in [\underline{\theta}, \bar{\theta}]$. Suppose that for all these types the value of the relationship is such that, under symmetric information, trade is profitable under some relational contract. Although the analysis becomes more cumbersome, the main insights remain unchanged relative to the two-type case. Full revelation of information cannot occur through contracts, and revelation must involve default or rejection. As a result, now revelation is always partial; otherwise, a contract with a zero probability of being honored or accepted should be implemented. This implies that the inefficiency persists over time. A characterization of revelation in equilibrium can be performed by using the density of principal types that honor the payments or accept the agent’s offer. Since revelation is never fully induced, this density is always strictly positive. 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 posterior is updated. Revelation may also be delayed, with a relatively high density of types honoring in the early periods. Lastly, if revelation is induced through rejection, it is induced all at once.

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

³³Athey and Bagwell (2008) study how the persistence of private information affects equilibrium outcomes in a dynamic Bertrand game.

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 . Intuitively, given risk neutrality, it is optimal to provide the strongest possible incentives, which come from specifying the maximum and minimum bonus payments that satisfy the self-enforcement constraint. For the case of asymmetric information, one can show that optimal contracts are also one-step. If no revelation is induced in equilibrium, or if revelation is induced only through rejection, the result follows immediately, as the role of bonus payments is exclusively to provide incentives. If revelation is induced through default, on the other hand, bonus payments are used not only to provide incentives but also to induce the high type to renege. However, all that matters is the total expected probability with which the high type honors a given contract, and this expected probability is independent of how many steps are defined in the contract.

7 Consequences of default, rejection, and unexpected offers

In this section, I discuss the restrictions on strategies introduced above and how the outcomes of the model change if they are relaxed. The appendix provides formal statements and proofs for the results discussed here.

Throughout the paper, I have assumed that following default the parties end the relationship with positive probability, and continue on the Pareto-optimal frontier otherwise (assumption A2). This assumption is without loss of generality—if 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, since then, as explained above, specifying the worst punishment for default is optimal. Now suppose that a default occurs in equilibrium. Then the continuation play after default must involve no trade with positive probability; otherwise, the two principal types would always receive the same payoffs and thus have the same incentives to renege, so a default would not occur in equilibrium. But this implies that assuming termination with positive probability following default is without loss. Further, suppose there exists an equilibrium where, if a default occurs, the relationship ends with probability $1 - \gamma$ and continues on an inefficient path of play with probability γ . Then it is not hard to show that, for some γ' , there exists an equilibrium where, if a default occurs, the relationship ends with probability $1 - \gamma'$ and continues on an efficient path of play with probability γ' , and where the parties' expected payoffs are the same. (See Proposition B1 in Appendix B.)

The other main restriction introduced in the paper is that, if no default occurs, the relationship always remains on the Pareto-optimal frontier (assumption A3). This means that the parties cannot terminate the relationship or switch to an inefficient level of trade with any positive probability upon observing a rejection or an unexpected offer. This assumption is not without loss; as I describe below, there exist Pareto-optimal equilibria with expected payoffs to the parties that cannot be replicated under this assumption. On the other hand, as already noted, 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. (See Proposition B2 in Appendix B.)

The motivation for assuming Pareto-optimal play while no default has occurred is to study situations where the 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 formally punished by law. The latter are no more than part of the strategic bargaining; neither a rejection nor an unexpected proposal would constitute a reason for legal action. How effectively parties can punish these deviations may also be related to market conditions. For example, in a lender-borrower setting (with informal credit contracts), borrowers may not be able to get a loan after they have 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 of the paper change when this restriction is relaxed, consider first allowing the parties to impose inefficient punishments following a rejection. (For now, continue assuming that they cannot do so following an unexpected offer.) For the same reasons as in the paper, full revelation through contracts cannot occur in equilibrium. However, since the agent may now threaten to walk away upon observing rejection, he can always induce revelation through rejection when making the offer, regardless of the discount factor. This has two main implications. First, if the bargaining power structure is such that both revelation through default and through rejection are feasible, I find that the agent always prefers inducing revelation through rejection. Second, and as a consequence of this, I show that the parties' individual expected payoffs are now everywhere increasing in their bargaining powers. (See Proposition B3 in Appendix B for these two results.) This result does not hold under the assumptions of the paper because, for a relatively high discount factor, revelation through rejection is then not possible, so revelation can be induced only through default and hence only if the principal has some bargaining power. Therefore, a conclusion from this analysis is that the extent to which the parties may want to give up some bargaining power depends on the extent to which their ability to credibly punish a deviation varies with the nature of the deviation.

Next, suppose that the parties can also impose inefficient punishments following an unexpected offer. The main implication for the results is that now a contract-separating equilibrium exists, since the agent can threaten to walk away if a contract that gives him the expected future value of the relationship with the low type, and leaves the principal with an expected payoff equal to the low type's outside option, is not offered. In such a case, the low type would be indifferent between this and any other contract, while the high type would strictly prefer to offer a contract that leads to termination of the relationship. (See Proposition B4 in Appendix B.) Indeed, this equilibrium exists for any bargaining power distribution with $\lambda > 0$. But then, one may argue, this prediction is not very "reasonable:" the agent captures all the surplus from the relationship even if the principal has more, or even all, the bargaining power. Realistically, one may think that the

principal would “refuse to play this equilibrium;” no matter what her type is, the principal is better off in an equilibrium that induces no revelation. More generally, when the 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, note that, independently of the restrictions on strategies, the result that informational asymmetries about the value of the relationship always lead to inefficiencies remains true. This follows from the fact that, while trade is profitable with both principal types, separation of principal types requires that trade cease with positive probability (or for some period of time) on the equilibrium path, be it after a default, a rejection, or an unexpected offer. (See Proposition B5 and Corollary B1 in Appendix B.) On the other hand, note that this result would not be true in a setting where the value of the relationship for the high type is low enough that no trade with this type is possible. Under such an assumption, which is the one adopted by the literature on relationship building (see Section 2), an equilibrium that achieves the symmetric-information expected surplus exists whenever revelation through rejection or through contracts is feasible.

8 Concluding remarks

In informal contractual relationships, parties may not perfectly know how much the other party values the relationship. This paper studies how this asymmetry of information affects contract design in a general agency setting. I develop a theoretical framework that can address this issue and provide a characterization of the form of the optimal contract. Here I conclude by going back to the employment example described in the introduction and discussing some empirical implications of the results. The model may also offer predictions for informal credit contracts, inter-firm agreements, supply-chain relationships, and other applications where contracting tends to be informal and information is incomplete.

The paper begins by showing that, under certain assumptions, the form of the contract can neither be used for signaling nor for screening. Intuitively, this says that firms cannot credibly convey their intentions through the payments or rewards they promise to give the workers. This result suggests, as a consequence, that a firm’s promised payments will not be a strong predictor of the probability with which the firm honors or reneges on these payments *ex post*. Additionally, the contract-pooling result tells us that in negotiations over worker pay, a union will not find it convenient to offer a menu of contracts to the firm, and will rather propose a single, take-it-or-leave-it compensation package.

