

# LEARNING IN RELATIONAL CONTRACTS

Rumen Kostadinov\*  
McMaster University

Aditya Kuvalekar†  
U Carlos III de Madrid

October, 2019

## Abstract

We study relational contracts between a firm and a worker with mutual uncertainty about match quality. The worker's actions are publicly observed and generate both output and information about the match quality. We show that the relational contracts may be inefficient. We characterize the inefficiency through a holdup problem on the contemporaneous output. In the frequent action limit, these inefficiencies persist if and only if information degrades at least at the same rate at which impatience vanishes. We characterize optimal relational contracts and show that they involve actions that yield both a lower payoff and less information than another action.

*Keywords:* Relational contracts, dynamic agency, career concerns

JEL codes: C73,D21,D23,D83,M55

---

\*E-mail: [kostadir@mcmaster.ca](mailto:kostadir@mcmaster.ca) and [akuvalek@eco.uc3m.es](mailto:akuvalek@eco.uc3m.es).

†We are especially grateful to Joyee Deb and Elliot Lipnowski for countless discussions and comments. We thank Debraj Ray, David Pearce and Ennio Stacchetti for valuable insights and constant encouragement. We thank Pierpaolo Battigalli, Dhruva Bhaskar, Rahul Deb, Matthias Fahn, William Fuchs, Daniel Garrett, Johannes Hörner, Aaron Kolb, Antoine Loeper, Chiara Margaria, Alessandro Pavan, João Ramos, Nishant Ravi, Evan Sadler, Maher Said and Nikhil Vellodi for useful comments. We would also like to thank the seminar audiences at Bocconi University, Boston University, Calgary University, Norwegian School of Economics, UC San Diego, University of Edinburgh and various conferences. Aditya Kuvalekar gratefully acknowledges support from the Ministerio Economía y Competitividad, Maria de Maeztu grant (MDM 2014-0431), and Comunidad de Madrid, MadEco-CM (S2015/HUM-3444).

# 1 Introduction

Two features are ubiquitous in employment relationships. First, there is mutual uncertainty about their match quality, a crucial determinant of productivity.<sup>1</sup> The second feature is the incompleteness of contracts. Contracts are rarely conditioned on all observable aspects of an employee's performance: some actions can be difficult to verify in an outside court, or prohibitively costly to contract on. However, a benefit of the dynamic nature of employment relationships is that the parties form relational contracts—informal agreements linking good performance to bonuses and promotion. The goal of this paper is to study the nature of relational contracts when the parties face uncertainty about their match quality.

For example, suppose Bob and his manager, Ann, perceive that their match quality is relatively poor. Bob is on the verge of getting fired unless he succeeds in his next project. If Bob works hard—a costly activity for him—he is more likely to succeed. If he shirks, indeed, being successful is less likely, but at least Bob would save his effort costs. Should Bob work hard in order to maximize his chances of staying employed or shirk, since the relationship is likely to end? Can Ann provide incentives for hard work?

We answer these questions by studying a standard relational contracting setting of perfect information between a Ann (the firm) and Bob (the worker) where the parties are symmetrically uninformed about their match quality. In this framework, Bob's work plays a dual role: it produces immediate revenue (output) for Ann. But his output and effort also produce information about their match quality. Our first contribution is to highlight the mechanism that leads to systematic distortions in the relational contracts between Ann and Bob. In short, if the immediate revenue is a strong reason why Ann would have Bob work hard at the brink of separation, the outcomes are inefficient. On the other hand, if the primary motive is to acquire information, then the outcomes are efficient. So what should the parties do when they cannot attain efficiency? Our second contribution answers this question by providing a full characterization of the optimal relational contract. Most notably, this characterization reveals that Ann may assign Bob to a less ambitious project which will not require hard work but will also reveal less information about the future prospects of the relationship. These relational contracts resemble probationary periods or “performance improvement plans (PIPs)” that we will expand upon later.

**Summary of the framework and results:** We study an infinite horizon, discrete time game between a firm and a worker. In each period, the firm can fire the worker or offer him

---

<sup>1</sup><https://hbr.org/2017/06/when-leaders-are-hired-for-talent-but-fired-for-not-fitting-in>

a salary. If the worker accepts the salary, he chooses an *observable* action. The output is stochastically determined by the worker’s action and a relationship-specific match quality. We study the set of all relational contracts, i.e., Perfect Bayesian Equilibria of this game.

Notice that an action produces both output and also information about the match quality. Hence, actions can be ranked according to their informativeness (in the Blackwell sense) and their profitability—expected output minus the effort cost of the action. For most of the paper we analyze an environment with two actions: high and low. We assume that the high action is more informative as well as more profitable than the low action.

We begin by studying the efficiency of relational contracts. The efficient benchmark is the solution to the experimentation problem of a planner who maximizes the joint surplus of the firm and the worker ignoring incentive constraints. It is characterized by a simple policy—terminate the relationship below a certain cutoff belief about the match quality and choose the high action above the cutoff.<sup>2</sup>

We obtain a full characterization of efficiency driven by the interaction between contract incompleteness and learning. Consider the efficient cutoff belief where the planner is indifferent between the outside options and experimentation. At this belief, the cost of the worker’s action is exactly offset by the value of experimentation over and above the outside options. The value of experimentation consists of two parts—contemporaneous output and the option value of learning. In relational contracts, incentives to choose the high action are provided via the option value of learning alone. Therefore, if the contemporaneous output is larger than the outside options, the option value of learning will not offset the costs of effort. As a result, relational contracts are inefficient. Otherwise, they are efficient. The way we interpret this, in our original Ann and Bob example, is that whenever the immediate revenue (contemporaneous output) is sufficiently attractive for Ann, relational contracts are inefficient. On the other hand, if the option value of learning is sufficiently large, relational contracts are efficient. One implication of this result is that inefficiency can be mitigated by increasing unemployment benefits which would strengthen the outside options.

We proceed to characterize the optimal relational contract in Proposition 3. The contract exhibits a three-region structure (Figure 1). The players take their outside options at low beliefs. The worker chooses the low action at intermediate beliefs and the high action at high beliefs. Despite its inferior profitability and informativeness, the low action offers a

---

<sup>2</sup>This is a correlated three-arm bandit problem, the three arms being the high action, the low action and the outside option. In general, such problems can be analytically challenging. However, since the high action unambiguously dominates the low action, obtaining the optimal policy is straightforward.

better alternative to terminating the relationship when the firm cannot credibly promise to reimburse the worker for the cost of the high action.

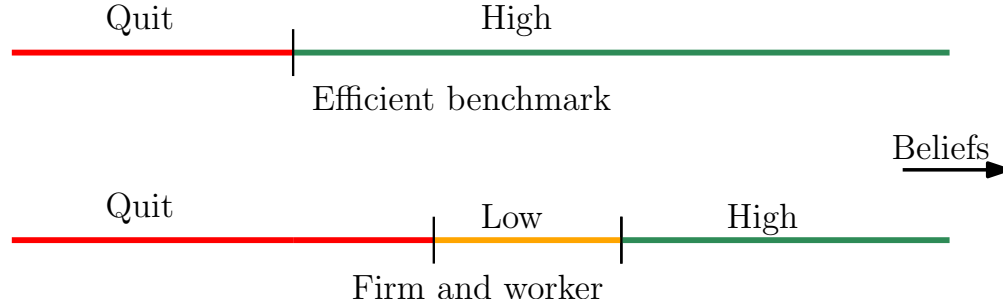


Figure 1: Difference between the efficient benchmark and optimal relational contract

As mentioned earlier, this characterization resembles the performance improvement plans (PIPs) that are often used in consulting firms.<sup>3</sup> In PIPs, struggling employees are assigned less important roles as a chance to redeem themselves and stay employed before getting fired. In the optimal relational contract, the low action region can be interpreted as a PIP. We would like to point out that, unlike other models of relationship building (Watson, 1999, 2002; Halac, 2014), an important implication of our model is that the worker may move in and out of a PIP multiple times. This is a testable implication of our model.

The intuition behind Proposition 3 is that the high action is used whenever it is incentive compatible, since it is more informative and more profitable. Hence, the low action is only optimal at lower beliefs where the high action cannot be incentivized. This intuition is a correct starting point but is incomplete. In the efficient benchmark, the low action is never used *because*, absent incentive constraints, more information is always desirable. In a relational contract incentive constraints may make information undesirable at some beliefs.<sup>4</sup> This in turn may render the low action superior *because* it is less informative.

The proof of Proposition 3 obtains the three-region structure by using the algorithmic procedure by Abreu, Pearce, and Stacchetti (1990), APS hereafter. We adapt the APS algorithm to show that iterative application of an operator on an appropriate initial function yields, as a fixed point, the equilibrium surplus function (maximum achievable surplus at

<sup>3</sup>See <https://www.shrm.org/resourcesandtools/tools-and-samples/how-to-guides/pages/performanceimprovementplan.aspx>

<sup>4</sup>Preference for information is equivalent to the convexity of the surplus function for the optimal relational contract. As we will see later, the surplus function in the optimal relational contract is generally not convex. In a special case where all actions are equally informative, such non-convexities do not pose any technical hurdles and a three-region structure is immediate. See Section 7.4 for details.

each belief). We demonstrate the use of this approach in two steps—(i) Applying the operator once on the function obtained in the efficient solution yields a three-region structure; and (ii) due to our assumptions on the ranking of payoff and informativeness, we obtain that the three-region structure is preserved upon every subsequent operation. Therefore, the three-region structure is obtained as a fixed point property. We are optimistic that this approach could be useful in other settings.

A natural point of inquiry for our model is the effect of increasing patience. There are typically two kinds of limits that are studied towards this goal. Folk theorems study the limit as players become arbitrarily patient, leaving the rest of the environment unchanged. Another source of patience is frequent interaction, i.e., reducing the length of each period. The literature originating in Abreu, Milgrom, and Pearce (1991) (AMP hereafter) studies the limiting outcomes of repeated games with imperfect monitoring as the length of a period converges to 0.<sup>5</sup> Shorter periods increase the effective patience of players, despite a fixed rate of time discounting. However, they have the additional effect of reducing information per period. The frequent action limit and the limit in standard folk theorems are equivalent in settings of perfect monitoring and complete information. In our setting—perfect monitoring with uncertainty about match quality—the two limits may no longer be equivalent. Therefore, we have a rare opportunity to explore both limits in an environment with perfect monitoring.

Being agnostic about which of the two limits is more natural, in Section 6, we study both approaches to the patient limit in a unified framework. We consider a sequence of instances of our model with players’ patience increasing along the sequence while the quality of information may degrade, both measured in per-period terms. We say that the learning problem is trivial if the limiting optimal policy associated with the efficient outcome involves experimentation at any positive belief. Theorem 1 provides a sharp characterization of the limiting outcome in the efficient benchmark and in the optimal relational contract. First, the holdup problem characterizing inefficiency from Proposition 2 persists in the limit if and only if the learning problem is non-trivial. Second, the triviality of the learning problem is characterized by the relationship between the rate at which informativeness, as measured by the variance of the posterior belief, deteriorates, and the rate at which players become perfectly patient along the sequence of models. Whenever information decays at a

---

<sup>5</sup>AMP consider complete information and imperfect monitoring, while our setting has incomplete information and perfect monitoring. This distinction gives rise to qualitative differences discussed in detail in Section 6.

slower rate than the rate at which impatience vanishes, the learning problem is trivial. Consequently, there is no holdup problem and relational contracts are efficient, consistent with a folk theorem.<sup>6</sup> When information decays at an equal or faster rate, the learning problem is nontrivial and the holdup problem persists in the limit. Hence, the rate of information decay determines whether inefficiency can be a robust limiting feature of our model.

The remainder of the paper proceeds as follows. Section 2 discusses the connection of our work with some existing literature. In Section 3, we describe our game and the equilibrium concept. Section 4 sets up the preliminary analysis leading to our main results presented in Section 5. In Section 6 we analyze the limit of our model as the parties become patient and information may deteriorate. Section 7 details some extensions of the model.

## 2 Related Literature

The early literature on relational contracts established their efficiency in various settings of complete and perfect information when the parties are sufficiently patient (Bull (1987), Thomas and Worrall (1988), MacLeod and Malcomson (1989)). This is no longer true in settings with private information such as moral hazard and adverse selection (Levin (2003), Halac (2012)). Our paper shows that incomplete yet *public* information alone can also create inefficiencies in relational contracts. Hörner and Samuelson (2013) obtain a similar type of inefficiency in the version of their model with observable actions, which is a special case of the class of frequent action models that we consider. In Hörner and Samuelson (2013) inefficiency always obtains due to specific parametric assumptions (outside options equal to zero). Our characterization of efficiency shows that this is not true in general and it identifies the force behind (in)efficiency.

Persistent uncertainty in relational contracts has also been studied by Thomas and Worrall (2010), Kwon (2016) and Malcomson (2016).<sup>7</sup> Recently, DeVaro et al. (2018) have studied relational contracts with moral hazard where a publicly known, exogenously evolving state affects the future value of the relationship but not the productivity of effort. They find that the effort is increasing in the state in the optimal relational contract, similar to our three-region characterization. Unlike these settings, the worker controls the flow of information through his effort in our environment.

---

<sup>6</sup>The closest existing folk theorem to our setting is due to Wiseman (2005) discussed in Section 6.

<sup>7</sup>Beyond relational contracts, in a repeated Bertrand game with persistent, private cost shocks, Athey and Bagwell (2008) show that high persistence relative to patience leads to inefficiencies.

Our conclusions share some similarities with the literature on relationship building that posits that favourable outcomes are achieved by establishing trust through repeated interactions. Typically, these dynamics are driven by asymmetric information (Ghosh and Ray, 1996; Watson, 1999, 2002; Halac, 2014). McAdams (2011) studies a stochastic partnership game with an exogenously evolving relationship capital that yields similar dynamics as our optimal relational contract. While we discuss the contrast in detail after Proposition 3, the evolution of relationship capital is endogenous in our setting.<sup>8</sup>

Symmetric uncertainty has traditionally been explored in the career concerns literature beginning with Holmström (1999). Typically, the agent is paid his marginal product every period, which leaves no room for dynamic incentive provision, e.g. backloaded compensation. Also, the career concerns literature focuses on the *signal jamming* incentives where the worker tries to influence the market's beliefs. When the agent's action is unobservable, market driven incentives can give rise to divergent beliefs and, subsequently the ratchet effect, recently demonstrated by Cisternas (2018). Observable action rules out both signal jamming and the ratchet effect. In a continuous time environment without transfers, Kuvalekar and Lipnowski (2016) study a firm-worker relationship with symmetric uncertainty about match quality and observable actions. They show that even with observable actions, a non-trivial agency problem ensues in a world without commitment. Besides the lack of transfers, the actions in their environment are costless and they focus on Markovian equilibria.

Symmetric uncertainty has also garnered attention in the dynamic contracting literature. With full commitment, the efficient outcome is trivially achieved in our setting. Hence, the dynamic contracting literature is often concerned with moral hazard. Prat and Jovanovic (2014) analyzed a continuous time model with Brownian learning, and found that in order to avoid the ratchet effect (as in Laffont and Tirole (1988) and Bhaskar (2014)) the contract induces the agent to shirk early on, and when the information about his ability is more precise, the rewards are more tightly linked to the output. These dynamics are similar to our findings—under limited commitment a relationship with perceived low match quality results in low effort and transitions to high effort as prospects improve. Due to the binary state of the match quality in our model, the signals' precision does not improve over time, and we can see switches between high and low effort due to streaks of luck, while Prat and Jovanovic (2014) obtained a permanent shift to high effort.

---

<sup>8</sup>McAdams (2011) also considers a general version where the relationship capital evolves endogenously. However, he does not obtain a three-region structure as in our Proposition 3.

Uncertainty about match quality naturally leads to an experimentation problem connecting us with Keller et al. (2005), Bolton and Harris (1999) who also have underexperimentation in the strategic interaction. However, in our setting there is no free-rider effect, which is the main driver of inefficiencies in that literature.

### 3 Model

We consider an infinitely repeated interaction between a firm (she) and a worker (he), both risk-neutral. The productivity of their relationship is determined by the quality of their match, which can be either good or bad. The parties are symmetrically uninformed: they begin with a common prior  $p_1 \in (0, 1)$  that the match quality is good. Time is discrete and denoted by  $t = 1, 2, \dots$ . The firm and the worker have outside options of  $\underline{v}$  and  $\underline{u}$  per period, respectively. Let  $\underline{s} = \underline{u} + \underline{v}$  be the sum of their outside options.

The timing within a period  $t$  is shown in Figure 2. At the beginning of the period the firm and the worker share a common belief  $p_t$  that the match quality is good. The firm decides whether to continue interacting with the worker ( $d_t^f = 1$ ) or to terminate the relationship by firing the worker ( $d_t^f = 0$ ). If she interacts, she offers a wage  $w_t$  to the worker. The worker decides whether to accept the offer ( $d_t^w = 1$ ) or terminate the relationship by quitting the firm ( $d_t^w = 0$ ). If the parties terminate the relationship, they receive their outside options in all remaining periods, starting from the current one.<sup>9</sup> If, instead, a wage offer is accepted, the wage is paid to the worker and he chooses a level of effort ( $e_t$ ), which can be High ( $H$ ) or Low ( $L$ ) and is observed by the firm.<sup>10</sup> Effort is not contractible, so the worker is free to choose any level of effort, regardless of the wage he accepted. The cost of  $H$  to the worker is  $c > 0$ , while the cost of  $L$  is 0.

Effort \ Match quality	Good	Bad
	High	$\gamma_H$
Low	$\gamma_L$	$\beta_L$

Table 1: Probability of a success given effort and match quality.

<sup>9</sup>Allowing the parties to suspend the relationship for a finite number of periods, instead of terminating it, would not affect the equilibrium outcome. This follows from the existence of an equilibrium where both parties receive their outside options in all periods.

<sup>10</sup>Allowing the firm to pay a bonus at the end of the period does not change the model, as the bonus can be factored into the next period's wage.



