# Policy Commitments in Relational Contracts

Daniel Barron
Northwestern University

Michael Powell
Northwestern University

October 2014

**Abstract**

How should an organization choose policies to strengthen its relationships with employees and partners? We explore how biased policies arise in relational contracts using a flexible dynamic game between a principal and several agents with unrestricted vertical transfers and symmetric information. If relationships are publicly observed, then optimal policies are never biased—they are always chosen to maximize total continuation surplus. In contrast, if relationships are bilateral—each agent observes only his own output and pay—then the principal may systematically choose backward-looking, history-dependent policies to credibly reward an agent who performed well in the past. We first show that biased policies are prevalent in a broad class of settings. Then we argue that biases can manifest in interesting ways in several simple examples. For instance, hiring may lag demand, promotions may be awarded in biased tournaments, and allocation decisions may "stick with" an inefficient worker.

# 1 Introduction

Business relationships often rest upon parties' goodwill rather than the contracts they sign–fear of destroying future surplus can motivate individuals both to perform well and to reward strong performance by their partners. In the canonical relational-incentive contracting models that capture this intuition (Bull, 1987; MacLeod and Malcomson, 1988; Levin, 2003), the principal's only role is to promise and pay monetary compensation to her agents. She is otherwise entirely passive.

Yet in any real-world enterprise, managers make a host of decisions that affect how a group of individuals contribute to the firm's objectives. Supervisors assign tasks to team members. Supply-chain managers source from suppliers. Executives allocate capital to divisions. Human-resource managers hire and fire employees. These decisions make certain individuals more integral and others less integral to the firm. And importantly, these decisions are often made on the basis of past performance, even when doing so harms future prospects. Supervisors bias promotions, CFOs bias capital allocations, and supply-chain managers bias future business toward those who saw past success (Peter and Hull, 1969; Graham, Harvey, and Puri, 2013; Asanuma, 1989). If the firm can compensate employees with monetary bonuses, then in principle it should be able to reward past successes without inefficiently tainting future decisions. Why, then, are biased decisions such a widespread feature in organizations?

In this paper, we argue that backward-looking policies can arise in optimally managed relationships among a principal and her agents. Biased decisions lead to lower continuation surplus. However, a principal who promises to bias future decisions towards an agent can credibly motivate that agent today using monetary payments. To make this point, we develop a general framework that builds upon Levin (2003)'s repeated principal-agent model with moral hazard, transferable utility, and risk-neutral parties. We extend Levin's framework to accommodate persistent public states and multiple agents. The key feature of our model is that the principal can make a public **decision** in each period that influences how agents' choices affect the firm's output. A **policy** is a complete decision plan for the relationship. A policy is **backward-looking** if it involves decisions that do not maximize continuation surplus. We say that such decisions are **biased**.

We first show that backward-looking policies never arise if relationships are public–that is, if all players commonly observe the history of past play. In this setting, the agents can coordinate to jointly punish the principal if she does not uphold her promises. In effect, the future surplus produced by all of the agents is at stake in each relationship. Biased decisions both decrease continuation surplus and weaken the principal's incentives to uphold her promises, and therefore have no place in a surplus-maximizing relationship. Instead, the principal "settles up" with her agents at the end of each period and makes decisions to maximize continuation surplus.

In contrast, backward-looking policies arise naturally if relationships are bilateral–that is, if each agent cannot observe the principal's interactions with other agents. In this setting, players cannot coordinate punishments or rewards. A decision that makes an agent more integral to the principal ensures that the principal and that agent have more to lose if they do not uphold their promises to one another. In particular, the principal can promise larger rewards to an agent who is expected to produce more future surplus. Decisions biased toward an individual therefore complement more generous reward schemes for that individual but also negatively affect the firm's overall performance in the future.

As an example of how backward-looking policies might optimally emerge, consider hiring decisions made by the owner of an up and coming business. Achieving early success requires sacrifice from early employees, and motivating this sacrifice requires the owner to promise rewards of either compensation or future goodwill. But these promises are only credible if maintaining relationships with early employees is important for the business. One way to ensure that early employees remain valued is for the owner to adopt a policy of being slow to hire following an increase in demand for the firm's products, which would make existing workers relatively more indispensable for the firm. Such a policy is not without costs, as orders may go unfulfilled, but these costs may be worth incurring in order to establish cooperative behavior early on. We explore this example in more detail in Section 6.

To formally argue that backward-looking policies arise in surplus-maximizing relational contracts, we define self-enforcing relational contracts in a game with imperfect private monitoring. We consider belief-free equilibrium (BFE) of the dynamic game. This solution concept provides a tractable approach that highlights why backward-looking policies arise.