The paper shows that information may still be revealed in equilibrium, and offers predictions on how relationships will induce and react to information revelation. The analysis suggests that, unlike in models where the value of the relationship is commonly known, informal relationships may follow non-stationary paths, may experience events of default and rejection, and may sometimes be dissolved. The results can help understand the evolution of incentives, the length of employment relationships, the probability of breakup, and the speed with which relationships may converge to stationary paths in different environments.

A main point of the paper is that the path of a relationship will depend on the allocation of bargaining power. In an industry where the union has all the bargaining power, I find that union negotiations may instantly lead to a strike or a permanent increase in wages. This occurs when the union can induce revelation through rejection. However, the model suggests that this will generally be unlikely, and that the employment relationship will follow a stationary path characterized by relatively low incentives and a low level of surplus. In this case, an increase in the firms' bargaining power, due for instance to an increase in the supply of workers from a declining industry or overseas, could promote efficiency. This shift in bargaining power would allow for discretionary payments from the firm to the union, and thus allow the parties to induce revelation through default. That is, we would observe an increase in incentives which, while making the relationship more susceptible to dissolution, would allow the parties to potentially move up to a new, more efficient stationary path. Paradoxically, I find that these events may increase wages and be more beneficial for the union than for the firm. On the other hand, the results show that strengthening the bargaining position of firms further may delay or even impede revelation of information, thus causing efficiency losses. In terms of policy implications, the analysis suggests a potential role for measures that shift bargaining power from one party to the other, such as regulations that affect firms' ability to hire replacement workers during a strike, or regulations on severance packages. Designing an optimal policy, though, appears to be a delicate task.

The model also stresses how the degree of asymmetric information and the prior distribution of types impacts the relationship. The analysis suggests that in industries where firms' prospects are uncertain, and not very likely to be good relative to alternative opportunities, revelation of information will not be induced, so the probability of breakup will be low and the path of relationships stationary. Naturally, relationships will also be unlikely to be dissolved in industries where firms are large, have significant industry-specific assets, and their value of the relationship is then easy to assess. It is in new, fast-growing industries where firms' prospects are unsure but likely to be good where the model predicts a higher probability of breakup. In these industries, following a phase of typically increasing incentives, relationships will either converge to stationary paths with a high level of surplus or be dissolved. How fast these events occur will depend on the perceived probability that firms have a high value of the relationship. When this probability is relatively high, relationships will become stationary, or be dissolved, relatively fast. When this probability is relatively low, parties will tend to postpone these events, implementing contracts that minimize the probability of conflict in the early stages of the relationship. Interestingly, the results then show that an improvement in an industry's outlook, or the release of favorable information about a firm, may precipitate events of default that would not have occurred or would have been delayed otherwise. As for policy implications, the model suggests that stricter information disclosure requirements may increase efficiency, although it also makes clear why firms—both firms with low as well as high relationship values—may be against such requirements.

Appendix A

This appendix provides proofs for the results stated formally in the paper.

Proof of Lemma 1. If the agent's belief is $p = 1$, under A1, h 's optimal strategy is to renege on payments \bar{b}_ℓ when the principal makes the offer and output is \bar{y} , and to reject when the agent makes the offer. Then,

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

Hence,

$$\begin{aligned}\pi_h(1) - \pi_\ell(1) &= \frac{\lambda f_\ell \delta (\pi_\ell(1) - r_\ell + r_h - \pi_\ell(1)) + (1 - \lambda)(1 - \delta)(r_h - r_\ell)}{1 - \lambda(1 - f_\ell)\delta - (1 - \lambda)\delta} \\ &= \left(1 - \frac{\lambda(1 - \delta)}{1 - \delta(1 - \lambda f_\ell)}\right) (r_h - r_\ell).\end{aligned}$$

□

Proof of Lemma 2. Suppose that b' is implemented and induces full revelation through default. Then since $r_h > r_\ell$, it must be that h reneges and ℓ honors (otherwise, if ℓ reneges, h would also want to renege). Then, by A3, after output is $y = \bar{y}$, b_h is implemented from then on if a default occurs and the relationship continues (with probability γ'), and b_ℓ if no default occurs. But then, if (E_ℓ) does not hold, ℓ prefers to renege rather than to honor, and if (D_h) does not hold, h prefers to honor rather than to renege. □

Proof of Lemma 3. Suppose that (b', r'_p) is implemented and induces full revelation through rejection. Then since $r_h > r_\ell$, it must be that h rejects and ℓ accepts. Then, by A3, after this period, b_h is implemented from then on if a rejection occurs, and b_ℓ if no rejection occurs. But then, if the first condition given in the lemma does not hold, ℓ prefers to reject rather than to accept, and if the second condition does not hold, h prefers to accept rather than to reject. □

Proof of Proposition 2. 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) 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 always honor, information is never revealed.

Next, I show that any PPBE with no revelation implements b_h . First, note that this PPBE cannot implement $b' < b_h$; regardless of beliefs (given A3), both types can increase their expected payoffs by deviating to b_h . Second, any PPBE that implements $b'' > b_h$ induces revelation. Suppose not. Then either both types always renege or both always honor. But the former cannot occur in a PPBE, and since $b'' > b_h$, h optimally reneges in equilibrium. □

Proof of Proposition 3. 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 ℓ with probability one; if $y = \bar{y}$ and the principal honors, $b^*(p)$ 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 $y = \underline{y}$, $b^*(p)$ 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_ℓ) . 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 ℓ would want to deviate to b_h regardless of beliefs (given A3). And by the definition of $b^*(p)$, if (3) does not hold for any k_h with $k_h(p) \in [0, 1)$ and γ , then there is no contract-pooling WMPBE that induces revelation. (See the text for a discussion of program (P2) which defines $b^*(p)$.)

Finally, I show that if a contract-pooling WMPBE with revelation does not exist, a contract-pooling PPBE with revelation does not exist. I prove the contrapositive. I proceed in three steps.

Step 1: Suppose there exists a PPBE with revelation with $\tilde{\mu}_{\tilde{t}}(p) \neq \tilde{\mu}_{\tilde{t}+1}(p)$ for some \tilde{t} . (Of course, the beliefs $\tilde{\mu}_{\tilde{t}}$, $\tilde{\mu}_{\tilde{t}+1}$ would differ only for off-the-equilibrium-path contracts. More precisely, the claim considers a PPBE with $\tilde{\mu}_{\tilde{t}}(p|\tilde{b}) \neq \tilde{\mu}_{\tilde{t}+1}(p|\tilde{b})$ where either $g_\ell(\tilde{\mu}_{\tilde{t}}, \tilde{b}) = g_h(\tilde{\mu}_{\tilde{t}}, \tilde{b}) = 0$, or $g_\ell(\tilde{\mu}_{\tilde{t}+1}, \tilde{b}) = g_h(\tilde{\mu}_{\tilde{t}+1}, \tilde{b}) = 0$, 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 $\tilde{k}_{h\tilde{t}}(\mu(p), b) \neq \tilde{k}_{h\tilde{t}+1}(\mu(p), b)$ or $\tilde{\gamma}_{\tilde{t}}(\mu(p), b) \neq \tilde{\gamma}_{\tilde{t}+1}(\mu(p), b)$ for some \tilde{t} . But by the same reasoning as above, then there exists a PPBE with revelation with $\tilde{k}_{ht}(\mu(p), b) = \tilde{k}_{h\tilde{t}+1}(\mu(p), b)$ and $\tilde{\gamma}_t(\mu(p), b) = \tilde{\gamma}_{\tilde{t}+1}(\mu(p), b)$ for $t = \tilde{t}, \tilde{t} + 1$ and, thus, given steps 1 and 2, a WMPBE with revelation exists. \square