The outcome of the worker's effort is either a success (high output) ( $y_t = y^h$ ) or a failure (low output) ( $y_t = y^l < y^h$ ). The probability of success when the worker's effort is  $e$  is given by  $\gamma_e$  when the match quality is good and  $\beta_e$  when the match quality is bad, as shown in Table 1. Given a prior  $p$  and effort  $e \in \{H, L\}$ , let  $\pi_e(p)$  and  $R_e(p)$  denote, respectively, the expected probability of success and the expected output in the current period. Then

$$\begin{aligned}\pi_e(p) &= p\gamma_e + (1-p)\beta_e \\ R_e(p) &= \pi_e(p)y^h + (1-\pi_e(p))y^l\end{aligned}$$

Since all actions are publicly observable, the firm and the worker form a common posterior  $p_{t+1}$  at the end of period  $t$ . Let  $\phi_e^+(p)$  (resp.  $\phi_e^-(p)$ ) be the posterior belief following effort  $e$  and success (failure) when the prior at the beginning of the period was  $p$ . According to Bayes' rule,

$$\phi_e^+(p) = \frac{p\gamma_e}{p\gamma_e + (1-p)\beta_e} \quad \phi_e^-(p) = \frac{p(1-\gamma_e)}{p(1-\gamma_e) + (1-p)(1-\beta_e)}. \quad (1)$$

In addition, we allow for public randomization. Formally, at the beginning of each period the parties observe the realization  $x_t$  of a random variable uniformly distributed on  $[0, 1]$ . This is a common convenience in the literature, as it convexifies the equilibrium set of payoffs.

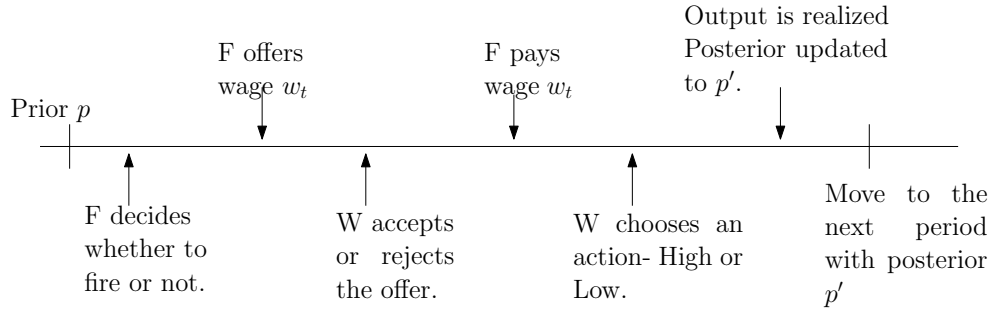


Figure 2: Timeline of one period

### 3.1 Assumptions

The following assumptions define our baseline model and will be maintained throughout Sections 4, 5, and 6. They are visualized in Figure 3

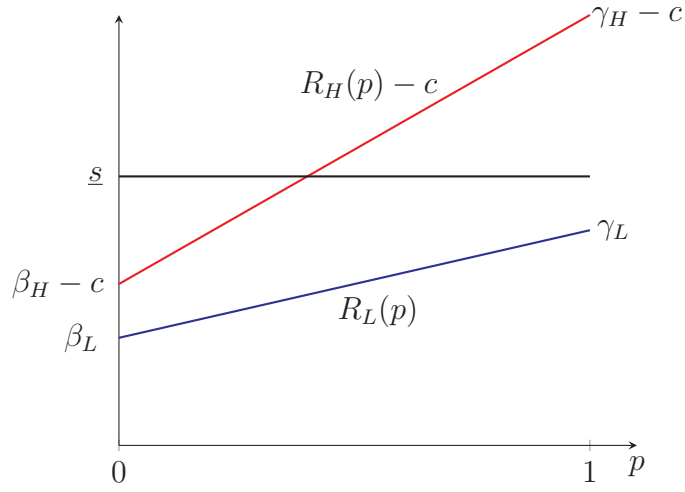


Figure 3: Myopic surplus as a function of beliefs.

**Assumption 1.**  $R_H(1) - c > \underline{s} > R_H(0) - c > R_L(0)$  and  $\underline{s} > R_L(1) > R_L(0)$ .

**Assumption 2.** If  $\beta_H = \beta_L = 0$ , then  $\gamma_H > \gamma_L$ . Otherwise,  $\frac{\gamma_H}{\beta_H} > \frac{\gamma_L}{\beta_L}$ .

The assumptions imply the following:

1.  $H$  generates more expected surplus (output net of cost of effort) than  $L$  regardless of the match quality.
2.  $H$  generates more (less) expected surplus than the outside options when the match quality is good (bad).
3.  $L$  generates less expected surplus than the outside options regardless of the match quality.
4.  $H$  is more informative than  $L$  in the sense of Blackwell (1953), which means that

$$\phi_H^+(p) \geq \phi_L^+(p) \geq p \geq \phi_L^-(p) \geq \phi_H^-(p) \quad \text{for all } p.$$

Finally, we assume that players are patient enough to sustain  $H$  in equilibrium when the match quality is good. Without this assumption the unique equilibrium would involve the players taking the outside option at every belief.

**Assumption 3.**  $\delta [R_H(1) - c - \underline{s}] \geq (1 - \delta)c$ .

## 3.2 Discussion of the model

In the baseline model we assume that the match quality, realized at the beginning of the relationship, is static. This need not be the case. The match quality can change when firms undergo technology shocks, changes in management, or market conditions. To model these scenarios we could assume that the match quality evolves exogenously according to some Markov process. Such evolution would not change the structure of our analysis and the results. This is because, even though the environment is changing, the players observe the same signals regarding its evolution. Therefore, they share the same beliefs regarding match quality at all points in time.

Our model can also accommodate a setting, wherein the worker is assigned projects by the firm. Suppose the firm can assign either a big or a small project to the worker each period. Big projects carry more responsibility, e.g. managing teams, dealing with clients etc. Small projects can be back-office jobs. The worker chooses his effort on the project. Big projects, due to greater responsibility, can succeed only if the worker exerts high effort. In that case, the probabilities of success are  $\gamma_H$  and  $\beta_H$  respectively depending on whether the match is good or bad. The probability of success for the small projects is  $\gamma_L$  and  $\beta_L$  depending on whether the match is good or bad *regardless of the choice of action*. In this framework, the analysis and the results remain identical to the baseline model.

## 3.3 Strategies and Equilibrium

We denote by  $h_t := \{x_t, w_t, e_t, y_t\}$  the history in stage  $t$  if the relationship has not been terminated. Let  $h^t := \{h_\tau\}_{\tau=1}^{t-1}$  be the history at the beginning of period  $t$  if the relationship has not been terminated. Let  $\mathcal{H}^t$  be the set of histories up to time  $t$ .

A strategy for the firm is a sequence of functions  $\{D_t^f, W_t\}_{t=1}^\infty$ , where  $D_t^f : \mathcal{H}^t \times [0, 1] \rightarrow \{0, 1\}$  and  $W_t : \mathcal{H}^t \times [0, 1] \rightarrow \mathbb{R}$  determine the interaction decision and the wage offer as a function of the history in previous stages and the public random variable in the current stage. Similarly, a strategy for the worker is a sequence of functions  $\{D_t^w, E_t\}_{t=1}^\infty$ , where  $D_t^w : \mathcal{H}^t \times [0, 1] \times \mathbb{R} \rightarrow \{0, 1\}$  and  $E_t : \mathcal{H}^t \times [0, 1] \times \mathbb{R} \rightarrow \{H, L\}$  determine his interaction decision and effort level as a function of the history in the previous stages, the current public random variable, and the firm's wage offer.

Let  $D_t := d_t^f d_t^w$ . Following history  $h^t$  and a posterior belief  $p_t$ , the respective continu-

ation payoffs to the firm and the worker from strategies  $(\sigma^P, \sigma^A)$  are given by

$$v_t(p_t|h^t, (\sigma^P, \sigma^A)) := \mathbb{E} \left[ (1 - \delta) \sum_{\tau=t}^{\infty} \delta^{\tau-t} \left( D_t(y_t - w_t) + (1 - D_t) \frac{v}{1 - \delta} \right) \right]$$

$$u_t(p_t|h^t, (\sigma^P, \sigma^A)) := \mathbb{E} \left[ (1 - \delta) \sum_{\tau=t}^{\infty} \delta^{\tau-t} \left( D_t(w_t - c\mathbb{1}_{e_t=H}) + (1 - D_t) \frac{u}{1 - \delta} \right) \right]$$

where the expectation is taken over nonterminal histories that are continuations of  $h^t$  given the belief  $p_t$  and the continuations of strategies  $(\sigma^P, \sigma^A)$ .<sup>11</sup>

Following the literature, we define a relational contract as a Perfect Bayesian Equilibrium (PBE) with public randomization of the dynamic game with the restriction that beliefs are updated recursively according to Bayes rule (1) after every period (henceforth referred to as equilibrium). This means that the public beliefs are updated only through the realization of the output given the worker's action *on and off path*. The motivation behind this is that in a PBE we can assign arbitrary beliefs off the equilibrium path. For example, if the worker chose the high action when he was supposed to choose the low action, the posterior belief can be arbitrary and unrelated to the output realization. This is obviously unsatisfactory given the structure of our game. Ideally, we want to embed the notion that players cannot signal what they do not know.<sup>12</sup> Our restriction achieves this purpose.

Let  $\mathcal{E}(p)$  denote the set of equilibrium payoffs of the game starting at prior belief  $p$ .

## 4 Equilibrium Characterization

### 4.1 Efficiency benchmark

The sole conflict in our environment is the separation between the entity exerting effort (worker) and the entity collecting the output (firm). Therefore, our natural notion of efficiency is one where effort is contractible. Alternatively, we could view this as an experimentation problem wherein a planner, possessing the same information about match quality as the firm and the worker, chooses effort to maximize the total surplus of the firm and the worker. This is a multi-armed bandit problem with two correlated arms ( $H$  and  $L$ ), and one safe arm given by the outside options.

---

<sup>11</sup>Note that if the relationship is terminated at time  $t$  ( $D_t = 0$ ), the game ends and no histories of length greater than  $t$  occur in equilibrium. The players receive  $\frac{u}{1-\delta}$  and  $\frac{v}{1-\delta}$  at time  $t$ , which is equivalent to receiving their outside options in each subsequent period.

<sup>12</sup>As Battigalli (1996) pointed out, this implies a notion of strategic independence.

Given a value function  $f$ , let  $T_e f$  denote the surplus from choosing an effort  $e \in \{H, L\}$  today and drawing continuations from  $f$  according to the probability distribution induced by  $e$ .

$$T_e f(p) := (1 - \delta)(R_e(p) - \mathbb{1}_{e=H}c) + \delta \mathbb{E}^e[f(p)]$$

where  $\mathbb{E}^e[f(p)] = \pi_e(p)f(\phi_e^+(p)) + (1 - \pi_e(p))f(\phi_e^-(p)) \quad \forall e \in \{H, L\}$

The optimal policy for this problem specifies an effort level or the outside options at each belief  $p$ . Let  $G(p)$  denote the payoff associated with the optimal policy. The value function  $G$  is the unique fixed point of the following operator  $T : B[0, 1] \rightarrow B[0, 1]$  where  $B[0, 1]$  denotes the set of bounded real-valued functions on  $[0, 1]$ :

$$Tf(p) = \max\{T_H f(p), T_L f(p), \underline{s}\}. \quad (2)$$

Proposition 1 uses standard dynamic programming arguments to establish that  $G$  is convex. Thus, the planner prefers more informative arms. Since  $H$  offers more information and more surplus than  $L$ , the planner never chooses the latter. Hence, the optimal policy has a simple structure. There is a cutoff belief  $p^{FB}$  above which the planner uses  $H$  and below which the planner chooses the outside options.

**Proposition 1.** *The efficient value function is given by the unique fixed point  $G$  of the operator  $T$ .  $G$  is increasing and convex. There exists a unique belief  $p^{FB} \in (0, 1)$  such that the optimal policy is to choose  $H$  above  $p^{FB}$  and the outside options below  $p^{FB}$ .*

All proofs are presented in the appendix.

## 4.2 Recursive Characterization

We proceed to analyse  $\mathcal{E}(p)$ , the set of equilibrium payoffs of the game starting with prior  $p \in [0, 1]$ . Techniques from Abreu, Pearce, and Stacchetti (1990) (APS) are easily adapted to show that equilibria admit a recursive structure with the belief acting as a state variable. Any equilibrium payoff  $(u, v) \in \mathcal{E}(p)$  can be decomposed into an action profile for the current period and continuation payoffs subject to incentive constraints. Continuation payoffs must belong to  $\mathcal{E}(p')$ , where  $p'$  is the updated belief. For each  $(u, v) \in \mathcal{E}(p)$ , the payoffs are decomposed in one of the following three ways.

- (High effort) There exist continuation payoffs  $(u^+, v^+) \in \mathcal{E}(\phi_H^+(p))$ ,  $(u^-, v^-) \in \mathcal{E}(\phi_H^-(p))$  and a wage  $w$  such that,

$$\begin{aligned}
v &= (1 - \delta)(R_H(p) - w) + \delta[\pi_H(p)v^+ + (1 - \pi_H(p))v^-] \\
u &= (1 - \delta)(w - c) + \delta[\pi_H(p)u^+ + (1 - \pi_H(p))u^-] \\
v &\geq \underline{v}, \quad u \geq \underline{u} \\
\delta [\pi_H(p)u^+ + (1 - \pi_H(p))u^- - \underline{u}] &\geq (1 - \delta)c
\end{aligned} \tag{3}$$

- (Low effort) There exist continuation payoffs  $(u^+, v^+) \in \mathcal{E}(\phi_L^+(p))$ ,  $(u^-, v^-) \in \mathcal{E}(\phi_L^-(p))$  and a wage  $w$  such that

$$\begin{aligned}
v &= (1 - \delta)(R_L(p) - w) + \delta[\pi_L(p)v^+ + (1 - \pi_L(p))v^-] \\
u &= (1 - \delta)w + \delta[\pi_L(p)u^+ + (1 - \pi_L(p))u^-] \\
v &\geq \underline{v} \quad u \geq \underline{u}
\end{aligned}$$

- (Termination) The firm does not interact or the worker rejects the offer.

$$v = \underline{v} \quad u = \underline{u}$$

The above characterization relies on the fact that players act sequentially and publicly, and that there is an equilibrium that gives both parties their outside options (see Lemma 2 in the appendix). This equilibrium is used to threaten deviators.

In light of this punishment, the incentive constraint for the firm is straightforward. Suppose the firm deviates by offering  $w'$ . If  $w' \geq \underline{u}$ , the worst continuation equilibrium for the firm is one where the worker accepts and exerts  $L$  and the outside options are taken in the following period. The resulting payoff for the firm is strictly less than  $\underline{v}$ . If  $w' < \underline{u}$ , then the worker quits giving the firm a payoff of  $\underline{v}$ . Therefore, the firm's incentives amount to a participation constraint.

For the worker, an additional constraint (3) is that he needs to be compensated through future continuation value if he is exerting  $H$ . If he deviates by choosing  $L$ , he still collects the wage, saves on the cost  $c$  and is punished by receiving  $\underline{u}$  tomorrow.

### 4.3 Optimal Relational Contracts through Surplus Maximization

Notice that the contemporaneous wages can be used to redistribute surplus without affecting incentives, as in Levin (2003). Hence, an optimal relational contract maximizes the sum of the firm's and the worker's payoffs and is characterized by a surplus function  $S$  defined below.

$$S(p) := \max_{(u,v) \in \mathcal{E}(p)} \{u + v\}$$

In a surplus-maximizing relational contract, it is without loss of generality to keep the firm to its outside option and give the worker the entire future surplus over and above the firm's outside option by shifting the current wage in the firm's favour. This offers the strongest incentives for the worker without impacting the incentives of the firm. This observation helps us characterize  $S$  as a fixed point of the operator  $T_*$  defined in Lemma 1 below.  $T_*$  resembles the operator  $T$  in (2) used to obtain the efficient surplus, but with the addition of an incentive constraint (4) for using  $H$ .

**Lemma 1.**  $S$  is a fixed point of the operator  $T_* : B[0, 1] \rightarrow B[0, 1]$  given by

$$T_* f(p) := \begin{cases} \max\{T_H f(p), T_L f(p), \underline{s}\} & \text{if } \delta(\mathbb{E}^H[f(p)] - \underline{s}) \geq (1 - \delta)c \\ \max\{T_L f(p), \underline{s}\} & \text{otherwise} \end{cases} \quad (4)$$

and

$$S = \lim_{n \rightarrow \infty} T_*^n G$$

where  $T_*^n G \geq T_*^{n+1} G$  and the limit is taken pointwise.

The incentive constraint in  $T_*$  implies that  $T_*$  is not a contraction.<sup>13</sup> Despite this, we can still obtain  $S$  by iteratively applying  $T_*$  on the efficient surplus function  $G$ . This result exploits the algorithmic result of APS. The APS algorithm obtains the equilibrium payoff set by an iterative application of a set-valued operator on a large enough payoff set. Since our setting has a separate equilibrium payoff set for each belief, we adapt the algorithm to operate on payoff correspondences. Then we exploit the fact that the equilibrium set at any belief  $p$  can be characterized entirely by  $S(p)$ . Thus, the APS algorithm can be reduced

<sup>13</sup>To see this, note that  $S$  is only one of many fixed points of  $T_*$ . Another fixed point is a function which equals  $\underline{s}$  for all beliefs in  $[0, 1]$ .

to operating on surplus functions. This characterization, stated in Lemma 1 below, will let us establish a number of key properties of  $S$ , including a complete characterization of the optimal relational contract.

## 5 Results

### 5.1 (In)efficiency of relational contracts

We say that relational contracts are **inefficient** if the optimal relational contract cannot achieve the efficient surplus at some belief i.e.  $S(p) < G(p)$  for some  $p \in [0, 1]$ . In the absence of any private information, the only tension is the separation between the entity incurring the cost of effort (worker) and the entity collecting the output (firm). In an environment with known match quality, MacLeod and Malcomson (1989) showed that when the players are sufficiently patient, relational contracts are efficient. The uncertainty about match quality, even though symmetric, may overturn this result. Proposition 2 provides a necessary and sufficient condition for inefficiency.