We develop a set of straightforward necessary and sufficient conditions for a policy to be part of a self-enforcing relational contract. Using these conditions, we first consider a broad class of environments and show that backward-looking policies are typically part of surplus-maximizing relational contracts. Indeed, unless players are very patient or very impatient, decisions are biased with positive probability in nearly every period. We show that policies favor those agents who have performed well in the past at the expense of those who have not. In the resulting relational contract, agents compete to secure future decisions that are biased towards them.

Finally, we explore several examples and show that backward-looking policies can arise in a host of realistic settings. The inefficiencies that occur in these examples are of potential independent interest. Revisiting the hiring example, we confirm that additional hiring may optimally lag an increase in demand. We also argue that a firm might promote workers who have performed well in the past, even if they are not best-suited to the promoted position. And we show that a firm might inefficiently stick with an employee after learning that he is worse than the alternatives.

**Literature Review** Our paper is closely related to the literature on sequential inefficiencies in optimal contracts. The seminal contribution by Fudenberg, Holmstrom, and Milgrom (1990, henceforth FHM) considers sequential efficiency in long-term formal contracts. FHM identify several reasons why an optimal formal contract may entail inefficient continuation play; we highlight two here. First, the principal might only be able to punish the agent by simultaneously harming herself, for instance by firing the worker. Second, players might have asymmetric information about future payoffs. Under either of these conditions, the optimal formal contract may entail inefficiencies that arise over the course of play.

Within the relational contracting literature, Bull (1987), Baker, Gibbons, and Murphy (1994), Levin (2003), Kranz (2011), and many others study models in which the conditions from FHM hold. In these settings, stationary relational contracts are optimal and no sequential inefficiencies arise. A recent and growing literature, partially surveyed in Malcomson (2013), explores dynamic relational contracts that evolve based on past play. For instance, Fong and Li (2012) consider relational contracts if the agent has limited liability and show that the principal might inefficiently suspend production to punish poor performance. Li, Matouschek, and Powell (2014) show that if transfers are limited but the principal can reward and punish the agent with future control rights, she may permanently alter the firm's organization away from what maximizes continuation surplus. Board (2011) considers a setting with limited liability in which a principal chooses to trade with a single agent in each period. He argues that because a principal backloads incentive payments, she distorts this allocation decision to favor agents with whom she has traded in the past. Halac (2012), Malcomson (2014), and others study how relational contracts evolve if the players have asymmetric information about the future. Relational concerns influence dynamics in these papers. However, FHM's discussion suggests that the optimal formal contract in these settings could also entail history-dependent inefficiencies.

This paper takes a different approach. We focus on an environment in which the formal contract would not exhibit any history-dependent inefficiencies. Despite this, we show that history-dependent inefficiencies may arise in an optimal relational contract. Driven entirely by relational considerations, the principal may bias her decisions to favor some agents over others. Biased decisions are required to credibly motivate the agents, even though all parties are risk-neutral and have deep pockets. This intuition is closely related to Andrews and Barron (2014), who analyze optimal allocation dynamics in a supply chain, and Calzolari and Spagnolo (2011), who consider procurement auctions. The goal of our analysis is to extend the basic intuition of these papers and provide a general framework for analyzing backward-looking policies in relational contracts.

Our dynamic game has imperfect private monitoring. More precisely, we assume that one agent cannot observe the actions of the other agents. This assumption is similar to Segal's (1999) analysis of private offers in formal contracts, though our biases are quite different because they are driven by relational concerns. As discussed in Kandori (2002) and elsewhere, games with private monitoring are technically challenging because equilibrium payoffs depend on players' beliefs and so are not necessarily recursive. In this paper, we consider belief-free equilibrium (as in Ely, Horner, and Olszewski (2005)), which are recursive and so allow us to highlight the intuition behind biases in surplus-maximizing relationships.

# 2 Why Do Inefficient Policies Arise? An Example

In this section, we informally introduce the key ideas of our model in the context of an example.

Consider a principal who interacts with two agents in periods $t = 0, 1, ....$ In $t = 0$, the principal and each agent pay one another wages. Neither party is constrained by limited liability. Denote by $w_{i,t} \in \mathbb{R}$ the net wage to agent $i$. After this payment, each agent $i$ privately chooses a binary effort $e_{i,0} \in \{0, 1\}$ at cost $ce_{i,0}$. Agent $i$'s effort determines his output $y_{i,0} \in \{0, H_i\}$, with $H_1 > H_2 > 0$. The probability that $y_{i,0} = H_i$ equals $pe_{i,0}$. After output is realized, the principal and each agent exchange bonuses, with the net bonus to agent $i$ denoted $\tau_{i,0} \in \mathbb{R}$. At the start of $t = 1$, the principal chooses one of the two agents. He repeatedly plays this stage game with the chosen agent, but has no further interactions with the other agent. Players share a common discount factor $\delta \in (0,1)$. The principal and agent $i$ respectively earn $(1-\delta)\sum_{i=1}^{2}(y_{i,t} - w_{i,t} - \tau_{i,t})$ and $(1-\delta)(w_{i,t} + \tau_{i,t} - ce_{i,t})$ in period $t$.