Proof of Proposition 4. Contract-semi-separation is feasible only if h offers some b' with probability one and reneges when $y = \bar{y}$ with some positive probability, and ℓ mixes between b' and some $b'' < b'$ and always honors both contracts. But if ℓ offers b' with probability $g'_\ell \in (0, 1)$, $\mu(p|b') = pg'_\ell/[pg'_\ell + (1 - p)] < p$, and thus, given k_h , $\chi(\mu, b') < \chi(p, b')$ and $e(\mu, b') < e(p, b')$. So if there exists a PPBE where both types offer b' given a belief $\mu(p|b')$, there exists a PPBE where both types offer b' given a belief p . \square

Proof of Proposition 5. By Proposition 3, if (3) holds, a contract-pooling PPBE with revelation exists, which also implies that both types prefer this PPBE to the pooling one. Suppose that a contract-semi-separating PPBE implementing b' , $b'' < b'$ exists. Then by Proposition 4, a contract-pooling PPBE implementing b' also exists. But then, using the proof of Proposition 4, it is straightforward to see that there exists a contract-pooling PPBE that Pareto-dominates the contract-semiseparating PPBE. Finally, by Corollary 4, if (3) does not hold, the pooling PPBE is unique, and thus Pareto optimal. \square

Proof of Proposition 6. Suppose first that, whenever full revelation can be induced, bonus \bar{b}_ℓ can be implemented. Then postponing full revelation does not allow to implement a higher bonus, and it requires h to be indifferent between reneging and honoring in the present. Thus, given

$\delta < 1$, postponing revelation reduces h 's expected payoff. Hence, n^{\min} is the length of the shortest sequence \mathbf{k}_h for which revelation is feasible, and the results follow from Proposition 3. For p_0 sufficiently low, (3) implies $\pi_\ell(p_0, b) < \pi_h(0)$ for all $b \leq b_\ell$ if $k_h(p_0) = 0$, and since $\chi(p) \rightarrow k_h(p)$ as $p_0 \rightarrow 0$, it must be that $k_h(p) = 0$ whenever $b(p) > b_h$. This gives $K = \emptyset$ for $p_0 < \underline{p}$. For p_0 sufficiently high, $\pi_\ell(p_0, b) \geq \pi_h(0)$ for some $b \leq b_\ell$ and $k_h(p_0) = 0$. This gives $n^{\min} = 1$ for p_0 high, say $p_0 \geq p'$. Lastly, it remains to be shown that $\underline{p} < p'$ (instead of $\underline{p} = p'$). For this, take $p'' = p' - \varepsilon$ for some small $\varepsilon > 0$. By construction, $\pi_\ell(p'', b) < \pi_h(0)$ for all $b \leq b_\ell$ if $k_h(p'') = 0$, but $\pi_\ell(p', b') \geq \pi_h(0)$ for some b' and $k_h(p') = 0$. Then let b'' satisfy (I_h) for $\gamma(p'') = 0$ and $p_b = p'$: $\bar{b}'' = \frac{\delta}{1-\delta}[\pi_h(p', b') - r_h]$. Note that $\pi_h(p', b') > \pi_h(0)$, and hence $\bar{b}'' > \bar{b}_h$. Let $k_h'' \in (0, 1)$ be such that for $k_h(p'') = k_h''$, $\phi(p'' | w'' + \bar{b}'') \equiv p_b = p'$. As $\varepsilon \rightarrow 0$, $k_h'' \rightarrow 1$. But then for ε sufficiently small and $k_h(p'') = k_h''$, $\pi_\ell(p'', b'') \geq \pi_h(0)$.

Next, to show the results for n^{\max} , take for simplicity some p_0 such that $n^{\min} = 1$ (the reasoning is analogous for $n^{\min} > 1$). From above, h 's payoff is highest when $n_{\mathbf{k}} = 1$. However, ℓ 's payoff may be highest when $n_{\mathbf{k}} > 1$. To see this, consider two contract-pooling WMPBE implementing \mathbf{b}' and \mathbf{b}'' , with corresponding sequences \mathbf{k}'_h and \mathbf{k}''_h . Suppose $k'_h(p_0) = 0$, $k''_h(p_0) > 0$, $k''_h(p_{0b'}) = 0$. Then $\chi'(p_0) = p_0 + (1-p_0)k'_h(p_0) = p_0$ and $b'(p_0)$ satisfies (D_h) , (E_ℓ) , while $\chi''(p_0) = p_0 + (1-p_0)k''_h(p_0) > p_0$ and $b''(p_0)$ satisfies (I_h) , (E_ℓ) . Increasing $n_{\mathbf{k}}$ then increases $\chi(p_0)$, from p_0 to $\chi''(p_0)$, but reduces $\bar{b}(p_0)$, from $\bar{b}'(p_0)$ to $\bar{b}''(p_0)$, to leave h indifferent between renegeing and honoring. For p_0 close to one, the former effect is small and thus $\pi_\ell(p_0, b') > \pi_\ell(p_0, b'')$. For p_0 sufficiently small, the increase in $\chi(p_0)$ is large enough that $\pi_\ell(p_0, b') < \pi_\ell(p_0, b'')$. Of course, $n^{\max} < \infty$, as increasing $n_{\mathbf{k}}$ further requires a further decline \bar{b} , eventually leading to a decline in $\pi_\ell(p_0, b)$. For $p_0 = \underline{p}$, $n^{\max} = n^{\min}$. Otherwise, if $n^{\max} > n^{\min}$, there would exist some small $\varepsilon > 0$ such that a Pareto-optimal WMPBE with $n_{\mathbf{k}} = n^{\max}$ would be feasible for $\underline{p} - \varepsilon$. But this contradicts the definition of \underline{p} .

Finally, suppose that for some parameter values, full revelation can be induced only if the bonus payment is not too high. That is, given some p and $k_h(p) = 0$, $\pi_\ell(p, b') \geq \pi_h(0)$ for some $\bar{b}'(p) < \bar{b}_\ell$, but $\pi_\ell(p, b'') < \pi_h(0)$ for $\bar{b}''(p) > \bar{b}'(p)$. This may occur if $\pi_\ell(p, b)$ is decreasing in $\bar{b}(p)$ for low values of $\chi(p)$ (see Proposition 9). In this case, postponing full revelation allows to implement a higher bonus when full revelation is induced, as it allows to increase p , and thus $\chi(p)$, prior to inducing full revelation. As a result, n^{\min} may be greater than the length of the shortest sequence \mathbf{k}_h for which revelation is feasible, since h may prefer to postpone full revelation to implement a higher bonus, and then it follows that ℓ must also prefer to postpone full revelation. However, since $\pi_\ell(p, b)$ is increasing in $\bar{b}(p)$ for $\chi(p)$ sufficiently high, and the highest bonus that can be implemented when full revelation is induced is then increasing in p , the comparative statics of n^{\min} and n^{\max} with respect to p_0 are as above. \square