**Proposition 2.** *Relational contracts are inefficient if and only if  $R_H(p^{\text{FB}}) > \underline{s}$ .*

We proceed to illustrate the argument. Since every belief admits an equilibrium where the parties receive their outside options, the efficient solution can be replicated in a relational contract at every belief  $p < p^{\text{FB}}$ . Hence, to establish efficiency, it suffices to show that every belief  $p \geq p^{\text{FB}}$  admits an equilibrium where the worker chooses  $H$ . This amounts to showing that  $H$  can be incentivized at  $p^{\text{FB}}$  since relational contracts generate more surplus at higher beliefs. Since  $G(p^{\text{FB}}) = \underline{s}$ , we have

$$(1 - \delta) (R_H(p^{\text{FB}}) - c) + \delta \mathbb{E}^H[G(p)] = \underline{s}.$$

This means that the cost of  $H$  is exactly compensated by the sum of two terms: the contemporaneous output and continuation surplus in excess of the sum of the outside options:

$$\underbrace{(1 - \delta) c}_{\text{Today's cost}} = \underbrace{(1 - \delta)(R_H(p^{\text{FB}}) - \underline{s})}_{\text{Today's excess surplus}} + \underbrace{\delta(\mathbb{E}^H[G(p)] - \underline{s})}_{\text{Tomorrow's excess surplus}}.$$

In equilibrium, the firm cannot commit to rewarding the worker with the output in the current period. This creates a holdup problem when today's excess surplus must be



used to compensate the worker for his effort in the efficient solution. The worker can be incentivized to choose  $H$  only when tomorrow's excess surplus exceeds today's cost, as evidenced by (4). If  $R_H(p^{\text{FB}}) > \underline{s}$ , tomorrow's excess surplus must be less than today's cost, i.e., the incentive constraint (4) at  $p^{\text{FB}}$  fails even if the continuation surplus in relational contracts  $S$  equals the efficient surplus  $G$ . Since  $S \leq G$ , relational contracts are inefficient. Conversely, if  $R_H(p^{\text{FB}}) \leq \underline{s}$ , then the incentive constraint at  $p^{\text{FB}}$  is satisfied assuming that the continuation surplus is given by  $G$ . A fixed point argument using Lemma 1 then establishes that relational contracts are efficient.<sup>14</sup>

Since the source of inefficiency is a holdup problem on today's excess surplus, one might think that the introduction of output-contingent spot contracts would restore it. However, as we show in Section 7, as long as the worker is protected by limited liability, the inefficiency persists. Thus, the inefficiency is not a direct consequence of unknown match quality but rather its subtle interaction with incomplete contracts.

While Proposition 2 places a condition on the cutoff belief in the efficient outcome which is an endogenous object, Corollary 1 below gives a sufficient condition for inefficiency, based only on the primitives of the model.

**Corollary 1.** *If  $R_H(0) \geq \underline{s}$ , then relational contracts are inefficient.*

Moreover, if the condition above fails, i.e.  $R_H(0) < \underline{s}$ , then, relational contracts are efficient for a sufficiently high  $\delta$ . When  $\delta$  is fixed, it is also possible to generate sets of parameters with positive measures for which relational contracts are efficient or inefficient.

## 5.2 Optimal relational contracts

We now give a characterization of the optimal relational contracts.

**Proposition 3.** *An optimal relational contract is described by two cutoff beliefs:  $\underline{p}$  and  $\bar{p}$  such that  $\bar{p} \geq \underline{p} \geq p^{\text{FB}}$ . At any belief  $p$*

- *The firm and the worker interact if and only if  $p \geq \underline{p}$ .*
- *The worker exerts  $L$  if  $p \in [\underline{p}, \bar{p})$ .*
- *The worker exerts  $H$  if  $p \geq \bar{p}$ .*

---

<sup>14</sup>The nature of the holdup problem is not tied to the specific information structure we consider. The argument generalises to a setting with continuous output under sufficient regularity.

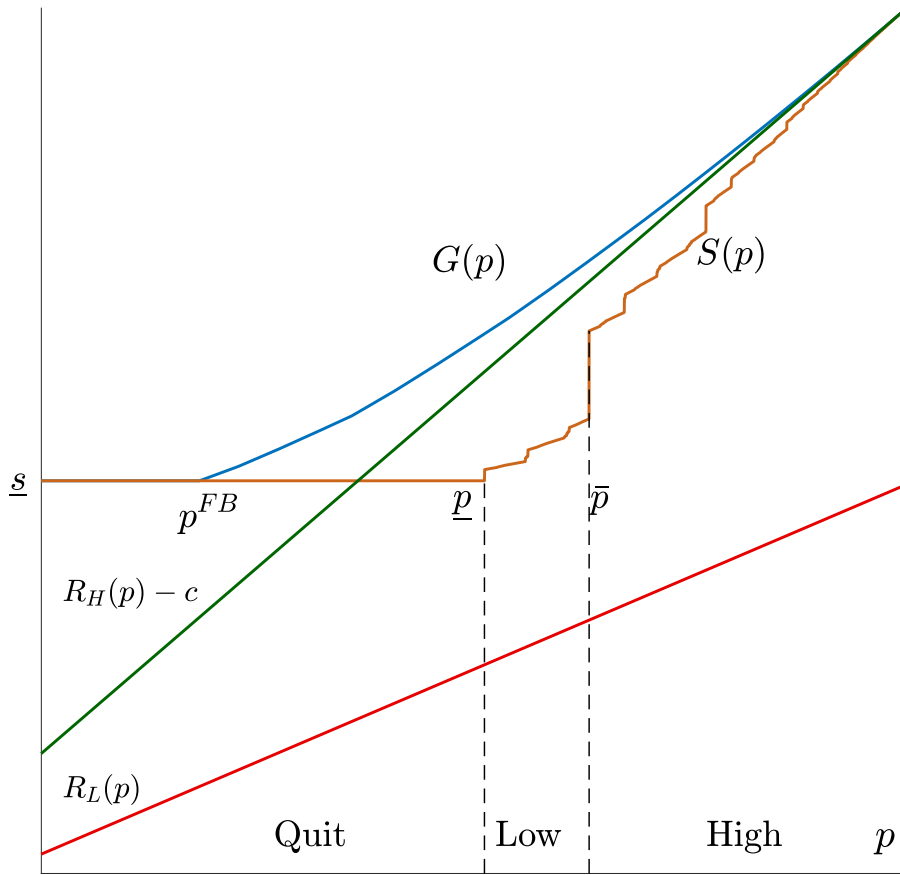


Figure 4: Plot of  $G$  and  $S$

The structure of optimal relational contracts can be interpreted as follows. When the players are sufficiently pessimistic about the match quality, the (expected) value of the relationship is low and learning has a high opportunity cost, so it is optimal to fire the worker. When the players are highly optimistic, the future value of the relationship is sufficient to incentivize the worker to exert  $H$ . At intermediate beliefs there is not enough surplus to provide incentives, but the firm would like to retain the worker, as it is reasonably likely that the match quality is good. This is accomplished by the worker choosing  $L$  which is cheaper to incentivize and also offers some learning.

The structure of the optimal relational contracts resembles PIPs mentioned in the introduction. Typically, a worker not meeting the performance standards is placed on a PIP where he is monitored closely. In the event that his performance improves, the worker is moved out of a PIP; otherwise, he is fired. In our model, this corresponds to a belief

that dips below  $\bar{p}$  and climbs back above it. These dynamics are different from the ideas of “starting small” (Ghosh and Ray, 1996; Watson, 1999, 2002; Halac, 2014). In these models, parties build up the value of their relationship monotonically as long as it is not terminated. In our model, beliefs do not evolve monotonically so the value of the relationship may not improve with time. The parties realise that a drop in beliefs may happen even if the match quality is good, so the worker is given a chance to improve his performance, instead of being fired immediately. Hence, our model explains why a worker in good standing can be placed on a PIP, complete it successfully and continue his career with the firm.

While Proposition 3 does not guarantee that  $L$  is used in an optimal relational contract, Proposition 4 below gives a condition on the parameters of the model, under which this is the case. This condition implies the sufficient condition for inefficiency in Corollary 1.

**Proposition 4.** *If  $R_L(0) - \underline{s} + \delta\pi_L(0)(R_H(0) - \underline{s}) \geq 0$  then  $\underline{p} < \bar{p}$  in Proposition 3.*

We proceed to give an overview of the proof of Proposition 3. A starting intuition is that  $H$  should be preferable to  $L$  at any belief, so that the worker exerts  $L$  only if the incentive constraint for  $H$  cannot be satisfied. Incentive provision being easier at higher beliefs (as there is more future surplus to promise the worker), the three-region structure should result. However, this intuition is precluded by the nonconvexity of  $S$ , the surplus function in the optimal relational contract (Figure 4).<sup>15,16</sup> Hence, more informative actions are not necessarily preferable. Recall that  $L$  is never used in the efficient benchmark because more information is better, and  $H$  dominates  $L$  on both the informativeness and payoff dimensions. In a relational contract,  $L$  may be preferred to  $H$  at some beliefs where the gains from the inferior information in  $L$  outweigh the payoff losses. Therefore, we cannot dismiss the possibility of multiple switches between  $L$  and  $H$  as the beliefs improve.

The key to overcoming this obstacle lies in the APS algorithm as seen in Lemma 1. As  $S$  can be obtained through iterative application of  $T_*$  to the efficient surplus function  $G$ ,  $S$  inherits the properties of  $G$  that are preserved under  $T_*$ . While  $T_*$  does not preserve convexity (or even continuity), it preserves the weaker property that, barring incentive constraints,  $H$  is preferred to  $L$ .

<sup>15</sup>The parameters used for the plot are  $\gamma_H = 0.7$ ,  $\beta_H = 0.45$ ,  $\gamma_L = 0.35$ ,  $\beta_L = 0.226$ ,  $c = 0.22$ ,  $\delta = 0.85$ ,  $\underline{u} = 0.15$ ,  $\underline{v} = 0.2$ .

<sup>16</sup> $S$  exhibits a jump at  $\bar{p}$  where the optimal effort switches from  $L$  to  $H$ . The reason is that the surplus from  $H$  exceeds the surplus from  $L$  at any belief, which will be shown as part of the proof of Proposition 3. Even though  $H$  would create more surplus than  $L$  at beliefs lower than  $\bar{p}$ , the incentive constraint precludes its use, resulting in the jump. Consequently, there are jumps at beliefs, whose posteriors reach  $\bar{p}$ , or can be reached from  $\bar{p}$ .

**Definition 1.** A function  $f : [0, 1] \rightarrow \mathbb{R}$  exhibits a monotone policy if  $T_H f \geq T_L f$ .

If an increasing function  $f$  exhibits a monotone policy, then the optimal policy associated with  $f$  will exhibit the three-region structure of Proposition 3. Therefore, the proof boils down to showing that if an increasing function  $f$  exhibits a monotone policy, then  $T_* f$  is also an increasing function that exhibits a monotone policy. The details of this argument are provided in the appendix.

Proposition 3 resembles the characterization in the infinitely repeated stochastic partnership game studied by McAdams (2011). Therein, the state, which evolves exogenously, captures the temptation to cheat. Though outwardly different, these two models have a similar intuition behind obtaining the three-region characterization: providing incentives is difficult when future prospects of the relationship look grim. In such situations, parties interact but indulge in inefficient activities, hoping that the relationship could drift towards an optimistic regime. In our model, the state variable that affects the incentives is the belief about the match quality which evolves *endogenously*.

## 6 Asymptotic Behavior

In this section we explore the limit of our baseline model as the parties become arbitrarily patient. There are two typical approaches. Folk theorems analyze the limit as  $\delta \rightarrow 1$ , ceteris paribus. It is also possible to consider a frequent action limit in which the length of each period is reduced to 0 without changing patience per unit of time. In a seminal paper, Abreu, Milgrom, and Pearce (1991) (AMP) showed that these two approaches can produce radically different results in games of imperfect monitoring. In repeated games of perfect monitoring the two approaches are equivalent when information is complete. This equivalence breaks down in our richer setting due to incomplete information about match quality. For example, if information accumulates with calendar time, shortening the period length degrades the quality of information. Therefore, the presence of learning about match quality is a unique opportunity to explore differences between these two approaches in a game of perfect monitoring.

We do this by studying a sequence of (discrete time) instances of our model with  $\delta$  converging to 1 along the sequence. As  $\delta$  grows, the quality of information per period may (*but need not*) degrade. The main result of this section, Theorem 1, shows that the speed at which information degrades relative to the speed at which impatience vanishes

is crucial for the limit behavior of our model. If information degrades at a slower rate, a folk theorem obtains and relational contracts are efficient in the sense that the sup norm of  $G - S$  converges to 0 as  $\delta \rightarrow 1$ . Wiseman (2005) obtains a similar result when information is unchanged along the sequence by proving a folk theorem in general repeated games with learning.<sup>17,18</sup> Our result can be seen as a mild extension of this folk theorem to a case in which information can degrade, but it does so more slowly than the rate at which players become patient.

On the other hand, if information degrades at a rate equal to or higher than the rate of decay of impatience, relational contracts can be inefficient in the limit – the holdup problem of Proposition 2 persists even with patient players. This result stands in contrast to the folk theorem of Peski and Wiseman (2015) where the probability of a state change vanishes with period length. In our setting it is possible that the probability of a substantial change in state (beliefs) vanishes with period length, yet the belief always changes and a folk theorem may or may not hold.

The case where information and impatience degrade at the same rate is reminiscent of AMP. They study a two player prisoner’s dilemma with imperfect monitoring. A key take-away in the frequent action limit is that the sustainability of cooperation crucially depends on the nature of information. In their “good news” environment, signals arrive according to a Poisson distribution, and their rate is higher when players cooperate than when they defect. In the “bad news” environment, the rate of Poisson signals is higher when the players defect than when they cooperate. The bad news environment facilitates full cooperation when the players are sufficiently patient, while the good news environment offers no hope thereof. The source of this difference is imperfect monitoring, where information is used to statistically infer the actions played in the past. In our environment – learning with perfect monitoring – actions are public, and the information is only used to learn about future prospects. Therefore, regardless of the environment (good news or bad news), the holdup problem of Proposition 2 gives rise to inefficiencies when information degrades at the same rate as patience. Qualitatively there is no distinction between the two environments, in contrast to AMP. We illustrate this point with a simple example at the end of this

---

<sup>17</sup>Technically, the folk theorem of Wiseman (2005) does not apply to our setting because the set of individually rational payoffs in the game where match quality is known to be bad has an empty interior.

<sup>18</sup>For further work on folk theorems with learning see Fudenberg and Yamamoto (2010), Fudenberg and Yamamoto (2011), Wiseman (2012), and Yamamoto (2017). Well-known folk theorems for stochastic games (e.g. Hörner et al. (2011)) do not apply to our setting because the space of beliefs is not an irreducible state space due to the option of quitting.

section.

Formally, we consider a sequence of models  $(\mathcal{M}_k)_{k=1}^\infty$  with parameters indexed by  $k$  and cutoffs  $p_k^{\text{FB}}$ ,  $\underline{p}_k$ ,  $\bar{p}_k$  as in Proposition 1 and Proposition 3, such that:

1.  $\delta_k \rightarrow 1$  as  $k \rightarrow \infty$
2. None of  $c_k$ ,  $R_H^k$ ,  $R_L^k$ ,  $\underline{u}_k$ ,  $\underline{v}_k$  vary with  $k$ .
3. The limits  $p_*^{\text{FB}} = \lim_{k \rightarrow \infty} p_k^{\text{FB}}$  and  $\bar{p}_* = \lim_{k \rightarrow \infty} \bar{p}_k$  exist, as well as the (possibly infinite) limits of  $\frac{\gamma_H^k}{\beta_H^k}$ ,  $\frac{1-\gamma_H^k}{1-\beta_H^k}$ , and  $\frac{\gamma_H^k - \beta_H^k}{1-\delta_k}$  as  $k \rightarrow \infty$ .

The second assumption is merely a normalization of per-period payoffs.<sup>19</sup> Henceforth, payoffs will not be indexed by  $k$ , in order to economize on notation. The last assumption involves no loss of generality, as it amounts to considering a subsequence.

In addition, we assume that either the limiting posterior belief following  $H$  and success exceeds any interior prior or the limiting posterior belief following  $H$  and failure is below any interior prior. This is formally stated in Assumption 4 below.

**Assumption 4.** *At least one of the following holds:*

- (a)  $\lim_k \frac{\gamma_H^k}{\beta_H^k} \in (1, \infty]$
- (b)  $\lim_k \frac{1-\gamma_H^k}{1-\beta_H^k} \in [0, 1)$ .

We measure the informativeness of  $H$  via the variance of the posterior belief:

$$\text{Var}_H^k(p) = \pi_H^k(p) \left( \phi_{H,k}^+(p) - p \right)^2 + (1 - \pi_H^k(p)) \left( \phi_{H,k}^-(p) - p \right)^2.$$

The limit of the ratio  $\frac{\text{Var}_H^k(p)}{1-\delta_k}$  measures the rate of decay of information per period relative to the rate at which impatience vanishes, when the prior belief is  $p$ . In the proof of Theorem 1 we show that this limiting ratio exists and is either finite for all  $p \in (0, 1)$  or infinite for all  $p \in (0, 1)$ . The relative rate of information decay plays a key role in the characterization of the limiting outcome presented below.

**Theorem 1.** *The following are equivalent:*

- (i)  $\lim_{k \rightarrow \infty} \frac{\text{Var}_H^k(p)}{1-\delta_k}$  is finite for all  $p \in (0, 1)$

---

<sup>19</sup>It is possible to normalize payoffs per unit of time to the same effect.

(ii)  $p_*^{FB} > 0$ .

(iii)  $\bar{p}_* > 0$ .

In addition, if  $R_H(p_*^{FB}) > \underline{s}$  then (i)-(iii) are also equivalent to

(iv)  $\bar{p}_* > p_*^{FB}$ .

Notice that  $R_H(p_*^{FB}) > \underline{s}$  is an analogue of the condition in Proposition 2, which characterizes the holdup problem in the baseline model.

The centerpiece of Theorem 1 is the class of models where the informativeness of  $H$  per period decays at the same rate as the players' discounting. These models approximate a natural environment in which information about match quality arrives in continuous time, but the players act at discrete points in time, with actions becoming arbitrarily frequent. Under this interpretation, information is proportional to the length of the period and Theorem 1 shows that the tensions of the baseline model persist in the limit: First, the learning problem is nontrivial, since the efficient solution exhibits an interior cutoff that separates outside options and experimentation through  $H$ . Second, there is an incentive problem associated with the use of  $H$  in relational contracts. The holdup problem from Proposition 2 which characterizes inefficiency in the baseline model also characterizes inefficiency in the limit. Two natural examples, outlined at the end of this section, are models in which success or failure are realized with probability proportional to the length of the period.

In addition to characterizing the limiting outcome of these frequent action models, Theorem 1 shows that they are at the boundary of two qualitatively different environments. When information decays faster than the rate of impatience, the holdup problem highlighted in Proposition 2—and thereby the resulting inefficiency—persists in the limit. On the other hand, when information decays at a comparatively slower rate, or does not vanish at all, we obtain a folk theorem: At any positive belief, the efficient solution and the optimal relational contract involve experimentation through  $H$  when the discount factor is high enough. The reason is that the opportunity cost of learning vanishes relative to its benefit.

In our baseline model, one consequence of inefficiency is that the use of  $L$  may increase the surplus of an optimal relational contract. However, this property is not preserved in the limit. Let  $S$  denote the pointwise limit of  $(S_k)$ , taking a subsequence, if necessary.<sup>20</sup> Proposition 5 below states that  $L$  does not create any surplus in the limit.

**Proposition 5.** *Let  $\hat{p} = \inf\{p \mid S(p) > \underline{s}\}$ . Then  $\bar{p}_* \leq \hat{p}$ .*

<sup>20</sup>The limit exists by Helly's selection theorem and Lemma 6.

## 6.1 Example: Good and bad news Poisson models

We illustrate the results of this section with an example of a good news Poisson model in which the players become patient at the same rate as the rate of information decay. Consider a sequence  $\Delta_k$  of period lengths converging to 0. The discount factor is  $\delta_k = e^{-\Delta_k}$ . The probabilities of success with known match quality in model  $\mathcal{M}_k$  are given by

$$\begin{aligned}\gamma_H^k &= \gamma_H \Delta_k & \gamma_L^k &= \gamma_L \Delta_k \\ \beta_H^k &= \beta_H \Delta_k & \beta_L^k &= \beta_L \Delta_k.\end{aligned}$$

This approximates a Poisson arrival of success as  $\Delta_k \rightarrow 0$ . Such an information structure is interpreted as a “good news” model because the rare signal makes players more optimistic about their match quality causing their belief to jump upwards.<sup>21</sup>

The values of success and failure are  $y_k^h = \frac{1}{\Delta_k}$  and  $y_k^l = 0$  respectively, so the revenue from effort  $e$  is  $R_e(p) = p\gamma_e + (1 - p)\beta_e$ .

The proof of Theorem 1 shows that the variance of the posterior is proportional to  $\gamma_H^k - \beta_H^k$  which is of the order of  $\Delta$ . Hence,  $\lim_{k \rightarrow \infty} \frac{\text{Var}_H^k(p)}{1 - \delta_k}$  is finite for all  $p \in (0, 1)$ . Theorem 1, therefore, implies that the learning problem in the efficient solution is non-trivial. In addition, the inefficiency of relational contracts can persist in the limit, whenever the holdup problem of Proposition 2 arises. A sufficient condition for inefficiency is  $R_H(0) > c$ , similarly to Corollary 1.

Alternatively, we can consider a bad news model where the probability of failure is proportional to  $\Delta$ . To be more precise, the probabilities of failure with known match quality in model  $\mathcal{M}_k$  are given by

$$\begin{aligned}1 - \gamma_H^k &= \gamma_H \Delta_k & 1 - \gamma_L^k &= \gamma_L \Delta_k \\ 1 - \beta_H^k &= \beta_H \Delta_k & 1 - \beta_L^k &= \beta_L \Delta_k.\end{aligned}$$

The value of a success and failure are  $y_k^h = 0$ ,  $y_k^l = -\frac{1}{\Delta}$  respectively.

Unlike the good news model, here the “news” is a failure and, upon its arrival the posterior jumps discretely downward. Also, in contrast to the imperfect monitoring setting of AMP—both the models exhibit inefficiencies due to a holdup problem even asymptotically.

Lastly, a very special case of the good news model is the “fully revealing good news”

<sup>21</sup>In the experimentation literature “good news” and “bad news” is often conclusive, i.e. one signal establishes the match quality. This is a special case of our model when  $\beta_H = \beta_L = 0$ .



model where  $\beta_H = \beta_L = 0$ . Here, any success conclusively establishes that the match quality is good. This model parallels Hörner and Samuelson (2013) and similar inefficiencies arise in this environment as in theirs.

## 7 Extensions

Our model can be extended in various directions. We briefly discuss some natural directions in which our results generalize.

### 7.1 Spot contracts contingent on output

As we saw in Section 5.1, the source of inefficiency is a holdup problem arising from the separation between the firm that collects the output and the worker, who exerts effort. A natural question is whether this inefficiency would disappear if the parties could write one period contracts contingent on output. As we show below, limited liability plays a key role in the answer. If the worker is protected by limited liability, output contingent contracts cannot achieve efficiency unless it was already achievable with constant wages. But, when the worker is not protected by limited liability, relational contracts achieve efficiency by spreading the wages for success and failure to induce high effort.

**Proposition 6.** *Suppose the firm can commit to one-period output-contingent contracts. Then efficiency can be achieved if the worker is not protected by limited liability. If the minimum wage the firm must offer the worker for any level of output is  $\underline{w}$ , then relational contracts are inefficient iff  $R_H(p^{FB}) > \underline{w}$ .*

### 7.2 Combination of verifiable and non-verifiable actions

In most organizations, there are some contractible aspects of a worker’s behavior. Hence, the lack of contractibility in our model may seem extreme. However, it is also common that some aspects of the worker’s behavior are not verifiable in court. For example, a teaching assistant’s wage can be contingent on the number of sections she teaches but not on how well she answers questions. As we show here, the holdup problem from the baseline model remains as long as there is some non-contractibility of the worker’s actions. Moreover, in any optimal relational contract, the parties can find the use of dominated actions beneficial for reasons similar to Proposition 3.

We model a mix of contractible and non-contractible actions by considering two-dimensional effort  $e \in \{L, H\} \times \{L, H\}$ . We assume that the first component of  $e$  is verifiable in court, while the second is not. The cost of effort is separable in the two components, and  $H$  is more costly than  $L$  on each of them. Effort is ranked informationally as well as in payoff as follows:

$$(H, H) \succ (H, L) \succ (L, H) \succ (L, L)$$

Naturally, the efficient outcome involves using  $(H, H)$  above a certain cutoff belief and taking the outside option below. The argument for inefficiency in Proposition 2 is virtually unchanged, as the same holdup problem arises on the non-contractible dimension. Moreover, an analogous result to our Proposition 4 holds, showing that an optimal relational contract would involve using  $L$  on the non-contractible dimension.

### 7.3 Continuous effort

So far we have analyzed a setting with only two efforts, one of which is both more informative and more profitable. A natural question is whether a continuous choice of effort would produce similar inefficiencies. In this section we enrich the model by making the choice of effort continuous. Also, a more informative effort may not necessarily be more profitable. As mentioned previously, such correlated bandit problems are very difficult to analyze in discrete time once the information and payoffs ranking do not coincide. However, we show that our central findings are true in this model as well: (i) relational contracts can be inefficient and (ii) efforts that are worse for both information and payoff are used in an optimal relational contract.

To this end, consider a modification of the baseline model where effort lies in  $[0, 1]$ . Let there be two strictly increasing, concave functions  $\gamma$  and  $\beta$  defining the probability of success conditional on the choice of effort ( $e$ ) and match quality ( $\theta$ ) as follows:

$$\begin{aligned}\gamma(e) &= \text{Prob}(y^h | \theta = G, \text{effort} = e) \\ \beta(e) &= \text{Prob}(y^h | \theta = B, \text{effort} = e).\end{aligned}$$

Let  $c(e)$  be the cost of effort  $e$  where  $c$  is strictly increasing and convex. The myopic profit

from effort  $e$  at prior  $p$  is given by

$$h(p, e) := p[\gamma(e)y^h + (1 - \gamma(e))y^l] + (1 - p)[\beta(e)y^h + (1 - \beta(e))y^l] - c(e).$$

We assume that  $\gamma$ ,  $\beta$ , and  $c$  are differentiable and satisfy the following assumption.

**Assumption 5.** 1.  $\frac{\gamma(e)}{\beta(e)}$  is increasing in  $e$ , and  $\gamma'(e) > \beta'(e)$

2.  $c(0) = c'(0) = 0$ .

3.  $\gamma'(1) < c'(1)$

The myopically efficient action at belief  $p$  is denoted by  $\hat{e}(p)$  given by,

$$\frac{\partial h(p, \hat{e}(p))}{\partial e} = [p\gamma'(\hat{e}(p)) + (1 - p)\beta'(\hat{e}(p))][y^h - y^l] - c'(\hat{e}(p)) = 0.$$

Assumption 5 implies that  $\hat{e}$  is increasing and  $0 < \underline{e} := \hat{e}(0) < \hat{e}(1) =: \bar{e} < 1$ . That is, conditional on the match quality being good (bad), the efficient action is  $\bar{e}$  ( $\underline{e}$ ) which is strictly less than 1 (larger than 0).

The measure of informativeness for an effort  $e$  is the fraction  $\frac{\gamma(e)}{\beta(e)}$ , which is increasing due to Assumption 5. Hence, higher effort levels are more informative, but not necessarily more profitable.

As in the baseline model, the optimal policy for the efficient outcome is characterized by a cutoff  $p^{\text{FB}}$ , below which the outside options are taken. Also, the value function is convex. Therefore, efforts dominated in both payoff and informational content are not used. In particular, efforts lower than  $\underline{e}$  are never used.

The holdup problem in the baseline model arises because the action chosen at the efficient cutoff is costly to the worker. Here too, since the lowest effort chosen is  $\underline{e}$ , the choice of effort at  $p^{\text{FB}}$  will involve a cost to be borne by the worker. Therefore, the qualitative nature of the holdup problem is identical to the case of two efforts resulting in efficiency not being achievable in any relational contract. Moreover, a condition analogous to our Proposition 4 shows that efforts that are worse for both information and payoff will be used in the optimal relational contract.

While we have highlighted the form of inefficiency resulting in the use of efforts that are inferior for information as well as payoff, in this richer model, the optimal relational contract may involve the worker choosing an inefficiently higher effort. This can be seen numerically as shown in Figure 5. Notice that, for a range of beliefs in the middle, the

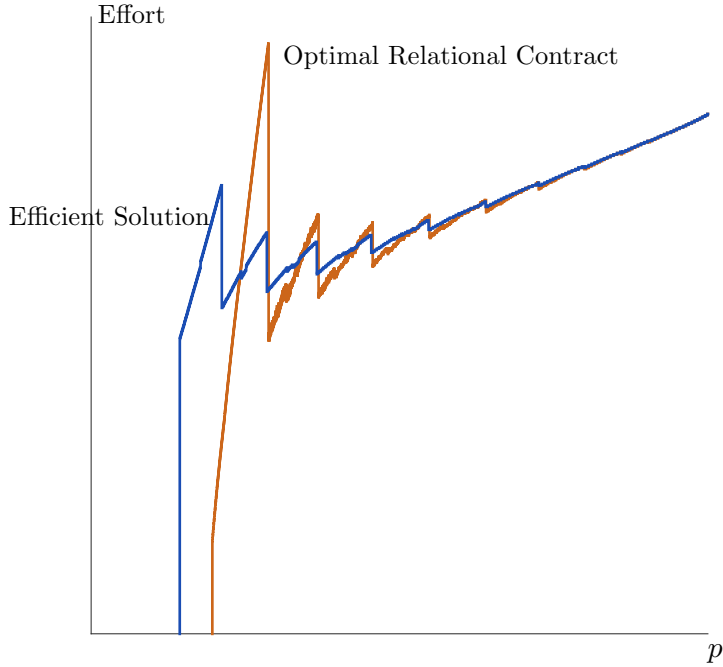


Figure 5: Plot of effort in the planner's and agency problem.  
Parameters:  $\gamma(e) = 0.95e, \beta(e) = \frac{2}{7}\sqrt{e}, c(e) = 0.9e^2 - 0.009, e \in [0.1, 1]$ .

worker chooses an effort that is higher than the efficient benchmark. Also, as we can see, the optimal effort is non-monotonic in both the efficient benchmark and the optimal relational contract. This non-monotonicity arises because the informational and payoff rankings do not coincide. This is precisely why correlated bandits are difficult to analytically understand in such problems.

## 7.4 Exogenous Learning

The three-region structure obtained in Proposition 3 suggests that higher effort is used at higher beliefs when efforts have the same informational and profitability rankings. Although this property does not generalize to a setting with more than two effort levels, monotonic effort can be recovered in a simpler model where each effort is equally informative.<sup>22</sup>

To formalize this model, suppose the set of effort levels is  $[0, 1]$  and the probabilities of success given good and bad match quality are  $\gamma$  and  $\beta$ , respectively, regardless of the effort. Let  $g : [0, 1] \rightarrow \mathbb{R}_+$  and  $c : [0, 1] \rightarrow \mathbb{R}_+$  be strictly increasing,  $C^2$  functions describing the

<sup>22</sup>McAdams (2011) also considers exogenous learning with relational contracts set in a partnership game with two effort levels and shows a similar monotonicity of effort.

revenue from success and the cost of effort, respectively. We assume  $g$  is concave and  $c$  is convex with  $c(0) = 0$ . The revenue from failure is 0, regardless of the effort. Hence, the myopic profit of effort  $e$  at a prior  $p$ , is given by

$$h(p, e) = \underbrace{[p\gamma + (1-p)\beta]}_{\pi(p)} g(e) - c(e).$$

In particular, if we assume that  $\pi(p)g'(1) - c'(1) > 0$  for all  $p$ , then the efficient solution would naturally involve the outside options below a cutoff  $p^{\text{FB}}$  and effort  $e = 1$  above it.

Now consider the surplus from effort  $e$  at prior  $p$  in a relational contract:

$$(1 - \delta)h(p, e) + \delta \mathbb{E}^e S(p).$$

Since all efforts are equally informative, they all bring the same expected continuation surplus  $\mathbb{E}^e[S(p)]$ . Hence, higher effort is more profitable. Its use is precluded only by the incentive constraint:

$$\delta[\mathbb{E}^e S(p) - \underline{s}] \geq (1 - \delta)c(e).$$

Since the cost of effort is increasing, the set of incentive compatible effort levels at prior  $p$  is an interval  $[0, \bar{e}(p)]$  where  $\bar{e}$  is increasing. The optimal relational contract involves the outside options below a threshold belief  $\underline{p}$  and effort  $\bar{e}(p)$  at any higher belief. Hence, the effort is smoothly increasing in the belief. As in the continuous effort model in the previous section, our main results regarding inefficiency and the use of dominated efforts in equilibrium can be shown to hold.

## 8 Conclusion

In this paper, we studied a dynamic firm-worker relationship with symmetric uncertainty about match quality and limits to contractibility. There are two main takeaways. First, despite the absence of private information, uncertainty about the match quality can preclude efficiency. Second, the presence of actions that produce both worse information and worse payoffs offer a partial remedy, as they may be cheaper to incentivize. The inefficiency is robust to various extensions of the model, and persists in a frequent action limit. Moreover, separating the effects of higher patience and shorter period length reveals that

limiting efficiency depends on the relationship between the rates of decay of information and impatience.

In our view, the two key features in our environment – unknown match quality and incompleteness of contracts – are pervasive. Therefore, our model could serve as a building block toward studying various other economically important scenarios, such as obtaining the nature of relational contracts when there is partial contractibility in the form of spot contracts on output or even long run contracts on output. Another natural direction is to endogenize the outside option by embedding our model in a labor market with search frictions. Also, as mentioned in Section 3.2, our model is formally equivalent to a setting where the firm assigns jobs to the worker. This interpretation relates us to Pastorino (2018) who finds, in a setting with worker not being strategic, that learning about ability accounts for one quarter of cumulative wage growth of a worker. We believe that our framework can be used to understand the effect of learning on wages when the worker is strategic.

## 9 Appendix

### Proofs from Section 4.1

**Proof of Proposition 1.**  $G$  is the unique solution of the Bellman equation  $f = Tf$ . Standard contraction mapping arguments establish that  $G$  is increasing. To establish the convexity of  $G$ , let  $f \in B[0, 1]$  be convex. It suffices to show that  $T_H f$  and  $T_L f$  are convex, since a maximum of convex functions is convex. To this end, let  $p_1, p_2 \in [0, 1]$ ,  $\lambda \in [0, 1]$  and  $\tilde{p} = \lambda p_1 + (1 - \lambda)p_2$  and notice that

$$\begin{aligned} f(\pi_H^+(\tilde{p})) &= f\left(\frac{\lambda\pi_H(p_1)}{\pi_H(\tilde{p})} \frac{\gamma_H p_1}{\pi_H(p_1)} + \frac{(1-\lambda)\pi_H(p_2)}{\pi_H(\tilde{p})} \frac{\gamma_H p_2}{\pi_H(p_2)}\right) \\ &\leq \frac{\lambda\pi_H(p_1)}{\pi_H(\tilde{p})} f(\pi_H^+(p_1)) + \frac{(1-\lambda)\pi_H(p_2)}{\pi_H(\tilde{p})} f(\pi_H^+(p_2)) \\ \text{and } f(\pi_H^-(\tilde{p})) &\leq \frac{\lambda(1-\pi_H(p_1))}{1-\pi_H(\tilde{p})} f(\pi_H^-(p_1)) + \frac{(1-\lambda)(1-\pi_H(p_2))}{1-\pi_H(\tilde{p})} f(\pi_H^-(p_2)). \end{aligned}$$

It follows that

$$\begin{aligned} T_H f(\tilde{p}) &= (1 - \delta)(R_H(\tilde{p}) - c) + \delta \left( \pi_H(\tilde{p}) f(\pi_H^+(\tilde{p})) + (1 - \pi_H(\tilde{p})) f(\pi_H^-(\tilde{p})) \right) \\ &\leq (1 - \delta)(R_H(\tilde{p}) - c) + \delta \left( \lambda \pi_H(p_1) f(\pi_H^+(p_1)) + (1 - \lambda) \pi_H(p_2) f(\pi_H^+(p_2)) \right) \\ &\quad + \delta \left( \lambda (1 - \pi_H(p_1)) f(\pi_H^-(p_1)) + (1 - \lambda) (1 - \pi_H(p_2)) f(\pi_H^-(p_2)) \right) \\ &= \lambda \left( (1 - \delta)(R_H(p_1) - c) + \delta \mathbb{E}^H[f(p_1)] \right) + (1 - \lambda) \left( (1 - \delta)(R_H(p_2) - c) + \delta \mathbb{E}^H[f(p_2)] \right) \\ &= \lambda T_H f(p_1) + (1 - \lambda) T_H f(p_2) \end{aligned}$$

as required. The argument for the convexity of  $T_L f$  is analogous.