Proof of Proposition 7. That h 's preferred WMPBE has $n_{\mathbf{k}} = n^{\min}$ and ℓ 's $n_{\mathbf{k}} = n^{\max}$ follows from Proposition 6. To show the probationary contract result, suppose parameter values are such that in the equilibrium with $n_{\mathbf{k}} = n^{\min}$, bonus $\bar{b}(p) = \bar{b}_\ell$ can be implemented when full revelation is induced. This is true unless output is relatively insensitive to incentives and the agent's posterior belief when full revelation is induced relatively low (see Proposition 9). Consider for simplicity the contract-pooling WMPBE implementing \mathbf{b}' and \mathbf{b}'' as in Proposition 6, with corresponding sequences \mathbf{k}'_h and \mathbf{k}''_h . As above, assume $n_{\mathbf{k}'} = 1 = n^{\min}$, and suppose $n_{\mathbf{k}''} = 2 = n^{\max}$. Note that p_0 is such that $\pi_\ell(p_0, b'') > \pi_\ell(p_0, b') \geq \pi_h(0)$. But then, if $k''_h(p_0) < \bar{k}_h$, increasing $k_h(p_0)$ to \bar{k}_h must increase ℓ 's expected payoff, as the benefit from increasing $k_h(p_0)$ is constant in $k_h(p_0)$ while the cost is decreasing in $k_h(p_0)$. This follows from $\partial^2 \chi(p_0) / \partial (k_h(p_0))^2 = 0$, $\partial^2 \chi(p_{0b}) / \partial (k_h(p_0))^2 > 0$, and the fact that, by assumption, setting $k''_h(p_0) < \bar{k}_h$ does not allow for further increases in $\bar{b}''(p_{0b})$. Thus, ℓ 's preferred WMPBE has $k''_h(p_0) = \bar{k}_h$. More generally, if $n_{\mathbf{k}'} = n^{\min}$, $n_{\mathbf{k}''} = n^{\max} > n^{\min}$, and $\eta \equiv n^{\max} - n^{\min}$, ℓ 's preferred WMPBE has $k_h(p) = \bar{k}_h$ for the first η entries of \mathbf{k}''_h .

If parameter values are such that \bar{b}_ℓ cannot be implemented when inducing full revelation in the equilibrium with $n_{\mathbf{k}} = n^{\min}$, ℓ may not want to implement a probationary contract. The reason is that letting $k_h(p) < \bar{k}_h$ during the delay phase would allow to implement a higher bonus when full revelation is induced and, given h 's indifference condition, to implement a higher bonus in the previous periods. However, for this to be optimal for ℓ , it must be that $k_h(p)$ cannot be lowered in later periods instead. \square

Proof of Proposition 8. Take $\mathbf{k}_h \in K$ and corresponding γ, \mathbf{b} . Consider first the claim on $k(p)$. If $n_{\mathbf{k}} \leq 2$, there is nothing to prove. Then take $n_{\mathbf{k}} > 2$. For concreteness, consider p_0 and $p_{0b} > p_0$. Suppose first that ℓ 's expected payoff is increasing in the bonus payments. Then if $n_{\mathbf{k}} > n^{\min}$, $k_h(p_0) \geq k_h(p_{0b})$ follows from the reasoning in Proposition 7 (which shows that increasing $k_h(p_0)$ as much as possible is optimal). And if $n_{\mathbf{k}} = n^{\min}$, $k_h(p_0) \geq k_h(p_{0b})$ follows by construction. (If inducing revelation once the belief is updated to p_{0b} requires $k_h(p_0) < k_h(p_{0b})$ for this belief to be high enough, then revelation given the prior p_0 cannot be induced.) Next, suppose that ℓ 's expected payoff is decreasing in the bonus payments when the agent's belief is low. Then lowering $k(p_0)$ may allow to implement a higher bonus payment given p_{0b} , and in turn a higher bonus given p_0 . But if $k_h(p_0)$ needs to be reduced below $k_h(p_{0b})$ to increase $\bar{b}(p_{0b})$, it follows that ℓ will not be willing to induce revelation given p_0 . Thus, $k_h(p_0) \geq k_h(p_{0b})$ in equilibrium. Finally, the claim on $b(p)$ and $\gamma(p)$ follows from condition (I_h) . \square

Proof of Proposition 9. To show that $\gamma(p) = 0$ for $p \geq \hat{p}$ or $k_h(p) \geq \hat{k}_h(p)$, note that if $p \rightarrow 1$ or $k_h(p) \rightarrow 1$, $\chi(p) \rightarrow 1$, so the principal's expected payoff is always increasing in the bonus payments (for both types) and the worst punishment for default is optimal. To show that $\gamma(p) > 0$ may be optimal for $p < \hat{p}$, $k_h(p) < \hat{k}_h(p)$, note that if $\chi(p)$ is relatively low and the expected output relatively insensitive to incentives, then lowering the bonus payments increases $\pi_\ell(p, b)$. Thus, increasing $\gamma(p)$ above zero increases $\pi_\ell(p, b)$. Finally, for the last part, note that h 's continuation payoff after output is high is always larger than ℓ 's. Thus, for this continuation payoff fixed and given some fixed $n_{\mathbf{k}}$, $\partial\pi_h(p, b)/\partial\bar{b}(p) > \partial\pi_\ell(p, b)/\partial\bar{b}(p)$, as a higher $\bar{b}(p)$ increases the probability that output is high. Further, from conditions (I_h) , (D_h) , (E_ℓ) in program (P2), when $\gamma(p)$ falls (and thus $\bar{b}(p)$ increases), h 's continuation payoff after output is high either does not change or increases (see footnote 26). As a result, given $n_{\mathbf{k}}$, h prefers a lower $\gamma(p)$ than ℓ . \square

Proof of Proposition 10. Let an offer be denoted by the proposed bonus payment and expected transfer to the principal, $(\underline{b}(p), r_P(p))$. I proceed in three steps. First, I show that the equilibrium implements $(\underline{b}_\ell, r_P^{**})$ if (4) and (5) hold. Second, I show that it implements (\underline{b}_h, r_h) if (4) does not hold. Third, I show that it implements (\underline{b}_h, r_h) if (5) does not hold.

Step 1: Suppose that (4) and (5) hold. If the agent offers $(\underline{b}_\ell, r_P^{**})$, (4) implies that ℓ accepts and h rejects (see Lemma 3). Condition (5) implies $u(p_0, \underline{b}_\ell | r_P^{**}) \geq u(0)$, and, because by definition \underline{b}_ℓ maximizes enforceable incentives and r_P^{**} minimizes r_P , $u(p_0, \underline{b}_\ell | r_P^{**}) \geq u(p_0, \underline{b}' | r'_P)$ for any contract (\underline{b}', r'_P) that induces immediate revelation. Finally, condition (5) implies $u(p_0, \underline{b}_\ell | r_P^{**}) \geq u(p_0, \underline{b}' | r'_P)$ for any contract (\underline{b}', r'_P) 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 probability $a'_h \rightarrow 1$ in the current period. But then, if $u(p_0, \underline{b}_\ell | r_P^{**}) = u(0)$, $\underline{b}' = -(\delta/(1-\delta))(u(p_0, \underline{b}_\ell | r_P^{**}) - r_A) = \underline{b}_h$, and thus $u(p, \underline{b}') \rightarrow u(0)$. And if $u(p_0, \underline{b}_\ell | r_P^{**}) > u(0)$, $\underline{b}' < \underline{b}_h$, but these higher incentives must result in a less than proportional increase in the expected surplus, as otherwise \underline{b}_h would not be optimal. Thus, $u(p, \underline{b}') < u(p_0, \underline{b}_\ell | r_P^{**})$.