The convexity and monotonicity of  $G$  imply that  $T_H G > T_L G$ , so  $L$  is never used in the optimal policy. By Assumption 1,  $R_H(1) > \underline{s} > R_H(0)$ , so  $G(1) = \gamma_H - c$  and  $G(0) = \underline{s} > T_H G(0)$ . By continuity of  $G$ , there exists a maximal belief  $p^{\text{FB}} \in (0, 1)$  such that  $G(p^{\text{FB}}) = \underline{s}$ . Since  $T_H G$  is strictly increasing,  $G(p) > \underline{s}$  for all  $p > p^{\text{FB}}$ . Hence, the optimal policy is to take the outside options at beliefs  $p \leq p^{\text{FB}}$  and to choose  $H$  at higher beliefs. Clearly,  $p^{\text{FB}}$  is unique.  $\square$

## Proofs from Section 4.2

Here, we adapt the results of APS to our setting. We begin by obtaining the worst equilibrium payoffs.

**Lemma 2.**  $(\underline{u}, \underline{v}) \in \mathcal{E}(p)$  for all  $p$ .

*Proof.* Let  $p \in [0, 1]$ . Consider the following strategies:

- The firm never interacts with the worker.
- The worker accepts wage offer  $w$  if and only if  $w \geq \underline{u}$  and always chooses  $L$ .

As the worker is never rewarded for choosing  $H$ , he chooses  $L$  whenever he accepts the wage offer. Since he expects to receive his outside option in the future he only accepts wage offers above  $\underline{u}$ . Given that the worker chooses  $L$ , the firm does not get enough expected output to justify paying the worker  $\underline{u}$  since  $R_L(p) \leq R_L(1) < \underline{s}$ . Hence, the strategies constitute a PBE with payoffs  $(\underline{u}, \underline{v})$  at belief  $p$ .  $\square$

Let  $\mathcal{C}_0$  denote the set of bounded payoff correspondences  $W : [0, 1] \rightrightarrows \mathbb{R}^2$ . Let  $\mathcal{C}$  denote the set of compact-valued, upper hemicontinuous correspondences  $W \in \mathcal{C}_0$  such that  $(\underline{u}, \underline{v}) \in W(p)$  for all  $p$ .

We say payoffs  $(u, v)$  are enforceable with respect to  $W \in \mathcal{C}_0$  at belief  $p \in [0, 1]$  if they can be decomposed by an action profile in the initial period and continuation payoffs as described in the recursive formulation in Section 4.2 except that the continuation payoffs are drawn from  $W$  instead of  $\mathcal{E}$ .

**Definition 2.** Let  $\mathcal{BW}(p)$  be the convex hull of the set of payoffs  $(u, v)$  which are enforceable with respect to  $W \in \mathcal{C}$  at belief  $p$ .

It is straightforward to check that  $\mathcal{BW} \in \mathcal{C}$  whenever  $W \in \mathcal{C}$ .

**Definition 3.** A correspondence  $W \in \mathcal{C}_0$  is self-generating if  $W(p) \subseteq \mathcal{BW}(p)$  for all  $p$ .

**Proposition 7 (Self-generation).** If  $W \in \mathcal{C}_0$  is self-generating, then  $\mathcal{BW}(p) \subseteq \mathcal{E}(p)$  for all  $p$ .

*Proof.* Let  $p \in [0, 1]$  and  $(u, v) \in \mathcal{BW}(p)$ . If  $(u, v) = (\underline{u}, \underline{v})$  then Lemma 2 implies  $(u, v) \in \mathcal{E}(p)$ . Otherwise,  $(u, v)$  is enforced by effort  $e$ , wage  $w$  and continuation payoffs  $(u^+, v^+) \in \mathcal{BW}(\phi_e^+(p))$  and  $(u^-, v^-) \in \mathcal{BW}(\phi_e^-(p))$  since  $W$  is self-generating. Construct



inductively equilibrium strategies as follows: in the initial period, the firm offers wage  $w$ , the worker accepts and exerts effort  $e$ . Continuation strategies after any deviations follow the PBE in Lemma 2 with payoffs  $(\underline{u}, \underline{v})$ . Continuation strategies on the equilibrium path are constructed by decomposing the continuation payoffs inductively in the same manner. An application of the one-shot deviation principle shows that the strategies form an equilibrium. Boundedness of  $W$  guarantees that the resulting equilibrium payoffs equal  $(u, v)$ , as required.  $\square$

**Proposition 8** (Factorization).  $\mathcal{BE}(p) = \mathcal{E}(p)$  for all  $p$ .

*Proof.* Since equilibrium payoffs are bounded by the outside options and the production technology,  $\mathcal{E} \in \mathcal{C}_0$ . Let  $p \in [0, 1]$ ,  $(u, v) \in \mathcal{E}(p)$ . By Proposition 7 it suffices to show that  $(u, v) \in \mathcal{BE}(p)$ . If  $(u, v) = (\underline{u}, \underline{v})$ , this is trivial. Otherwise, consider without loss of generality a PBE  $\sigma$  at belief  $p$  with payoffs  $(u, v)$  without public randomization in the initial period. The payoffs  $(u, v)$  are enforceable with respect to  $W$  at  $p$ , since they can be decomposed by the first-period effort and wage, and the on-path continuation payoffs following success and failure in  $\sigma$ . To see this, note that  $(u, v) \geq (\underline{u}, \underline{v})$  since the players are free to exercise their outside options. Moreover, the incentive constraint for  $H$  given by (3) holds, since otherwise the agent can profitably deviate from  $\sigma$  in the first period by exerting  $L$  and taking his outside option in the next period. Hence,  $(u, v) \in \mathcal{BE}(p)$ , as required.  $\square$

**Proposition 9** (APS algorithm). Let  $W_0 \in \mathcal{C}$  satisfy  $W_0(p) \supseteq \mathcal{BW}_0(p)$  and  $W_0(p) \supseteq \mathcal{E}(p)$  for all  $p$ . Inductively define  $W_n = \mathcal{BW}_{n-1}$  for all  $n \in \mathbb{N}$ . Let  $W^*(p) = \bigcap_{n \in \mathbb{N}} W_n(p)$  for all  $p$ . Then  $W^*(p) = \mathcal{E}(p)$  for all  $p$ .

*Proof.* Observe that  $\mathcal{B}$  is monotonic: any correspondences  $U, U' \in \mathcal{C}_0$  with  $U(p) \subseteq U'(p)$  for all  $p$  satisfy  $\mathcal{BU}(p) \subseteq \mathcal{BU}'(p)$  for all  $p$ . Hence  $W_{n-1}(p) \supseteq W_n(p)$  for all  $n, p$  and consequently,  $W^*(p) \supseteq \mathcal{E}(p)$ . To show the converse inclusion it suffices to show  $W^*$  is self-generating. To this end, let  $p \in [0, 1]$  and  $(u, v) \in W^*(p)$ . Then  $(u, v) \in \mathcal{BW}_n(p)$  for all  $n$ . We want to show  $(u, v) \in \mathcal{BW}^*(p)$ . It is sufficient to consider  $(u, v) \neq (\underline{u}, \underline{v})$ . Then, there exists an effort level  $e$  such that  $(u, v) \in \mathcal{BW}_n(p)$  for infinitely many  $n$ . Along this subsequence, let  $w_n$  denote the wage offered,  $(u_n^+, v_n^+)$  and  $(u_n^-, v_n^-)$  be the continuation payoffs following success and failure respectively. Without loss of generality these sequences converge to  $w$ ,  $(u^+, v^+)$  and  $(u^-, v^-)$  respectively. Compactness of  $W_n(\phi_e^+(p))$  and  $W_n(\phi_e^-(p))$  and monotonicity in  $n$  imply  $(u^+, v^+) \in W^*(\phi_e^+(p))$  and

$(u^-, v^-) \in W^*(\phi_e^-(p))$ . Hence,  $(u, v)$  is enforceable with respect to  $W^*$  at belief  $p$  through wage  $w$  and continuation payoffs  $(u^+, v^+)$  and  $(u^-, v^-)$ , that is  $(u, v) \in \mathcal{BW}^*(p)$ .  $\square$

**Corollary 2.**  $\mathcal{E}$  is compact-valued and upper hemicontinuous.

*Proof.* This follows from the fact that each  $W_n$  in the statement of Proposition 9 is compact-valued and upper hemicontinuous.  $\square$

### Proofs from Section 4.3

It will be useful to operate on payoff correspondences that, like the equilibrium correspondence, are fully described by the maximum sum of payoffs that can be achieved at every belief.

**Definition 4.** Let  $\mathcal{D}$  be the set of all correspondences  $W \in \mathcal{C}$  such that

$$W(p) = \{(u, v) \mid (u, v) \geq (\underline{u}, \underline{v}), u + v \leq S^W(p)\} \quad \text{for all } p$$

for some  $S^W : [0, 1] \rightarrow [\underline{s}, \gamma_H - c]$ .

**Lemma 3.** If  $W \in \mathcal{D}$ , then  $\mathcal{BW} \in \mathcal{D}$ .

*Proof.* Let

$$S^{\mathcal{BW}}(p) = \max_{(u, v) \in \mathcal{BW}(p)} u + v \quad \text{for all } p.$$

Let  $p \in [0, 1]$ . Notice that  $S^{\mathcal{BW}}(p) \geq \underline{s}$  by definition of the operator  $\mathcal{B}$ . Moreover,  $R_H(1) - c$  is an upper bound on the sum of payoffs in a single period so  $S^{\mathcal{BW}}(p) \in [\underline{s}, R_H(1) - c]$ . It only remains to show that if  $(u, v) \geq (\underline{u}, \underline{v})$  and  $u + v = S^{\mathcal{BW}}(p)$  then  $(u, v) \in \mathcal{BW}(p)$  (the rest follows from taking convex combinations of  $(\underline{u}, \underline{v})$  and efficient payoffs in  $\mathcal{BW}(p)$ ). The statement clearly holds if  $S^{\mathcal{BW}}(p) = \underline{s}$  so suppose  $S^{\mathcal{BW}}(p) > \underline{s}$  and consider any such  $(u, v)$ . There exists a payoff  $(u', v') \in \mathcal{BW}(p)$  with  $u' + v' = S^{\mathcal{BW}}(p)$  which does not involve public randomization, since  $\mathcal{BW}(p)$  is a compact, convex subset of  $[\underline{u}, \infty) \times [\underline{v}, \infty)$ . The payoff  $(u, v)$  can be supported by the same effort and continuation payoffs as  $(u', v')$  through an adjustment of the wage.  $\square$

Lemma 4 states that in any optimal relational contract it is without loss of generality to give the worker the highest equilibrium continuation payoff. This can only improve the

incentives for the worker to exert  $H$ , and the wage can be used to redistribute the surplus without violating incentives for the firm. The same result holds when  $L$  is enforced but is not needed.

**Lemma 4.** *Let  $W \in \mathcal{D}$ ,  $p \in [0, 1]$ . Consider a payoff  $(u, v)$  such that*

$$u + v = \max\{u' + v' \mid (u', v') \in \mathcal{BW}(p) \text{ is enforced by } H\}.$$

*Then  $(u, v)$  can be enforced by effort  $H$  and continuation payoffs  $\hat{u}^+ = S^W(\phi_H^+(p)) - \underline{v}$  and  $\hat{u}^- = S^W(\phi_H^-(p)) - \underline{v}$  for the worker.*

*Proof.* By definition,  $(u, v)$  is enforced by some wage  $w$  and continuation payoffs  $(u^i, v^i) \in W(\phi_H^i(p))$  for  $i \in \{+, -\}$ . Let  $\hat{v}^i = \underline{v}$  for all  $i$ . Since  $W \in \mathcal{D}$  we have  $(\hat{u}^i, \hat{v}^i) \in W(\phi_H^i(p))$  for all  $i$ . Consider a wage

$$\hat{w} = w - \frac{\delta}{1 - \delta} [\pi_H(p)(v^+ - \underline{v}) + (1 - \pi_H(p))(v^- - \underline{v})].$$

We want to show that  $(u, v)$  is also enforced by  $H$ , wage  $\hat{w}$  and continuation payoffs  $(\hat{u}^+, \hat{v}^+)$  and  $(\hat{u}^-, \hat{v}^-)$  following success and failure respectively.

Efficiency of  $(u, v)$  implies that  $u^i + v^i = S(\phi_H^i(p))$  for all  $i$  so the promise keeping constraints readily hold:

$$\begin{aligned} v &= (1 - \delta)[R_H(p) - \hat{w}] + \delta[\pi_H(p)\hat{v}^+ + (1 - \pi_H(p))\hat{v}^-] \\ u &= (1 - \delta)[\hat{w} - c] + \delta[\pi_H(p)\hat{u}^+ + (1 - \pi_H(p))\hat{u}^-]. \end{aligned}$$

It remains to check the incentive constraints. Clearly,  $(u, v) \geq (\underline{u}, \underline{v})$  so we only need to check the incentive constraint for effort. For any  $i$  we have

$$S^W(\phi_H^i(p)) = u^i + v^i = \hat{u}^i + \underline{v}$$

so  $v^i \geq \underline{v}$  implies  $\hat{u}^i \geq u^i$ . Since the incentive constraint holds in the original decomposition with continuation payoffs  $u^+$  and  $u^-$  we have

$$\begin{aligned} &\pi_H(p)\hat{u}^+ + (1 - \pi_H(p))\hat{u}^- - \underline{u} \\ &\geq \pi_H(p)u^+ + (1 - \pi_H(p))u^- - \underline{u} \geq \frac{1 - \delta}{\delta}c \end{aligned}$$

which completes the proof.  $\square$

The following result follows directly from Lemma 3 and Proposition 9.

**Corollary 3.**  $\mathcal{E} \in \mathcal{D}$ .

Now we are ready to show the equivalence of operating on the frontier of the equilibrium set through the APS operator  $\mathcal{B}$  and operating on the maximum equilibrium surplus function through  $T_*$ .

**Lemma 5.**  $S^{BW}(p) = T_*S^W(p)$  for all  $p$ .

*Proof.* It suffices to show that, for any belief  $p$

- The maximum sum of payoffs in  $BW(p)$  decomposed by  $L$ , if it exists, is equal to  $T_L S^W(p)$ . If no payoffs in  $BW(p)$  are decomposed by  $L$  we have  $T_L S^W(p) < \underline{s}$ .
- The maximum sum of payoffs in  $BW(p)$  decomposed by  $H$ , if it exists, equals  $T_H S^W(p)$  and the incentive constraint below holds.

$$\pi_H(p)S^W(\phi_H^+(p)) + (1 - \pi_H(p))S^W(\phi_H^-(p)) \geq \underline{s} + \frac{1 - \delta}{\delta}c \quad (5)$$

If no payoffs in  $BW(p)$  are decomposed by  $H$  we have either  $T_H S^W(p) < \underline{s}$  or the incentive constraint (5) does not hold.

Let  $p \in [0, 1]$ . The maximum sum of payoffs in  $BW(p)$  decomposed by  $e$ , if it exists, is given by  $T_e S^W(p)$ ; the effort pins down the flow payoff as the wage transfer cancels out and maximality implies that the players share the maximum future surplus, as higher payoffs do not hurt incentives. Moreover, in the case of  $e = H$ , Lemma 4 states that this maximum sum of payoffs can be achieved by setting continuation payoffs  $u^i = S^W(\phi_H^i(p)) - \underline{u}$  for  $i \in \{+, -\}$ . Therefore, the incentive constraint

$$\pi_H(p)u^+ + (1 - \pi_H(p))u^- \geq \underline{u} + \frac{1 - \delta}{\delta}c$$

becomes equivalent to (5).

Hence, it remains to show that when  $T_e S^W(p) \geq \underline{s}$ , and, in the case of  $e = H$ , (5) holds, there exist payoffs in  $BW(p)$  decomposed by effort  $e$ . This can be achieved by setting continuation payoffs  $v^i = \underline{v}$  and  $u^i = S(\phi_e^i(p)) - \underline{v}$  for all  $i$  and adjusting the wage

to ensure that each player receives at least his outside option. In the case of  $e = H$ , this composition of continuation payoffs implies that (5) is equivalent to the incentive constraint for effort. This completes the proof.  $\square$

Now we have all the ingredients to the proof of Lemma 1.

**Proof of Lemma 1.** Since  $\mathcal{E} \in \mathcal{D}$  and  $S \equiv S^{\mathcal{E}}$ , Lemma 5 and Proposition 8 imply  $S(p) = T_*S(p)$  for all  $p$ . Thus,  $S$  is a fixed point of  $T_*$ .