Step 2: Suppose that (4) does not hold. Then immediate revelation is not feasible, since ℓ rejects whenever h rejects. Gradual revelation is not feasible either. First note that a PPBE where h mixes between accepting and rejecting and ℓ accepts with probability one cannot exist, as again

ℓ rejects whenever h rejects. Consider then a PPBE where ℓ mixes between accepting and rejecting and h rejects with probability one. Given a prior p at time t , the agent offers $(\underline{b}(p), r_P(p))$. If the principal accepts, the agent learns that $\theta = \ell$ and $(\underline{b}_\ell, r_\ell)$ is implemented from then on. If the principal rejects, the agent updates his beliefs. Then there exists some p' low enough that the agent optimally offers (\underline{b}_h, r_h) from then on. But then the principal anticipates this, and when $p = p''$ such that $\mu(p''|\text{reject}) = p'$, if the agent offers $r_P(p'') < r_h$, ℓ rejects with probability one. So, when $p = p''$, the agent optimally offers (\underline{b}_h, r_h) . But then, when $p = p'''$ such that $\mu(p'''|\text{reject}) = p''$, ℓ again rejects any $r_P(p''') < r_h$ with probability one. Continuing with this reasoning gives that revelation is not feasible if immediate revelation is not feasible. (An equilibrium where both types mix between accepting and rejecting cannot exist either as it would require the agent to offer $r_P = r_h$ while information has not been fully revealed, and would not induce full revelation in finite time.) Thus, if (4) does not hold, the agent offers (\underline{b}_h, r_h) .

Step 3: Suppose that (5) does not hold. Then the agent is better off by offering (\underline{b}_h, r_h) than $(\underline{b}_\ell, r_P^{**})$ or any other contract that induces immediate revelation. And by the discussion in Step 1, it is straightforward that he is also better off by offering (\underline{b}_h, r_h) than any contract (\underline{b}', r'_P) that induces gradual revelation.

Finally, note that the proof can be rewritten with strategies as a function of time. The results are independent of the Markov assumption. \square

Preliminaries for the proofs of the results of Section 5.3. I analyze the conditions for revelation under assumption A5. For simplicity, I assume that the relationship ends with probability one whenever a default occurs. The proofs can be extended to allow for probabilistic termination.

First note that a PPBE with revelation exists if and only if a contract-pooling WMPBE with revelation exists. The proof is similar to that for the case of $\lambda = 1$ and thus omitted. For $i \in \{A, P\}$, $k_h^i(p) \in [0, 1]$, $\chi(p|b^i) \equiv \chi^i = p + (1-p)k_h^i(p)$, define the constraints

$$e^i(p) \in \arg \max_e f(e) \left[\chi^i \bar{b}^i - \underline{b}^i + \frac{\delta}{1-\delta} (\chi^i (u(p_{b^i}, b) - r_A) - (u(p, b) - r_A)) \right] - c(e), \quad (\text{IC}_A)$$

$$\bar{b}^i(p) = \frac{\delta}{1-\delta} [\pi_h(p_{b^i}, b) - r_h] \quad \text{if } k_h^i(p) > 0, \quad (\text{I}_h)$$

$$\bar{b}^i(p) \geq \frac{\delta}{1-\delta} (\pi_h(1) - r_h) \quad \text{if } k_h^i(p) = 0, \quad (\text{D}_h)$$

$$\bar{b}^i(p) \leq \bar{b}_\ell \quad (\text{E}_\ell)$$

$$\underline{b}^i(p) = -\frac{\delta}{1-\delta} [u(p, b) - r_A]. \quad (\text{E}_A)$$

If the agent induces revelation, he chooses \mathbf{b}^{A^*} to maximize $u^A(p, b)$ subject to $k_h^A(p) \in [0, 1]$, $\chi(p|b^A) = p + (1-p)k_h^A(p)$, (IC_A) , (I_h) , (D_h) , (E_ℓ) , (E_A) , the principal's participation constraints (IR_ℓ) , (IR_h) (in the text), the condition that $\mathbf{b}^{A^*}(p')$ be optimal for $p' = p_b, p_{bb}, \dots$, and the principal's strategy \mathbf{b}^{P^*} . Using the one-shot deviation principle, the agent induces revelation if and only if

$$u^A(p, \mathbf{b}^{A^*}, \mathbf{b}^{P^*}) \geq (1-\delta)(s_h - r_h) + \delta u(p, \mathbf{b}^{A^*}, \mathbf{b}^{P^*}). \quad (8)$$

Either (IR_h) or (IR_ℓ) must bind when $i = A$; otherwise, the agent can increase $w(p)$ and increase $u^A(p, \mathbf{b}^{A^*}, \mathbf{b}^{P^*})$. Also, the higher $\mu(p|\text{reject})$, the relatively more binding (IR_h) and less binding (IR_ℓ) . If $\mu(p|\text{reject}) = p$, (IR_h) binds and (IR_ℓ) may bind or not depending on \mathbf{b}^{P^*} and parameters; if $\mu(p|\text{reject}) = 0$, either (IR_ℓ) or (IR_h) or both may bind; and if $\mu(p|\text{reject}) = 1$, only (IR_h) binds.

If type θ induces revelation in a contract-pooling equilibrium, she chooses \mathbf{b}^{P^*} to maximize

$\pi_\theta^P(p, b)$ subject to $k_h^P(p) \in [0, 1]$, $\chi(p|b^P) = p + (1-p)k_h^P(p)$, (IC_A), (I_h), (D_h), (E_ℓ), (E_A), the agent's participation constraint (omitted), the condition that $b^{P^*}(p')$ be optimal for $p' = p_b, p_{bb}, \dots$, the agent's beliefs $\mu(p|b)$, and the agent's strategy \mathbf{b}^{A^*} . Given A3, θ induces revelation by offering $b^{P^*}(p)$ if and only if

$$\pi_\theta^P(p, b^{A^*}, b^{P^*}) \geq (1-\delta)(s_h - r_A) + \delta\pi_\theta(\mu(p|b_h), b^A, b^P), \quad (9)$$

$$\pi_\theta^P(p, b^{A^*}, b^{P^*}) \geq \pi_\theta^P(\mu(p|b^{P'}), b^A, b^{P'}) \quad \text{for all } b^{P'}(p) \in (b_h, b_\ell]. \quad (10)$$

The following result will be used below.