Let  $f : [0, 1] \rightarrow \mathbb{R}$  be a bounded function satisfying  $f \geq S$  and  $f \geq T_*f$  and  $W_0(p) = \{(u, v) \geq (\underline{u}, \underline{v}) | u + v \leq f(p)\}$  for all  $p$ . Hence,  $W_0 \in \mathcal{D}$ . Lemma 5 and  $f \geq T_*f$  imply  $W_0(p) \supseteq \mathcal{B}W_0(p)$ , whereas  $f \geq S$  implies that  $W_0(p) \supseteq \mathcal{E}(p)$  for all  $p$ . Thus,  $W_0$  satisfies the conditions in Proposition 9. In the notation of Proposition 9, we have  $W_n \in \mathcal{D}$  for all  $n$  by an inductive application of Lemma 3. Thus,  $S^{W_n}(p) = T_*^n S^{W_0}(p)$  for all  $n$  and  $p$  by Lemma 5. Since  $\mathcal{E} \in \mathcal{D}$  and  $W_n(p)$  converges to  $\mathcal{E}(p)$  for all  $p$ , we have that  $\lim_{n \rightarrow \infty} S^{W_n}(p) = S(p)$  for every  $p$ . Moreover,  $W_n(p) \supseteq W_{n+1}(p)$  implies  $T_*^n f \geq T_*^{n+1} f$ .

Finally, we can take  $f = G$  since the equilibrium cannot improve upon the efficient outcome ( $G \geq S$ ) and  $G = \max\{T_H G, T_L G, \underline{s}\} \geq T_*G$ .  $\square$

## Proofs from Section 5.1

**Proof of Proposition 2.** Rearranging

$$G(p^{\text{FB}}) = \underline{s} = (1 - \delta)(R_H(p^{\text{FB}}) - c) + \delta \mathbb{E}^H[G(p^{\text{FB}})]$$

we obtain

$$(1 - \delta)c = (1 - \delta)(R_H(p^{\text{FB}}) - \underline{s}) + \delta(\mathbb{E}^H[G(p^{\text{FB}})] - \underline{s}). \quad (6)$$

Consider the application of  $T_*$  to  $G$ . If  $R_H(p^{\text{FB}}) \leq \underline{s}$ , (6) implies

$$\delta(\mathbb{E}^H[G(p^{\text{FB}})] - \underline{s}) \geq (1 - \delta)c$$

so the incentive constraint in the definition of  $T_*$  is satisfied at all beliefs  $p \geq p^{\text{FB}}$ . Hence,  $T_*G(p) = \max\{T_H G(p), T_L G(p), \underline{s}\} = G(p)$  for all  $p \geq p^{\text{FB}}$ . Moreover,  $\underline{s} \leq T_*G(p) \leq G(p) = \underline{s}$  for all  $p \leq p^{\text{FB}}$ . Hence,  $T_*G = G$ . The algorithmic part of Lemma 1 implies  $S = G$  so relational contracts are not inefficient.

If, instead,  $\pi_H(p^{\text{FB}}) > \underline{s}$ , (6) implies

$$\delta(\mathbb{E}^H[G(p^{\text{FB}})] - \underline{s}) < (1 - \delta)c.$$

By continuity of  $G$  and the belief updates  $\phi_H^+$  and  $\phi_L^+$ , the above strict inequality holds for some  $p > p^{\text{FB}}$ , at which the incentive constraint for  $H$  is not satisfied. Then  $T_*G(p) = \max\{T_LG(p), \underline{s}\}$ . Since  $G(p) = T_HG(p) > T_LG(p)$ , Lemma 1 implies  $S(p) < G(p)$ . Hence, relational contracts are inefficient.  $\square$

## Proofs from Section 5.2

To prove Proposition 3 we first establish some properties of  $S$  similar to, but weaker than, the properties of  $G$  in Proposition 1.

**Lemma 6.**  *$S$  is increasing and upper semicontinuous with  $S(0) = \underline{s}$  and  $S(1) = R_H(1) - c$ . There exists a unique belief  $p^* \in (0, 1]$  such that  $S(p) = \underline{s}$  for  $p < p^*$  and  $S(p) > \underline{s}$  for  $p > p^*$ .*

*Proof.* The upper semicontinuity of  $S$  follows directly from the upper hemicontinuity of  $\mathcal{E}$  established in Corollary 2. We now turn to establishing that  $S$  is increasing. Since  $G$  is increasing, Lemma 1 implies it is sufficient to show that  $T_*$  maps increasing functions to increasing functions. Let  $f$  be one such function. Notice that the incentive constraint for  $H$  in the definition of  $T_*f$  can be rewritten as  $h(p) \geq A$  where  $A$  is a constant and  $h$  is increasing and usc. Moreover, other things equal,  $T_*f$  is larger if the incentive constraint is satisfied. Hence, it is sufficient to show  $T_Hf$  and  $T_Lf$  are increasing and usc. This follows directly from the fact that  $f$  has the same properties.

The algorithmic part of Lemma 1 also implies that the values of  $S$  at 0 and 1 can be found by showing that  $T_*G(0) = \underline{s}$  and  $T_*G(1) = R_H(1) - c$ . The former follows from  $\underline{s} \leq T_*G(0) \leq TG(0) = G(0) = \underline{s}$ , while the latter is a consequence of Assumption 3.

Since  $R_H(1) - c > \underline{s}$ , the monotonicity of  $S$  implies there exists a maximal belief  $p^*$  such that  $S(p) = \underline{s}$  if  $p < p^*$  and  $S(p) = \max\{T_HS(p), T_LS(p)\}$  if  $p \geq p^*$ . That  $p^*$  is unique follows from the strict monotonicity of  $T_HS$ ,  $T_LS$  and  $h$ .

$\square$

Next, we establish that the three-region structure in Proposition 3 is obtained if the surplus function exhibits a monotone policy.

**Lemma 7.** *If  $S$  exhibits a monotone policy there exist beliefs  $\underline{p}, \bar{p}$  such that*

$$S(p) = \begin{cases} \underline{s} & \text{if } p < \underline{p} \\ T_L S(p) & \text{if } \underline{p} \leq p < \bar{p} \\ T_H S(p) & \text{if } p \geq \bar{p} \end{cases}$$

*Proof.* By Lemma 6 there is a threshold belief  $p^*$  such that the players take their outside options at beliefs  $p < p^*$  and interact if  $p \geq p^*$ . The monotonicity of  $S$  implies that if the incentive constraint is satisfied at some belief it is satisfied for all higher beliefs. Let  $p^{\text{IC}}$  be the minimum belief at which the incentive constraint for  $H$  is satisfied. Note that such a belief exists due to  $S(1) = R_H(1) - c$  and the infimum is attained due to the upper semicontinuity of  $S$ . Hence,  $T_* f(p) = T_H f(p)$  iff  $p \geq \max\{p^*, p^{\text{IC}}\}$ . Putting  $\underline{p} = p^*$  and  $\bar{p} = \max\{p^*, p^{\text{IC}}\}$  completes the proof.  $\square$

The final piece of the proof of Proposition 3 is the following lemma which ensures that the monotone policy property is preserved under  $T_*$ .

**Lemma 8.** *Let  $f$  be an increasing function that exhibits a monotone policy. Then  $T_* f$  is increasing and exhibits a monotone policy.*

*Proof.* The proof of Lemma 6 showed that  $T_* f$  is increasing when  $f$  is increasing. Hence, there exists a maximal belief  $p^*$  such that  $T_* f(p) = \underline{s} \forall p < p^*$ . It remains to show that  $T_* f$  exhibits a monotone policy. Let  $p \in [0, 1]$ . We will show that  $T_H(T_* f)(p) > T_L(T_* f)(p)$ . We examine two cases depending on where the posterior beliefs lie relative to the cutoff  $p^*$ .

Firstly, if  $p^* > \phi_L^-(p)$ , we have  $T_* f(\phi_L^-(p)) = \underline{s}$ . Also,  $\phi_L^- \geq \phi_H^-$  implies that  $T_* f(\phi_H^-(p)) = \underline{s}$  since  $T_* f$  is increasing. Hence,

$$\begin{aligned} & T_H(T_* f)(p) - T_L(T_* f)(p) \\ & \geq (1 - \delta)(R_H(p) - c) + \delta[\pi_H(p)T_* f(\phi_H^+(p)) + (1 - \pi_H(p))\underline{s}] \\ & \quad - (1 - \delta)R_L(p) - \delta[\pi_L(p)T_* f(\phi_L^+(p)) + (1 - \pi_L(p))\underline{s}] \\ & \geq (1 - \delta)(R_H(p) - c - R_L(p)) > 0. \end{aligned}$$

where the second inequality follows from  $R_H(p) > R_L(p)$  and  $T_* f(\phi_H^+(p)) \geq T_* f(\phi_L^+(p)) \geq \underline{s}$ .

Secondly, suppose  $p^* \leq \phi_L^-(p)$ . Then  $T_* f(q') \in \{T_H f(q'), T_L f(q')\}$  for any  $q' \in \{\phi_L^+(p), \phi_L^-(p)\}$  which implies  $T_* f(q') \leq T_H f(q')$  since  $f$  exhibits a monotone policy.

Thus,  $T_L(T_*f)(p) \leq T_L(T_Hf)(p)$  and it follows that

$$\begin{aligned}
& T_H(T_*f)(p) - T_L(T_*f)(p) \\
& \geq T_H(T_Lf)(p) - T_L(T_Hf)(p) \\
& = (1 - \delta)(R_H(p) - c) + \delta(1 - \delta)(\mathbb{E}^H[R_L(p)]) + \delta^2\mathbb{E}^{H,L}[f(p)] \\
& \quad - \left( (1 - \delta)(R_L(p)) + \delta(1 - \delta)(\mathbb{E}^L[R_H(p) - c]) + \delta^2\mathbb{E}^{L,H}[f(p)] \right) \\
& = (1 - \delta)(R_H(p) - c) + \delta(1 - \delta)R_L(p) \\
& \quad - \left( (1 - \delta)(R_L(p) - c) + \delta(1 - \delta)(R_H(p) - c) \right) \\
& = (1 - \delta)^2(R_H(p) - c - R_L(p)) > 0
\end{aligned}$$

where

$$\mathbb{E}^{e_1, e_2}[f(p)] = \pi_{e_1}(p)\mathbb{E}^{e_2}[f(\pi_{e_1}^+(p))] + (1 - \pi_{e_1}(p))\mathbb{E}^{e_2}[f(\pi_{e_1}^-(p))]$$

for any  $e_1, e_2 \in \{H, L\}$ . This completes the proof.  $\square$

**Proof of Proposition 3.** We know from Lemma 1 that  $S = \lim_{n \rightarrow \infty} G_n$  where  $G_n = T_*^n G$ . Since  $G$  is increasing and exhibits a monotone policy (Proposition 1), Lemma 8 implies that  $G_n$  exhibits a monotone policy for all  $n$ . Taking limits yields  $T_H S(p) \geq T_L S(p)$  for all  $p$ , so  $S$  exhibits a monotone policy. Lemma 7 completes the proof.  $\square$

**Proof of Proposition 4.** Suppose, for the sake of contradiction, that  $\underline{p} = \bar{p}$ . Consider some belief  $p = \bar{p} - \varepsilon$  for some small  $\varepsilon > 0$ . It suffices to show that  $L$  generates more surplus than the outside options.

We know that  $S(\phi_L^+(p)) \geq S(\bar{p})$  whenever  $\varepsilon$  is small enough. So the joint surplus from  $L$  at belief  $p$  is at least

$$\begin{aligned}
& (1 - \delta)R_L(p) + \delta \left[ \pi_L(p)S(\bar{p}) + (1 - \pi_L(p))\underline{s} \right] \\
& \geq (1 - \delta)R_L(p) + \delta \left[ \pi_L(p) \left( (1 - \delta)R_H(\bar{p}) + \delta\underline{s} \right) + (1 - \pi_L(p))\underline{s} \right] \\
& = (1 - \delta)R_L(p) + \delta \left[ \pi_L(p)(1 - \delta) \left( R_H(\bar{p}) - \underline{s} \right) + \underline{s} \right] \\
& = \underline{s} + (1 - \delta) \left[ R_L(p) - \underline{s} + \delta\pi_L(p)(R_H(\bar{p}) - \underline{s}) \right]
\end{aligned}$$



This exceeds  $\underline{s}$  whenever

$$R_L(0) - \underline{s} + \delta\pi_L(0)(R_H(0) - \underline{s}) \geq 0$$

as required.  $\square$

## Proofs from Section 6

*Proof of Theorem 1.* Any limit in this proof is taken as  $k \rightarrow \infty$ .

We begin with some preliminaries regarding the efficient solution, which are a direct consequence of Proposition 1. For any  $k$ , define  $G_k^+(p) := G_k(\phi_{H,k}^+(p))$  and  $G_k^-(p) := G_k(\phi_{H,k}^-(p))$ . We have that  $p \geq p_k^{\text{FB}}$  if and only if

$$(1 - \delta_k)(R_H(p) - c) + \delta_k \left[ \pi_H^k(p)G_k^+(p) + (1 - \pi_H^k(p))\underline{s} \right] \geq \underline{s}$$

which can be rearranged as

$$\frac{\delta_k \pi_H^k(p)}{(1 - \delta_k)} (G_k^+(p) - \underline{s}) \geq \underline{s} - (R_H(p) - c) \quad (7)$$

and (7) holds at equality if  $p = p_k^{\text{FB}}$ . Since  $G_k$  is increasing, convex and  $G_k(p) \geq R_H(p) - c$ , each subgradient of  $G_k$  at any  $p$  is bounded above by  $R_H(1) - R_H(0)$ . Hence, the family  $(G_k)$  is equicontinuous and, by the Arzela-Ascoli theorem, it has a subsequence which converges uniformly. In what follows, we will assume that this is the original subsequence and use  $G$  to denote the limiting function.<sup>23</sup> In addition, notice that  $R_H(1) - c > \underline{s}$  and  $G_k(p) \geq R_H(p) - c$  for all  $p, k$  imply that  $p_*^{\text{FB}} < 1$  and  $G(p) > \underline{s}$  for some  $p < 1$ .

In the optimal relational contract, the following incentive constraint holds:

$$\delta_k \left( \pi_H^k(\bar{p}_k) (S_k^+(\bar{p}_k) - \underline{s}) + (1 - \pi_H^k(\bar{p}_k)) (S_k^-(\bar{p}_k) - \underline{s}) \right) \geq (1 - \delta_k)c$$

where  $S_k^+(p) := S_k(\phi_{H,k}^+(p))$  and  $S_k^-(p) := S_k(\phi_{H,k}^-(p))$ . A sufficient condition is

$$\delta_k \pi_H^k(\bar{p}_k) \left( S_k^+(\bar{p}_k) - \underline{s} \right) \geq (1 - \delta_k)c. \quad (8)$$

Lemma 6 and Helly's selection theorem imply that  $(S_k)$  has a subsequence which con-

<sup>23</sup>This is without loss of generality, since the only properties of the sequence we will need are  $p_k^{\text{FB}} \rightarrow p_*^{\text{FB}}$  and  $\bar{p}_k \rightarrow \bar{p}_*$ , which hold at any subsequence.

verges pointwise to some function  $S$ . Without loss of generality, we take the subsequence to be the original sequence.

We continue by describing the relationship between both parts of Assumption 4 and  $\lim \frac{\text{Var}_H^k(p)}{1-\delta_k}$ . The variance of the posterior at prior  $p$  can be expressed as

$$\text{Var}_H^k(p) = \left(p(1-p)\right)^2 \frac{(\gamma_H^k - \beta_H^k)^2}{\pi_H^k(p)(1 - \pi_H^k(p))}.$$

where

$$\frac{\gamma_H^k - \beta_H^k}{\pi_H^k(p)} = \frac{1}{p + \frac{1}{\frac{\gamma_H^k}{\beta_H^k} - 1}} \quad \text{and} \quad \frac{\gamma_H^k - \beta_H^k}{1 - \pi_H^k(p)} = \frac{1}{\frac{1}{1 - \gamma_H^k} - p}. \quad (9)$$

Suppose both parts of Assumption 4 hold. Then the limits of  $\frac{\gamma_H^k - \beta_H^k}{\pi_H^k(p)}$  and  $\frac{\gamma_H^k - \beta_H^k}{1 - \pi_H^k(p)}$  are finite for all interior  $p$ . Hence  $\lim \frac{\text{Var}_H^k(p)}{1-\delta_k} = \infty$ . In addition, Bayesian updating necessitates that  $0 < \liminf \pi_H^k(p) \leq \limsup \pi_H^k(p) < 1$  for all  $p$ .

Suppose Assumption 4(a), holds, but Assumption 4(b) does not. Then (9) implies that  $\frac{\gamma_H^k - \beta_H^k}{\pi_H^k(p)}$  converges to a finite limit for all interior  $p$ . Moreover, Bayes rule implies  $\pi_H^k(p) \rightarrow 0$  for all interior  $p$ , which implies  $\beta_H^k \rightarrow 0$ . Hence,  $\lim \frac{\text{Var}_H^k(p)}{1-\delta_k}$  is finite for all interior  $p$  iff the limit of

$$\frac{\gamma_H^k - \beta_H^k}{1 - \delta_k} = \frac{\gamma_H^k}{1 - \delta_k} \left(1 - \frac{\beta_H^k}{\gamma_H^k}\right) = \frac{1 - \beta_H^k}{1 - \delta_k} \left(1 - \frac{1 - \gamma_H^k}{1 - \beta_H^k}\right) \quad (10)$$

is finite. Let  $\pi^* = \lim \frac{\gamma_H^k}{1-\delta_k}$ . Hence,  $\lim \frac{\text{Var}_H^k(p)}{1-\delta_k}$  is finite for all interior  $p$  iff  $\pi^*$  is finite iff  $\lim \frac{\pi_H^k(p)}{1-\delta_k}$  is finite for all  $p$ .

If Assumption 4(b) holds but Assumption 4(a) does not hold, (9) implies that  $\frac{\gamma_H^k - \beta_H^k}{1 - \pi_H^k(p)}$  converges to a finite limit for all interior  $p$ . Bayes rule implies  $\pi_H^k(p) \rightarrow 1$  for all  $p$  and  $\beta_H^k \rightarrow 1$ . Hence,  $\lim \frac{\text{Var}_H^k(p)}{1-\delta_k}$  is finite iff  $\lim \frac{\gamma_H^k - \beta_H^k}{1-\delta_k}$  is finite iff  $\lim \frac{1-\beta_H^k}{1-\delta_k}$  is finite (see (10)).

We proceed to give separate proofs for the cases when parts (a) and (b) of Assumption 4 hold.