Result 1: A PPBE that induces revelation with $\mathbf{b} = \{b^{A^}, b^{P^*}\}$ exists only if for $p = p_0, p_{0b}, \dots$,*

$$\pi_h(p, b^{A^*}, b^{P^*}) \geq \pi_h(0) \quad (11)$$

$$u(p, b^{A^*}, b^{P^*}) \geq u(0) \quad (12)$$

Proof. By A3, regardless of beliefs, type h can always obtain $\pi_h(0)$ by offering b_h when the principal makes the offer and rejecting when the agent makes the offer, and the agent can always obtain $u(0)$ by offering b_h when he makes the offer and rejecting when the principal makes the offer. \square

Proof of Proposition 11. Step 1 (IE): To show that IE relaxes the agent's condition for revelation, note that the agent induces revelation if and only if he does so when $\mu(p|\text{reject}) = p$. This follows from (8), the observations made above about the IR constraints, and Result 1. (Since (IR_h) always binds for $\mu(p|\text{reject}) = p$, increasing $\mu(p|\text{reject})$ above p tightens this constraint; and given (11)-(12), if (8) holds for $\mu(p|\text{reject}) < p$, then it holds for $\mu(p|\text{reject}) = p$.) Next, note that given $\mu(p|\text{reject}) = p$, (8) is equivalent to

$$s_h^A(p, b^{A^*}, b^{P^*}) - \delta s_h(p, b^{A^*}, b^{P^*}) \geq (1-\delta)s_h. \quad (13)$$

As λ falls, (13) is relaxed, since (IR_ℓ) becomes less binding and eventually non-binding. Further, the contract that solves the agent's problem becomes steeper as (IR_ℓ) becomes less binding.

To show that IE relaxes the principal's condition for revelation, given the argument above, let $\mu(p|\text{reject}) = p$. Consider condition (9). Suppose first $\mu(p|b_h) = p$. Then, for $\theta = \ell$ and $\theta = h$ respectively, (9) is equivalent to

$$s_\ell^P(p, b^{A^*}, b^{P^*}) - \delta s_\ell(p, b^{A^*}, b^{P^*}) \geq (1-\delta)s_h, \quad (14)$$

$$s_h^P(p, b^{A^*}, b^{P^*}) - \delta s_h(p, b^{A^*}, b^{P^*}) \geq (1-\delta)s_h. \quad (15)$$

Note that for $\Delta s(p, b) \equiv s_h(p, b) - s_\ell(p, b)$,

$$s_h^P(p, b^A, b^P) - s_\ell^P(p, b^A, b^P) = \delta[f(e^P)\Delta s(p_{b^P}, b^A, b^P) + (1-f(e^P))\Delta s(p, b^A, b^P)],$$

and $\Delta s(p_{b^P}, b) \geq \Delta s(p, b)$, so (14) implies (15). Thus, in this case, it suffices to show that (14) is relaxed as λ falls. First take b^A fixed. As (IR_h) always binds, $b^A \geq b^P$. Then it is possible to show that the LHS of (14) (weakly) increases as λ falls. Second, note that b^A becomes steeper as λ falls (since (IR_ℓ) becomes less binding), also making the LHS of (14) increase. Suppose next that the principal's conditions for revelation are easiest to satisfy with some $\mu(p|b_h) = \mu' \neq p$. Then it must be that (9) with $\mu(p|b_h) = \mu'$ binds for $\theta = \ell$ if it binds for $\theta = h$. But then, as above, it is possible to show that IE relaxes (9) as λ falls.

Consider next condition (10). A contract-pooling PPBE implementing $b^{P^*} \in (b_h, b_\ell)$ requires $\mu(p|\tilde{b}^P) < p$ for $\tilde{b}^P > b^{P^*}$ to prevent deviations by h to \tilde{b}^P . But then, as shown in Step 2, ℓ

may want to deviate to \tilde{b}^P . IE makes this less likely by reducing ℓ 's benefit from signaling a high type: the stronger IE, the smaller the effects of a decline in the posterior belief on the amount of information the agent reveals.

Step 2 (CI): By the discussion in the text, if (i) an increase in the posterior increases $\pi_h(p, b)$ and decreases $\pi_\ell(p, b)$, and the difference between these two changes is sufficiently large, and (ii) the change in beliefs necessary to deter deviations by some type ($|\mu(p|b' \neq b^*) - p|$) is also sufficiently large, then a PPBE with revelation cannot be sustained. Now note that as λ falls, ignoring the effects of IE, it becomes more likely that both (i) and (ii) hold. For (i), note that while $\pi_h(p, b)$ is increasing in p for all $\lambda \in (0, 1)$, $\pi_\ell(p, b)$ becomes decreasing in p (and more strongly so) as λ falls. For (ii), note that as λ falls, even though \bar{b} falls (and $-\underline{b}$ increases), $\pi_h(p, b)$ also becomes less sensitive to changes in p . Thus, for an equilibrium pooling contract b^{P^*} , $\tilde{b}^P > b^{P^*}$, $\tilde{\mu}$ such that $\pi_h(p, b^A, b^{P^*}) = \pi_h(\tilde{\mu}, b^A, \tilde{b}^P)$, and λ close to zero, $p - \tilde{\mu}$ does not go to zero. An analogous analysis applies to $\tilde{b}^P < b^{P^*}$.

Even if the principal induces no revelation, CI may reduce revelation by the agent. Consider a PPBE with $b^{P^*} = b_h$, $b^{A^*} > b_h$. Such a PPBE requires $\mu(p|\tilde{b}^P) \equiv \tilde{\mu} < p$ for $\tilde{b}^P > b^{P^*}$ (otherwise, h would deviate to \tilde{b}^P). But then, if ℓ deviates to \tilde{b}^P , the agent updates his belief to $\tilde{\mu}$ and either strictly prefers not inducing revelation, or is indifferent between inducing and not inducing revelation. Further, if the agent induces revelation, then since $f(e(\tilde{\mu})) < f(e(p))$, revelation is now less likely to occur and the agent must offer $r_P(\tilde{\mu}) > r_P(p)$. Thus, ℓ may want to deviate to \tilde{b}^P .

Finally, to show formally that for λ high enough CI does not bind, note that if

$$(1 - \delta)(s_h - r_\ell - r_A) \geq \delta(1 - \lambda)(r_h - r_\ell), \quad (16)$$

then off-the-equilibrium-path beliefs $\mu(p|b^P < b_h) = p$, $\mu(p|b^P > b_h) = 0$ deter deviations by ℓ and h when $b^P = b_h$, $b^A > b_h$, and $r_P = r_\ell$. But then it follows immediately that CI is non-binding. Lastly, note that (16) holds for λ sufficiently high. \square

Proof of Lemma 4. If (6) holds, then for $\mu(p|\text{reject}) = 0$, (IR_ℓ) binds for all $\lambda \in (0, 1)$ and thus IE never binds. The agent's condition for revelation (8) when $\mu(p|\text{reject}) = 0$ is then

$$s_\ell^A(p, b^{A^*}, b^{P^*}) + (1 - \delta)(r_h - r_\ell) - \delta(u(p, b^{A^*}, b^{P^*}) + \pi_h(0)) \geq (1 - \delta)s_h$$

Given Result 1 and the fact that the sum of ℓ and the agent's per-period (flow) payoffs when $i = P$ is at least s_h , the condition given in the Lemma follows. (The principal's condition for revelation is always more stringent.) \square

Proof of Lemma 5. If (7) holds, then (16) holds for all $\lambda \in (0, 1)$. Then by the argument above, CI never binds. Further, for λ low enough and $\mu(p|\text{reject}) = p$, (IR_ℓ) is slack and the agent induces revelation if and only if (13) holds. The condition given in the Lemma follows. (The principal's condition for revelation is always more stringent.) \square

Proof of Proposition 12. The example in the text proves the proposition. \square

Proof of Proposition 13. The example in the text proves the proposition. \square

Appendix B

This appendix states and proves the results discussed in Section 7.

Proposition B1. *If a Pareto-optimal equilibrium exists, there exists a Pareto-optimal equilibrium that satisfies assumption A2 and gives the same expected payoffs to all the parties.*

Proof. Suppose that no default occurs in equilibrium. Then the worst punishment for default is optimal and terminating the relationship with probability one is without loss.

Suppose next that a default occurs in equilibrium. Then after default, ℓ and h 's continuation payoffs must be different; otherwise, for any given contract, both types would want to honor or both to renege, but this contradicts the assumption that a default occurs in equilibrium. Then, since ℓ and h only differ in their outside options, it must be that a default is followed by a contract that involves no trade with strictly positive probability with some type. And since $r_h > r_\ell$, such contract must involve no trade with h . But then note that ℓ never reneges in equilibrium, so assuming that the relationship ends with some positive probability after default is without loss.

Finally, suppose there exists a Pareto-optimal equilibrium where, after default, the relationship ends with probability $1 - \gamma > 0$ and continues on an inefficient path of play with probability γ . Consider a second equilibrium where, after default, the relationship ends with probability $1 - \gamma' > 0$ and continues on an efficient path of play with probability γ' . Let γ' be such that h 's continuation payoff after default is the same as in the first equilibrium. (It is straightforward to show that such γ' exists.) Then ℓ 's continuation payoff after default is lower than in the first equilibrium. Hence, the second equilibrium allows to implement the same or higher self-enforcing incentives as the first equilibrium while a default does not occur, and also to lower the punishment for default for h further conditional on ℓ 's enforcement constraint holding. The result follows. \square

Proposition B2. *If an equilibrium is Pareto optimal under assumptions A2-A4, then it is Pareto optimal when A2-A4 are not imposed.*

Proof. First note that A2 is without loss by Proposition B1. Next, note that any equilibrium under A3-A4 is also an equilibrium when A3-A4 are not imposed. Finally, suppose by contradiction that there exists a Pareto-optimal equilibrium under A3-A4 that is not Pareto optimal when A3-A4 are not imposed. Let the expected surplus generated by this equilibrium be s , and the expected payoffs u , π_ℓ , and π_h . Then it must be that, when A3-A4 are not imposed, there exists a Pareto-optimal equilibrium that (i) is not an equilibrium under A3-A4, (ii) generates expected surplus $s' > s$, and (iii) gives expected payoffs $u' \geq u$, $\pi'_\ell \geq \pi_\ell$, and $\pi'_h \geq \pi_h$. But then (i) and (ii) imply that such an equilibrium must induce separation of types by either (a) prescribing inefficient play following a rejection by the principal or (b) prescribing inefficient play following an unexpected offer by the principal. But in case (a), there exists an equilibrium where no inefficient play is prescribed after rejection and ℓ 's expected payoff is strictly higher than π'_ℓ (by (i), this is an equilibrium with no revelation), and in case (b), there exists an equilibrium where no inefficient play is prescribed after an unexpected offer and both ℓ and h 's expected payoffs are strictly higher than π'_ℓ and π'_h respectively (since $\lambda > 0$ in this case, both types' expected payoffs are strictly higher in an equilibrium with no revelation). But then, in any case, (iii) does not hold. Contradiction. \square

Proposition B3. *Suppose the parties may end the relationship with positive probability after a rejection. Then, (i) if an equilibrium where the agent induces revelation exists, there exists an equilibrium where the agent induces revelation through rejection, and the agent's expected payoff is higher in this equilibrium than in one in which he induces revelation through default; and (ii) the parties' expected payoffs are everywhere increasing in their bargaining powers.*

Proof. Since the agent can now walk away following a rejection by the principal, it is immediate that there always exists a contract that induces revelation through rejection when the agent makes the offer (for example, the agent offers a contract (b', r'_P) such that ℓ 's expected payoff following acceptance is equal to r_ℓ , and ends the relationship if the principal rejects). To prove (i), assume for simplicity that the relationship ends with probability one following a rejection or default, and normalize the agent's outside option to $r_A = 0$. If the agent induces revelation through rejection, given prior p_0 , his expected payoff is $p_0(s_\ell - r_\ell)$. If the agent induces revelation through default, his expected payoff is highest when $\lambda \rightarrow 0$, only (IR_h) binds, and revelation is induced immediately. Then, as $\lambda \rightarrow 0$, the agent's expected payoff goes to

$$\frac{(1 - \delta)(s(e) - r_h) + p_0 \delta f(e)(s_\ell - r_\ell)}{1 - \delta(1 - f(e))}$$

where $e \in \arg \max_{e'} f(e') p_0(\bar{b}_\ell - \underline{b}_\ell) - c(e')$. Then note that if $p_0(s_\ell - r_\ell) = s_h - r_h$, $p_0(\bar{b}_\ell - \underline{b}_\ell) = \bar{b}_h - \underline{b}_h$, and the agent is indifferent between inducing revelation through rejection, inducing revelation through default, and inducing no revelation. If $p_0(s_\ell - r_\ell) > s_h - r_h$, then $p_0(\bar{b}_\ell - \underline{b}_\ell) > \bar{b}_h - \underline{b}_h$, but the increase in the bonus payments must cause a less than proportional increase in $s(e)$, as otherwise b_h would not be optimal. Hence, the agent's expected payoff is higher when he induces revelation through rejection. Claim (ii) follows from the fact that the agent can induce revelation through rejection whenever he makes the offer (regardless of δ), and, by (i), he always prefers inducing revelation through rejection than through default. \square

Proposition B4. *Suppose the parties may end the relationship with positive probability after an unexpected offer. Then a contract-separating equilibrium exists for all $\lambda \in (0, 1]$.*

Proof. Let the agent's beliefs be $\mu(p_0|b_1, w_1) = 1$ for some contract (b_1, w_1) and $\mu(p_0|b_2, w_2) = 0$ for any contract $(b_2, w_2) \neq (b_1, w_1)$. Let $b_1 = b_\ell$ and w_1 be such that $\pi_\ell(p_0, b_1, w_1) = r_\ell$. Suppose that the agent ends the relationship with probability one if the principal offers (b_2, w_2) . Then it is immediate that ℓ is indifferent between (b_1, w_1) and (b_2, w_2) , while h strictly prefers (b_2, w_2) . The claim follows. \square

Proposition B5. *Regardless of the bargaining protocol, if $\theta \in \{\ell, h\}$ is privately observed by the principal, a separating or semi-separating equilibrium where trade occurs with probability one on the equilibrium path does not exist.*

Proof. Suppose by contradiction that trade occurs with probability one on the equilibrium path. Then ℓ and h 's expected payoffs are the same (recall that the two principal types differ only in their outside options). But then ℓ and h take the same actions in equilibrium, and no separation can occur. \square

Corollary B1. *Regardless of the bargaining protocol, if $\theta \in \{\ell, h\}$ is privately observed by the principal, a separating equilibrium without distortions does not exist.*