**Proof of Theorem 1 under Assumption 4(a).** Suppose Assumption 4(a) holds. The preliminary arguments establish that, if Assumption 4(b) holds,  $\lim \frac{\text{Var}_H^k(p)}{1-\delta_k} = \infty$  for all interior  $p$  and  $\pi^* = \infty$  since the probability of success is bounded below. If Assumption 4(b) does not hold, (i) holds iff  $\pi^*$  is finite. Hence, we conclude that (i) holds iff  $\pi^*$  is finite,

regardless of whether Assumption 4(b) holds.

(i)  $\Rightarrow$  (ii), (iii): Suppose  $\pi^*$  is finite and  $p_*^{\text{FB}} = 0$ . Then the RHS of (7) converges to  $\underline{s} - R_H(0) + c$ , which is positive by Assumption 1. Hence, when  $\pi^*$  is finite, we can show  $p_*^{\text{FB}} > 0$  by showing that the LHS of (7) at  $p_k^{\text{FB}}$  converges to 0. This implies  $\bar{p}_* > 0$ .

If  $\lim \frac{\gamma_H^k}{\beta_H^k} < \infty$ ,  $\phi_{H,k}^+(p_k^{\text{FB}}) \rightarrow 0$ , and, by the uniform convergence of  $(G_k)$ , we have  $G_k^+(p_k^{\text{FB}}) \rightarrow \underline{s}$ . Since  $\pi^*$  is finite, this implies that the LHS of (7) at  $p_k^{\text{FB}}$  converges to 0.

If  $\lim \frac{\gamma_H^k}{\beta_H^k} = \infty$ , the finiteness of  $\pi^*$  implies that  $\lim \frac{\beta_H^k}{1-\delta_k} = 0$ . Thus,  $\frac{\pi_H^k(p)}{1-\delta_k} \rightarrow p\pi^*$  for all  $p$ . It follows that the LHS of (7) converges to 0, as required.

(i)  $\Rightarrow$  (iv): Suppose  $\pi^*$  is finite,  $R_H(p_*^{\text{FB}}) > \underline{s}$ , and  $\bar{p}_* = p_*^{\text{FB}}$ . There exist  $\varepsilon > 0$  and  $K_1$  such that

$$\begin{aligned} \frac{\delta_k \pi_L^k(p_*^{\text{FB}} + \varepsilon)}{1 - \delta_k} \left( S_k(\phi_{L,k}^+(p_*^{\text{FB}} + \varepsilon)) - \underline{s} \right) &\leq \frac{\delta_k \pi_H^k(p_*^{\text{FB}} + \varepsilon)}{1 - \delta_k} \left( G_k(\phi_{L,k}^+(p_*^{\text{FB}} + \varepsilon)) - \underline{s} \right) \\ &\leq \lim \delta_k \frac{\pi_H^k(p_k^{\text{FB}})}{1 - \delta_k} \left( G_k(\phi_{L,k}^+(p_k^{\text{FB}})) - \underline{s} \right) + \varepsilon \\ &\leq \underline{s} - (R_H(p_*^{\text{FB}}) - c) + \varepsilon \\ &< \underline{s} - R_L(p_*^{\text{FB}} + \varepsilon). \end{aligned}$$

for all  $k \geq K_1$ . The first inequality follows from  $S_k \leq G_k$  and Assumption 1. The second inequality holds for any  $\varepsilon > 0$  small enough due to three limit results:  $G_k(\phi_{L,k}^+(p_k^{\text{FB}})) \rightarrow G(\phi_{L,k}^+(p_*^{\text{FB}}))$ ,  $\lim \frac{\pi_H^k(p_k^{\text{FB}})}{1-\delta_k} = \lim \frac{\pi_H^k(p_*^{\text{FB}})}{1-\delta_k} < \infty$ , and  $\delta_k \rightarrow 1$ . The first follows from the uniform convergence of  $\phi_{L,k}^+$ , and the uniform convergence of  $(G_k)$  to  $G$ . Towards the second limit, notice that

$$\pi_H^k(p) = \beta_H^k \left( 1 - p + p \frac{\gamma_H^k}{\beta_H^k} \right)$$

so Assumption 4(a) implies that  $\lim_k \frac{\pi_H^k(p_k^{\text{FB}})}{\pi_H^k(p_*^{\text{FB}})} = 1$ . Since  $\pi^*$  is finite,  $\lim_k \frac{\pi_H^k(p_*^{\text{FB}})}{1-\delta_k}$  exists, giving the desired result. The penultimate equality follows from (7) and  $\phi_{L,k}^+ \leq \phi_{H,k}^+$ . The final inequality is a result of Assumption 1 for an appropriate choice of  $\varepsilon$ .

The resulting inequality can be rearranged to show

$$\begin{aligned} (1 - \delta_k) \left( R_L(p) - \underline{s} \right) + \delta_k \left( \pi_L^k(p) S_k(\phi_{L,k}^+(p)) + (1 - \pi_L^k(p)) \underline{s} \right) &< \underline{s} \\ \text{for all } p \leq p_*^{\text{FB}} + \varepsilon, k \geq K_1. &\quad (11) \end{aligned}$$

Let  $K_2$  be such that  $\bar{p}_k < p_*^{\text{FB}} + \varepsilon$  for all  $k \geq K_2$  and let  $K = \max\{K_1, K_2\}$ . Suppose that there is a region where  $L$  is used in an optimal relational contract for some  $k \geq K$ . By Proposition 3 there exists  $p \leq p_*^{\text{FB}} + \varepsilon$ , at which the surplus from  $L$  is no less than  $\underline{s}$  and  $S_k(\phi_{L,k}^-(p)) = \underline{s}$ . This is a contradiction to (11). Hence, for all  $k \geq K$ , the optimal relational contract does not use  $L$  and the incentive constraint implies

$$\lim_{k \rightarrow \infty} \frac{\delta_k \pi_H^k(\bar{p}_k)}{1 - \delta_k} (G_k^+(\bar{p}_k) - \underline{s}) \geq c \quad (12)$$

Since  $\pi^*$  is finite,  $\frac{\delta_k \pi_H^k(\bar{p}_k)}{1 - \delta_k}$  converges to a finite number. Hence, the uniform convergence of  $(G_k)$  and (7) imply that the LHS of (12) converges to  $\underline{s} - (R_H(p_*^{\text{FB}}) - c)$ . A contradiction to  $R_H(p_*^{\text{FB}}) > \underline{s}$ . Hence, we conclude that  $p_*^{\text{FB}} < \bar{p}_*$ .

(ii)  $\Rightarrow$  (i): Suppose  $\pi^* = \infty$ . We want to show that  $p_*^{\text{FB}} = 0$ . Let  $\hat{p} = \sup\{p | G(p) = \underline{s}\} < 1$ . Suppose for a contradiction, that  $\hat{p} > 0$ . By Assumption 4 there exists  $\varepsilon > 0$  such that  $\phi_{H,k}^+(\hat{p} - \varepsilon) \geq \hat{p} + \varepsilon$  for all  $k$  large enough. Hence,  $G_k^+(\hat{p} - \varepsilon) \geq \underline{s} + \eta$  for some  $\eta > 0$ . Then (7) implies  $p_k^{\text{FB}} \leq \hat{p} - \varepsilon$  for high  $k$  and it follows that

$$G_k(p) = (1 - \delta_k)(R_H(p) - c) + \delta_k \left[ \pi_k(p) G_k^+(p) + (1 - \pi_k(p)) G_k^-(p) \right]$$

$$\frac{G_k(p) - G_k^-(p)}{p - \phi_{H,k}^-(p)} = \frac{1 - \delta_k}{p - \phi_{H,k}^-(p)} (R_H(p) - c - G_k^-(p)) + \delta_k \frac{\pi_H^k(p)}{p - \phi_{H,k}^-(p)} (G_k^+(p) - G_k^-(p))$$

for any  $p \in (\hat{p} - \varepsilon, \hat{p})$ . Notice that

$$\frac{\pi_H^k(p)}{p - \phi_{H,k}^-(p)} = \frac{\pi_H^k(p)(1 - \pi_H^k(p))}{p(1 - p)(\gamma_H^k - \beta_H^k)}.$$

Assumption 4(a) and (9) imply that  $\lim_{k \rightarrow \infty} \frac{\pi_H^k(p)}{\gamma_H^k - \beta_H^k}$  is positive and finite. The preliminaries of the proof show that if Assumption 4(b) holds, then  $1 - \pi_H^k(p)$  is bounded away from 0. If Assumption 4(b) does not hold,  $1 - \pi_H^k(p) \rightarrow 1$  for all interior  $p$ . Hence, there exists a constant  $\zeta > 0$  such that

$$\frac{1}{p(1 - p)} \zeta \leq \frac{\pi_H^k(p)}{p - \phi_{H,k}^-(p)} \leq \frac{1}{p(1 - p)} \frac{1}{\zeta}$$

for all interior  $p$ . Hence,  $\pi^* = \infty$  implies

$$\lim \frac{1 - \delta_k}{p - \phi_{H,k}^-(p)} = \lim \frac{1 - \delta_k}{\pi_H^k(p)} \frac{\pi_H^k(p)}{p - \phi_{H,k}^-(p)} = 0$$

for all  $p \in (\hat{p} - \varepsilon, \hat{p})$ . Since  $R_H(p) - c - G_k^-(p)$  is uniformly bounded, we have

$$\lim \frac{G_k(p) - G_k^-(p)}{p - \phi_k^-(p)} = \lim \delta_k \frac{\pi_H^k(p)}{p - \phi_k^-(p)} (G_k^+(p) - G_k^-(p))$$

Hence, there exists  $K$  such that for all  $k \geq K$  and  $p \in (\hat{p} - \varepsilon, \hat{p})$  we have

$$\frac{G_k(p) - G_k^-(p)}{p - \phi_k^-(p)} > \frac{1}{2} \zeta \eta \min_{p' \in [\hat{p} - \varepsilon, \hat{p}]} \frac{1}{p'(1 - p')}.$$

Consequently, there exists a constant  $\mu > 0$  such that  $G_k(\hat{p} - \frac{\varepsilon}{2}) > \underline{s} + \mu$  for all  $k \geq K$ , in contradiction to  $G(\hat{p} - \frac{\varepsilon}{2}) = \underline{s}$ .

We conclude that  $\hat{p} = 0$ . Hence, for any  $p > 0$ , there exists  $K$  such that  $G_k(p) > \underline{s}$  for all  $k \geq K$ . It follows that  $p_k^{\text{FB}} \rightarrow 0$ , as required.

(iii)  $\Rightarrow$  (i): Suppose  $\pi^* = \infty$ . We want to show that  $\bar{p}_* = 0$ . The strategy follows the argument for (ii)  $\Rightarrow$  (i) with some modifications.

Let  $\hat{p} = \sup\{p | S(p) = \underline{s}\}$ . Our first goal is to show that  $\hat{p} < 1$ . Let

$$f(p) = \begin{cases} \underline{s} & \text{if } p < p_1 \\ \underline{s} + (R_H(1) - c) \frac{p - p_1}{1 - p_1} & \text{if } p \geq p_1 \end{cases}$$

where  $p_1$  is large enough to ensure  $R_H(p) - c \geq f(p)$  for all  $p \geq p_1$ . Let  $T_*^k$  denote the operator  $T_*$  within model  $\mathcal{M}_k$ . Consider the function  $T_*^k f$  at belief  $p_1$ . Assumption 4(a) implies there exists a constant  $\eta > 0$  such that  $S_k^+(p_1) \geq \underline{s} + \eta$  for  $k$  large enough. Hence,  $\pi^* = \infty$  and (8) imply that the incentive constraint holds at all  $p \geq p_1$  for high  $k$ . The convexity of  $f$  implies that  $\mathbb{E}_k^H[f(p)] \geq f(p)$  for all such  $k$  and  $p$ , where  $\mathbb{E}_k^H$  denotes the  $\mathbb{E}^H$  operator in model  $\mathcal{M}_k$ . Hence,

$$T_*^k f(p) \geq (1 - \delta_k)(R_H(p) - c) + \delta_k \mathbb{E}_k^H[f(p)] \geq f(p).$$

for all  $p \geq p_1$ . It follows that  $T_*^k f(p) \geq f(p)$  for all  $p \in [0, 1]$  and  $k$  large enough. Hence,

the correspondence

$$W^f(p) = \{(u, v) | (u, v) \geq (\underline{u}, \underline{v}), u + v \leq f(p)\} \quad \text{for all } p$$

is self-generating and Proposition 7 implies that  $S_k \geq f$  for all  $k$  large enough. Hence,  $\hat{p} < 1$ .

Suppose that  $\hat{p} > 0$ . Define  $\varepsilon$  and  $\eta$  as in the proof of  $(ii) \Rightarrow (i)$ , so that  $S_k^+(\hat{p} - \varepsilon) \geq \underline{s} + \eta$ . Since  $\pi^* = \infty$ , the incentive constraint (8) holds for all  $k$  large enough. It follows that  $\bar{p}_k \leq \hat{p} - \varepsilon$  for high  $k$ . Replicating the remaining part of the argument for  $(ii) \Rightarrow (i)$  shows that we must have  $\hat{p} = 0$ . It follows that  $\bar{p}_* = 0$ , as required.

$(iv) \Rightarrow (i)$ : Consider a sequence of models where all parameters, except for  $\delta_k$  are constant and  $R_H(0) > \underline{s}$ . Then  $\pi^* = \infty$ , which implies  $\bar{p}_* = p_*^{\text{FB}}$ , but  $R_H(p_*^{\text{FB}}) = \underline{s}$ . Hence  $\pi^* = \infty$  implies that statement (iv) is false, as required.

**Proof of Theorem 1 under Assumption 4(b).** Suppose Assumption 4(b) holds. In light of the previous part of the proof, we will assume without loss of generality that Assumption 4(a) does not hold. Then  $\pi_H^k(p) \rightarrow 1$  for all  $p$ . Let  $\lambda = \lim \frac{\gamma_H^k - \beta_H^k}{1 - \delta_k}$ . It follows from the preliminary discussion that (i) holds iff  $\lambda$  is finite iff  $\lim \frac{1 - \beta_H^k}{1 - \delta_k}$  is finite. Notice that Assumption 4(b) implies that  $\lim \frac{1 - \gamma_H^k}{1 - \delta_k}$  is finite whenever  $\lim \frac{1 - \beta_H^k}{1 - \delta_k}$  is finite.

$(i) \Rightarrow (ii), (iii)$ : Suppose  $\lambda$  is finite and  $p_*^{\text{FB}} = 0$ . Since

$$\phi_{H,k}^+(p) - p = p(1 - p) \frac{\gamma_H^k - \beta_H^k}{\pi_H^k(p)}$$

and the bound on the subgradient of  $G_k$  implies that

$$\frac{G_k^+(p_k^{\text{FB}}) - \underline{s}}{\phi_{H,k}^+(p_k^{\text{FB}}) - p_k^{\text{FB}}} \leq R_H(1) - R_H(0)$$

for all  $k$ , it follows that

$$\begin{aligned} \frac{\delta_k \pi_H^k(p)}{1 - \delta_k} \left( G_k^+(p_k^{\text{FB}}) - \underline{s} \right) &\leq \left( R_H(1) - R_H(0) \right) \delta_k p_k^{\text{FB}} (1 - p_k^{\text{FB}}) \frac{\gamma_H^k - \beta_H^k}{1 - \delta_k} \\ &\leq \left( R_H(1) - R_H(0) \right) (\lambda + \varepsilon) p_k^{\text{FB}} (1 - p_k^{\text{FB}}) \end{aligned}$$

for large  $k$ . Hence,  $p_*^{\text{FB}} = 0$  implies that the LHS of (7) at  $p_k^{\text{FB}}$  converges to 0, while the RHS converges to  $\underline{s} - (R_H(0) - c) > 0$ . A contradiction. Hence,  $p_*^{\text{FB}} > 0$  and  $\bar{p}_* > 0$ ,

whenever  $\lambda$  is finite.

(i)  $\Rightarrow$  (iv): Suppose  $p_*^{\text{FB}} = \bar{p}_* = p_*$  and  $R_H(p_*) > \underline{s}$ . There exists  $\hat{p} > p_*$  such that  $\phi_{L,k}^-(p) < p_*$  for all  $p \in [p_*, \hat{p}]$ . Hence, for  $k$  high enough, the incentive constraint at  $\bar{p}_k$  is given by

$$\frac{\delta_k \pi_H^k(\bar{p}_k)}{1 - \delta_k} (S_k^+(\bar{p}_k) - \underline{s}) \geq c \quad (13)$$

For any  $p \geq \bar{p}_k$ , we have

$$\begin{aligned} S(p) &= (1 - \delta_k)(R_H(p) - c) + \delta_k \left[ \pi_H^k(p) S_k^+(p) + (1 - \pi_H^k(p)) S_k^-(p) \right] \\ &= S_k^+(p) - (1 - \pi_H^k(p))(S_k^+(p) - S_k^-(p)) \\ &\quad + (1 - \delta_k) \left( (R_H(p) - c) - \left( \pi_H^k(p) S_k^+(p) + (1 - \pi_H^k(p)) S_k^-(p) \right) \right) \end{aligned}$$

Hence, the following holds for all  $k$  high enough:

$$\begin{aligned} \frac{S_k^+(\bar{p}_k) - \underline{s}}{1 - \delta_k} &= \frac{1 - \pi_H^k(\bar{p}_k)}{1 - \delta_k} (S_k^+(\bar{p}_k) - \underline{s}) \\ &\quad - (R_H(\bar{p}_k) - c) + \left( \pi_H^k(\bar{p}_k) S_k^+(\bar{p}_k) + (1 - \pi_H^k(\bar{p}_k)) \underline{s} \right). \end{aligned} \quad (14)$$

Since Assumption 4(b) does not hold,  $\phi_{H,k}^+(p) \rightarrow p$  for all  $p$ . Hence, the uniform convergence of  $G$  implies that  $G_k^+(\bar{p}_k) \rightarrow \underline{s}$ . It follows that  $S_k^+(\bar{p}_k) \rightarrow \underline{s}$ . Since  $\lim \frac{1 - \beta_H^k}{1 - \delta_k}$  and  $\lim \frac{1 - \gamma_H^k}{1 - \delta_k}$  are finite,  $\lim \frac{1 - \pi_H^k(\bar{p}_k)}{1 - \delta_k}$  is finite. Hence, taking limits of (14), we obtain

$$\lim \delta_k \pi_H^k(\bar{p}_k) \frac{S_k^+(\bar{p}_k) - \underline{s}}{1 - \delta_k} = \lim \frac{S_k^+(\bar{p}_k) - \underline{s}}{1 - \delta_k} = \underline{s} - (R_H(p_*) - c) < c$$

which contradicts the limit of the incentive constraint (13).

(ii)  $\Rightarrow$  (i): Suppose  $\lambda = \infty$ . Let  $\Delta_k$  satisfy  $\delta_k = e^{-\Delta_k}$ . L'Hospital's rule implies  $\frac{\gamma_H^k - \beta_H^k}{\Delta} \rightarrow \infty$ .

Let  $\hat{p} = \inf\{p | G(p) > \underline{s}\}$ . Suppose, for the sake of contradiction, that  $\hat{p} > 0$ . There exists  $\varepsilon$  such that for all  $k$  high enough  $G(\hat{p} + \varepsilon) \geq \underline{s} + \varepsilon$ . Suppose the belief is  $\hat{p} - \varepsilon$  and

we want to find the number of consecutive successes needed to reach  $\hat{p} + \varepsilon$ . Since

$$\phi_{H,k}^+(p) - p = \frac{p(1-p)}{\pi_H^k(p)}(\gamma_H^k - \beta_H^k)$$

this number will be at most

$$N(k) = \frac{2\varepsilon}{\gamma_H^k - \beta_H^k} \max_{p \in [\hat{p}-\varepsilon, \hat{p}+\varepsilon]} \frac{p(1-p)}{\pi_H^k(p)} = \frac{\xi}{\gamma_H^k - \beta_H^k}$$

for high  $k$ , where  $\xi$  is a positive constant (since  $\pi_H^k \rightarrow 1$ ). Consider the following policy at  $\hat{p} - \varepsilon$ : in the next  $N(k)$  periods, exert  $H$  and, if failure occurs, take the outside options in the following period. If  $N(k)$  successes occur, continue with the optimal policy at the posterior belief, which must generate surplus of at least  $\underline{s} + \varepsilon$ . For  $k$  high enough, it must be that the value of this policy does not exceed the value of the outside options. Hence,

$$\underline{s} \geq \delta_k^{N(k)} \left[ \hat{\pi}^k (\underline{s} + \varepsilon) + (1 - \hat{\pi}^k) \underline{s} \right] + (R_H(0) - c)(1 - \delta_k^{N(k)})$$

where  $\hat{\pi}^k$  is the probability of getting  $N(k)$  successes in a row, starting from  $\hat{p} - \varepsilon$ . The flow payoffs during the first  $N(k)$  periods are either  $R_H(p) - c$  when  $H$  is exerted, or  $\underline{s}$ , both of which are more valuable than  $R_H(0) - c$  by Assumption 1. Rearranging, we obtain

$$(1 - \delta_k^{N(k)}) \underline{s} \geq \delta_k^{N(k)} \hat{\pi}^k \varepsilon + (R_H(0) - c)(1 - \delta_k^{N(k)})$$

We will violate this inequality for high  $k$  by showing that the LHS converges to 0, while the RHS converges to a positive constant. This follows from two limiting results:  $\delta_k^{N(k)} \rightarrow 1$  and  $\liminf \hat{\pi}^k > 0$ . Towards the first, notice that

$$\delta_k^{N(k)} = \exp\left(-\frac{\Delta}{\gamma_H^k - \beta_H^k} \xi\right)$$

converges to 1 since  $\frac{\gamma_H^k - \beta_H^k}{\Delta} \rightarrow \infty$ . Towards the second result, notice that

$$\hat{\pi}^k \geq \left(\pi_H^k(\hat{p} - \varepsilon)\right)^{N(k)}$$



and let

$$\begin{aligned}
B_k &= \frac{\log \left[ \left( \pi_H^k(\hat{p} - \varepsilon) \right)^{N(k)} \right]}{\xi} = \frac{\log \left[ \beta_H^k + (\gamma_H^k - \beta_H^k)(\hat{p} - \varepsilon) \right]}{\gamma_H^k - \beta_H^k} \\
&= \frac{\log \left[ 1 - (1 - \beta_H^k) - \left( (1 - \gamma_H^k) - (1 - \beta_H^k) \right) (\hat{p} - \varepsilon) \right]}{(1 - \beta_H^k) - (1 - \gamma_H^k)}
\end{aligned}$$

Given any  $\varepsilon > 0$  small enough, we have

$$(1 - \beta_H^k)(\lambda - \varepsilon) \leq (1 - \gamma_H^k) \leq (1 - \beta_H^k)(\lambda + \varepsilon)$$

for all  $k$  high enough. Hence,

$$B_k \geq \frac{\log \left[ 1 - (1 - \beta_H^k)(1 - (\lambda + \varepsilon - 1)(\hat{p} - \varepsilon)) \right]}{(1 - \beta_H^k)(1 - \lambda + \varepsilon)}$$

Since  $\lim \beta_H^k = 1$ , L'Hospital's rule implies

$$\liminf B_k \geq -\frac{1 - (\lambda + \varepsilon - 1)(\hat{p} - \varepsilon)}{1 - \lambda + \varepsilon}$$

Thus, we conclude that

$$\liminf \hat{\pi}^k \geq \liminf \exp(B_k \xi) > 0$$

as required. Hence,  $p_*^{\text{FB}} = 0$ .

(iii)  $\Rightarrow$  (i): Suppose  $\lambda = \infty$ . Let  $\hat{p} = \inf\{p | S(p) > \underline{s}\}$ . We can show  $\hat{p} < 1$  as in the proof under Assumption 4(a). Suppose  $\hat{p} > 0$ . Define  $\varepsilon$  and  $N(k)$  as in the previous part of the proof and recall the strategies considered at  $\hat{p} - \varepsilon$ : the worker exerts  $H$  for  $N(k)$  periods, unless a failure occurs which is followed by the outside options. In the event of  $N(k)$  consecutive successes, an efficient continuation equilibrium is played with surplus of at least  $\underline{s} + \varepsilon$ . As previously shown, these strategies create higher surplus than the outside options for high  $k$ . Hence, it only remains to show that the incentive compatibility constraint is satisfied in each of the initial  $N(k)$  periods. Since  $H$  is exerted only if success occurred in all of the previous periods, the expected continuation surplus from  $H$  in excess

of  $\underline{s}$  in the  $n$ -th period is at least

$$\delta \pi_H^k(\hat{p} - \varepsilon) \left[ \delta_k^{N(k)-n} \left( \pi_H^k(\hat{p} - \varepsilon) \right)^{N(k)-n} \varepsilon + (1 - \delta_k^{N(k)-n})(R_H(0) - c - \underline{s}) \right]$$

which is strictly positive for  $k$  high enough. Hence, we conclude that  $\hat{p} = 0$ . Hence  $\bar{p}_* = 0$ .

(iv)  $\Rightarrow$  (i): The proof of this part is identical to the proof under Assumption 4(b).  $\square$

**Proof of Proposition 5.** If  $\pi^* = \infty$ , the proof of (iii)  $\Rightarrow$  (i) in Theorem 1 shows that  $\underline{p}_k = \bar{p}_k$  for all  $k$  large enough. Hence, surplus above the outside options can only arise from  $H$  and we have  $\bar{p}_* \leq \hat{p}$ .

Suppose  $\pi^*$  is finite and let  $p > \hat{p}$ . Since  $\pi_H^k(p) \rightarrow 0$ , we have  $\phi_{H,k}^-(p) \rightarrow p$  as  $k \rightarrow \infty$ . Hence, there exists  $\varepsilon$  such that  $p - \varepsilon > \hat{p}$  and  $\phi_{H,k}^-(p) > p - \varepsilon$  for all  $k$  large enough. Moreover, there exists  $\eta > 0$  such that  $S_k(p - \varepsilon) \geq \underline{s} + \eta$  for large  $k$ . This implies the incentive constraint for  $H$  holds at  $p$  for all  $k$  large enough. Since  $\delta_k \rightarrow 1$ , the surplus from  $H$  at  $p$  exceeds  $\underline{s}$  for large  $k$ . Hence,  $\bar{p}_* \leq p$ . Since  $p > \hat{p}$  was arbitrary, the proof is complete.  $\square$

## Proofs from Section 7.1

**Proof of Proposition 6.** Let  $w^h$  and  $w^l$  be the wages upon success and failure and let  $\mathbb{E}[u']$  be the expected continuation value of the worker. The worker's incentive constraint to choose  $H$  over  $L$  is given by

$$(1 - \delta) \left[ \pi_H(p)w^h + (1 - \pi_H(p))w^l - c \right] + \delta \mathbb{E}[u'] \geq (1 - \delta) \left[ \pi_L(p)w^h + (1 - \pi_L(p))w^l \right] + \delta \underline{u}$$

$$(1 - \delta)(w^h - w^l)(\pi_H(p) - \pi_L(p)) + \delta(\mathbb{E}[u'] - \underline{u}) \geq (1 - \delta)c$$

Without limited liability, the firm can always increase the spread between  $w^h$  and  $w^l$  to satisfy the above incentive constraint without affecting the expected wage payment  $\pi_H(p)w^h + (1 - \pi_H(p))w^l$ . Hence, the incentive constraint for  $H$  can be satisfied whenever the relationship yields more surplus than  $\underline{s}$ . Therefore, the equilibrium is efficient.

With limited liability, however, the firm can no longer create an arbitrary wage spread since  $w^h, w^l \geq \underline{u}$ . To tackle the first case in the characterization of efficiency, suppose  $R_H(p^{\text{FB}}) \leq \underline{s}$ . It suffices to show that  $H$  is enforceable at belief  $p^{\text{FB}}$  when continuations

are drawn from  $G$ . Equation (6) implies

$$\delta \left( \mathbb{E}^H[G(p^{\text{FB}})] - \underline{s} \right) \geq (1 - \delta)c$$

so we can set the worker's continuation payoffs  $u^+, u^- \geq \underline{u}$  following success and failure so that

$$\delta \left( \pi_H(p^{\text{FB}})(u^+ - \underline{u}) + (1 - \pi_H(p^{\text{FB}}))(u^- - \underline{u}) \right) = (1 - \delta)c$$

which makes the continuation payoffs for the firm given by  $v^+ = G(\pi_H^+(p^{\text{FB}})) - u^+$  and  $v^- = G(\pi_H^-(p^{\text{FB}})) - u^-$  greater than or equal to  $\underline{v}$ . By setting  $w^h = w^l = \underline{u}$  the overall payoff of the worker is  $\underline{u}$  and the firm's payoff is  $\underline{v}$  since  $G(p^{\text{FB}}) = \underline{s}$  and the path of play is the same as in the first best. Hence, all the incentive constraints for  $H$  are satisfied at  $p^{\text{FB}}$  and we are done.

We now show that whenever  $R_H(p^{\text{FB}}) > \underline{s}$  the incentive constraint cannot be satisfied at  $p^{\text{FB}}$ . Suppose, for a contradiction, that it can be satisfied. Without loss of generality  $w^l = \underline{u}$ , otherwise the difference between wages for success and failure can be increased without affecting the expected wage while still satisfying the incentive constraint for  $H$ . Since  $S(\pi_H^-(p^{\text{FB}})) \leq G(\pi_H^-(p^{\text{FB}})) = \underline{s}$  the continuation payoffs following  $H$  and failure must be equal to the outside options. Let  $(u^+, v^+)$  denote the continuation payoffs following  $H$  and success. Equation (6) and the incentive constraint for  $H$  imply

$$\begin{aligned} & (1 - \delta)(\pi_H(p^{\text{FB}}) - \pi_L(p^{\text{FB}}))(w^h - \underline{u}) + \delta\pi_H(p^{\text{FB}})(u^+ - \underline{u}) \\ & \geq (1 - \delta)(R_H(p^{\text{FB}}) - \underline{s}) + \delta \left( \mathbb{E}^H[G(p^{\text{FB}})] - \underline{s} \right) \end{aligned}$$

and we obtain

$$(1 - \delta)(\pi_H(p^{\text{FB}}) - \pi_L(p^{\text{FB}}))(w^h - \underline{u}) \geq (1 - \delta)(R_H(p^{\text{FB}}) - \underline{s}) + \delta\pi_H(p^{\text{FB}})(v^+ - \underline{v}). \quad (15)$$

Notice that  $R_H(p^{\text{FB}}) > \underline{s}$  implies  $w^h - \underline{u} > 0$ . Now consider the firm's payoff:

$$\begin{aligned} v &= (1 - \delta)[\pi_H(p^{\text{FB}})(y^h - w^h) + (1 - \pi_H(p^{\text{FB}}))(y^l - \underline{u})] + \delta[\pi_H(p^{\text{FB}})v^+ + (1 - \pi_H(p^{\text{FB}}))\underline{v}] \\ v - \underline{v} &= (1 - \delta)[R_H(p^{\text{FB}}) - \underline{s} - \pi_H(p^{\text{FB}})(w^h - \underline{u})] + \delta\pi_H(p^{\text{FB}})(v^+ - \underline{v}). \end{aligned}$$

Using equation (15) we obtain

$$v - \underline{v} \leq -(1 - \delta) - \pi_L(p^{\text{FB}})(w^h - \underline{u}) < 0$$

which is a contradiction since the firm's payoff must exceed his outside option in any equilibrium. By continuity, the same argument can be extended to show that the incentive constraint cannot be satisfied at some belief  $p > p^{\text{FB}}$  so relational contracts are inefficient.  $\square$

## References

- Abreu, D., P. Milgrom, and D. Pearce (1991). Information and timing in repeated partnerships. *Econometrica* 59(6), 1713–1733.
- Abreu, D., D. Pearce, and E. Stacchetti (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 1041–1063.
- Athey, S. and K. Bagwell (2008). Collusion with persistent cost shocks. *Econometrica* 76(3), 493–540.
- Battigalli, P. (1996). Strategic independence and perfect bayesian equilibria. *Journal of Economic Theory* 70(1), 201–234.
- Bhaskar, V. (2014). The ratchet effect: A learning perspective. *Working Paper, University of Texas at Austin*.
- Blackwell, D. (1953). Equivalent comparison of experiments. *Annals of Mathematics and Statistics* 32, 265–272.
- Bolton, P. and C. Harris (1999). Strategic experimentation. *Econometrica* 67(2), 349–374.
- Bull, C. (1987). The existence of self-enforcing implicit contracts. *The Quarterly Journal of Economics*, 147–159.
- Cisternas, G. (2018). Two-sided learning and the ratchet principle. *The Review of Economic Studies* 85(1), 307–351.

- DeVaro, J., J.-H. Kim, and N. Vikander (2018). Non-performance pay and relational contracting: Evidence from ceo compensation. *The Economic Journal* 128(613), 1923–1951.
- Fudenberg, D. and Y. Yamamoto (2010). Repeated games where the payoffs and monitoring structure are unknown. *Econometrica* 78(5), 1673–1710.
- Fudenberg, D. and Y. Yamamoto (2011). Learning from private information in noisy repeated games. *Journal of Economic Theory* 146(5), 1733–1769.
- Ghosh, P. and D. Ray (1996). Cooperation in community interaction without information flows. *The Review of Economic Studies* 63(3), 491–519.
- Halac, M. (2012). Relational contracts and the value of relationships. *The American Economic Review* 102(2), 750–779.
- Halac, M. (2014). Relationship building: Conflict and project choice over time. *Journal of Law, Economics, and Organization* 30(4), 683–708.
- Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies* 66(1), 169–182.
- Hörner, J. and L. Samuelson (2013). Incentives for experimenting agents. *The RAND Journal of Economics* 44(4), 632–663.
- Hörner, J., T. Sugaya, S. Takahashi, and N. Vieille (2011). Recursive methods in discounted stochastic games: An algorithm for  $\delta \rightarrow 1$  and a folk theorem. *Econometrica* 79(4), 1277–1318.
- Keller, G., S. Rady, and M. Cripps (2005). Strategic experimentation with exponential bandits. *Econometrica* 73(1), 39–68.
- Kuvalekar, A. V. and E. Lipnowski (2016). Job insecurity. *SSRN working paper*.
- Kwon, S. (2016). Relational contracts in a persistent environment. *Economic Theory* 61(1), 183–205.
- Laffont, J.-J. and J. Tirole (1988). The dynamics of incentive contracts. *Econometrica*, 1153–1175.

- Levin, J. (2003). Relational incentive contracts. *The American Economic Review* 93(3), 835–857.
- MacLeod, W. B. and J. M. Malcomson (1989). Implicit contracts, incentive compatibility, and involuntary unemployment. *Econometrica*, 447–480.
- Malcomson, J. M. (2016). Relational incentive contracts with persistent private information. *Econometrica* 84(1), 317–346.
- McAdams, D. (2011). Performance and turnover in a stochastic partnership. *American Economic Journal: Microeconomics* 3(4), 107–142.
- Pastorino, E. (2018). The role of job assignment, learning, and human capital acquisition for careers in firms. Technical report, mimeo.
- Peski, M. and T. Wiseman (2015). A folk theorem for stochastic games with infrequent state changes. *Theoretical Economics* 10(1), 131–173.
- Prat, J. and B. Jovanovic (2014). Dynamic contracts when the agent’s quality is unknown. *Theoretical Economics* 9(3), 865–914.
- Thomas, J. and T. Worrall (1988). Self-enforcing wage contracts. *The Review of Economic Studies* 55(4), 541–554.
- Thomas, J. and T. S. Worrall (2010). Dynamic relational contracts with credit constraints. *Available at SSRN 1599525*.
- Watson, J. (1999). Starting small and renegotiation. *Journal of Economic Theory* 85(1), 52–90.
- Watson, J. (2002). Starting small and commitment. *Games and Economic Behavior* 38(1), 176–199.
- Wiseman, T. (2005). A partial folk theorem for games with unknown payoff distributions. *Econometrica* 73(2), 629–645.
- Wiseman, T. (2012). A partial folk theorem for games with private learning. *Theoretical Economics* 7, 217–239.
- Yamamoto, Y. (2017). Stochastic games with hidden states. *SSRN working paper*.