Proof. This result follows from Proposition B5. \square

References

- ABREU, D. (1988): “On the Theory of Infinitely Repeated Games with Discounting,” *Econometrica*, 56(6), 383–96.
- ADMATI, A., AND M. PERRY (1987): “Strategic Delay in Bargaining,” *Review of Economic Studies*, 54(3), 345–64.
- AGHION, P., AND P. BOLTON (1987): “Contracts as a Barrier to Entry,” *American Economic Review*, 77, 388–401.
- AGHION, P., AND B. HERMALIN (1990): “Legal Restrictions on Private Contracts Can Enhance Efficiency,” *Journal of Law, Economics, and Organization*, 6(2), 381–409.
- ATHEY, S., AND K. BAGWELL (2008): “Collusion with Persistent Cost Shocks,” *Econometrica*, 76(3), 493–540.
- AUSUBEL, L., AND R. DENECKERE (1989): “Reputation in Bargaining and Durable Goods Monopoly,” *Econometrica*, 57(3), 511–31.
- BAKER, G., R. GIBBONS, AND K. MURPHY (2002): “Relational Contracts and the Theory of the Firm,” *Quarterly Journal of Economics*, 117(1), 39–84.
- BARON, D., AND R. MYERSON (1982): “Regulating a Monopolist with Unknown Cost,” *Econometrica*, 50(4), 911–30.
- BÉNABOU, R., AND J. TIROLE (2003): “Intrinsic and Extrinsic Motivation,” *Review of Economic Studies*, 70, 489–520.
- BOWEN, R., L. DUCHARME, AND D. SHORES (1995): “Stakeholders’ Implicit Claims and Accounting Method Choice,” *Journal of Accounting and Economics*, 20, 255–295.
- BULL, C. (1987): “The Existence of Self-Enforcing Implicit Contracts,” *Quarterly Journal of Economics*, 102(1), 147–159.
- COASE, R. (1972): “Durability and Monopoly,” *Journal of Law and Economics*, 15(1), 143–49.
- DEANGELO, H., AND L. DEANGELO (1991): “Union Negotiations and Corporate Policy,” *Journal of Financial Economics*, 30, 3–43.
- ENDLICH, L. (1999): *Goldman Sachs: The Culture of Success*. New York: Alfred A. Knopf.
- FREIXAS, X., R. GUESNERIE, AND J. TIROLE (1985): “Planning under Incomplete Information and the Ratchet Effect,” *Review of Economic Studies*, 52, 173–91.
- FUCHS, W. (2007): “Contracting with Repeated Moral Hazard and Private Evaluations,” *American Economic Review*, 97(4), 1432–48.
- FUDENBERG, D., D. LEVINE, AND E. MASKIN (1994): “The Folk Theorem with Imperfect Public Information,” *Econometrica*, 62(5), 997–1039.
- FUDENBERG, D., D. LEVINE, AND J. TIROLE (1985): “Infinite-Horizon Models of Bargaining with One-sided Incomplete Information,” in *Game-Theoretic Models of Bargaining*, ed. by A. Roth. Cambridge: Cambridge University Press.

- FUDENBERG, D., AND J. TIROLE (1991): "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, 53(2), 236–60.
- GENICOT, G., AND D. RAY (2006): "Bargaining Power and Enforcement in Credit Markets," *Journal of Development Economics*, 79(2), 398–412.
- GHOSH, P., AND D. RAY (1996): "Cooperation in Community Interaction without Information Flows," *Review of Economic Studies*, 63, 491–519.
- GIBBONS, R. (2005): "Incentives Between Firms (and Within)," *Management Science*, 51(1), 2–17.
- GREEN, E., AND R. PORTER (1984): "Noncooperative Collusion under Imperfect Price Information," *Econometrica*, 52(1), 87–100.
- GROSSMAN, S., AND M. PERRY (1986): "Perfect Sequential Equilibrium," *Journal of Economic Theory*, 39, 97–119.
- HERMALIN, B. (2002): "Adverse Selection, Short-Term Contracting, and the Underprovision of On-the-Job Training," *Contributions to Economic Analysis and Policy*, 1(1), Article 5.
- HOLMSTRÖM, B. (1999): "The Firm as a Subeconomy," *Journal of Law, Economics, and Organization*, 15(1), 74–102.
- HUTCHENS, R. (1989): "Seniority, Wages and Productivity: A Turbulent Decade," *Journal of Economic Perspectives*, 3(4), 49–64.
- HYMOWITZ, C. (1990): "Kodak Gives 'Parachutes' to Workers," *Wall Street Journal*, January 12, B1.
- IPPOLITO, R., AND W. JAMES (1992): "LBOs, Reversions and Implicit Contracts," *Journal of Finance*, 47(1), 139–167.
- KLEIN, B., AND K. LEFFLER (1981): "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89(4), 615–641.
- LAFFONT, J.-J., AND J. TIROLE (1988): "The Dynamics of Incentive Contracts," *Econometrica*, 56(5), 1153–75.
- LAZEAR, E. (1979): "Why Is There Mandatory Retirement?," *Journal of Political Economy*, 87, 1261–84.
- LEAP, T. (1995): *Collective Bargaining and Labor Relations*. Englewood Cliffs, New Jersey: Prentice Hall.
- LEVIN, J. (2003): "Relational Incentive Contracts," *American Economic Review*, 93(3), 835–57.
- LEWIS, T., AND D. SAPPINGTON (1989): "Countervailing Incentives in Agency Problems," *Journal of Economic Theory*, 49, 294–313.
- MACLEOD, B. (2003): "Optimal Contracting with Subjective Evaluation," *American Economic Review*, 93(1), 216–40.
- MACLEOD, B., AND J. MALCOMSON (1989): "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 57(2), 447–80.

- MAILATH, G., AND L. SAMUELSON (2001): "Who Wants a Good Reputation?," *Review of Economic Studies*, 68, 415–41.
- ROGERSON, W. P. (1985): "The First-Order Approach to Principal-Agent Problems," *Econometrica*, 53(6), 1357–67.
- RUBINSTEIN, A. (1985): "A Bargaining Model with Incomplete Information About Time Preferences," *Econometrica*, 53(5), 1151–72.
- SCHMITZ, P. (2006): "Information Gathering, Transaction Costs, and the Property Rights Approach," *American Economic Review*, 96(1), 422–34.
- SHAPIRO, C., AND J. STIGLITZ (1984): "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*, 74(3), 433–444.
- SOBEL, J. (1985): "A Theory of Credibility," *Review of Economic Studies*, 52(4), 557–73.
- SPIER, K. (1992): "Incomplete Contracts and Signaling," *RAND Journal of Economics*, 23, 432–43.
- STEWART, J. (1993): "Taking the Dare," *The New Yorker*, July 26, 34–9.
- WATSON, J. (1999): "Starting Small and Renegotiation," *Journal of Economic Theory*, 85, 52–90.
- (2002): "Starting Small and Commitment," *Games and Economic Behavior*, 38, 176–99.
- WILLIAMSON, O. (2000): "The New Institutional Economics: Taking Stock, Looking Ahead," *Journal of Economic Literature*, 38(3), 595–613.