Suppose that all variables except for effort are publicly observed. Assume that whichever agent is chosen in period 2 can be motivated to work hard from then on. How might the principal motivate the two agents in period 0? Agent $i$ can be motivated by either the expectation of a bonus or fine today ($\tau_{i,0}$) or a continuation payoff in period 2 onwards ($U_{i,1}$). So agent $i$'s **total reward** for producing output $y_{i,0}$ equals

$$B_i(y_{i,0}) = E[(1-\delta)\tau_{i,0} + \delta U_{i,1} | h^1],$$

where $h^1$ is the realized history.

Agent $i$'s reward is constrained because players cannot commit to a reward schedule. In particular, agent $i$ can always earn 0 by choosing $e_{i,t} = 0$ in each period. So $B_i \geq 0$ in equilibrium. The principal can similarly "walk away" from both relationships by refusing to pay the agents. Therefore, the principal will not be willing to pay the agents more than the total continuation surplus produced by both of them. If $q_i \in [0,1]$ is the probability that agent $i$ is chosen in period 1, then the sum of both agents' rewards must satisfy $B_1 + B_2 \leq \delta[p(q_1 H_1 + q_2 H_2) - c]$. In this case, $q_1 = 1$ is clearly the surplus-maximizing choice. This decision maximizes total continuation surplus in periods $t = 2, ....$ It also relaxes the upper bound on the aggregate reward $B_1 + B_2$ and so permits the principal to credibly promise strong incentives in period 1.

Suppose instead that agent $i$ observes his own output $y_{i,t}$ and pay $\{w_{i,t}, \tau_{i,t}\}$, but not the other agent's output or pay. Under this **bilateral monitoring** assumption, we argue that the principal might choose to continue her relationship with agent 2 even though doing so leads to lower surplus in periods $t = 1, ....$ Moreover, the principal's decision optimally depends on the realized outputs in period 1.

As before, agent $i$ is motivated by his expected reward $B_i(y_{i,0})$. Because $i$ can walk away from the relationship, $B_i \geq 0$. However, now the principal can refuse to pay agent $i$ without alerting the other agent to this deviation. Moreover, the agents have no way to communicate with one another. So the principal is willing to pay an agent no more than **the total continuation surplus produced by that agent.** If the principal were asked to pay more, she would earn more by abandoning her relationship with that agent and continuing to trade with the other. So agent $i$'s reward must satisfy

$$0 \leq B_i(y_{i,0}) \leq \delta[q_i(pH_i - c)].$$

Suppose the principal chooses agent 1 in period 1, which maximizes total continuation surplus. Then $q_2 = 0$ and $B_2 = 0$. Intuitively, the principal cannot credibly offer agent 2 any payoff because they interact only once. The principal can either maximize total continuation surplus in periods $t = 1, ...$ **or** motivate agent 2 in period 1, but she cannot do both. As a result, the optimal relational contract might entail **biased decisions** if $H_1 - H_2$ is not too large.

What type of inefficiencies arise? One possibility is that the principal chooses randomly between the two agents. In that case, $q_1 = q_2 = \frac{1}{2}$ and so both agents can be given some reward following high output. However, note that $B_2 > 0$ only if $y_{2,0} = H_2$. Therefore, the principal can do even better by setting $q_2 > 0$ only if $y_{2,0} = H_2$. Such a history-dependent policy ensures that the principal can credibly reward agent 2 at exactly the histories in which agent 2's reward is constrained from above.

In short, in a relational contract with bilateral monitoring, the principal's optimal policy entails history-dependent dynamic inefficiencies. Importantly, agents are motivated to work hard by the prospect of present and future monetary rewards. These wages and bonuses are made credible by the principal's policy. That is, the policy does not serve as a direct incentive for effort, but instead determines what kinds of direct incentives are credible in a relational contract. Inefficient policies arise even though the parties could in principle "settle up" using transfers in each period.

While this example may seem artificial, we argue that the same basic intuition leads to biased policies in many settings. The rest of this paper analyzes a model that generalizes this intuition and applies that model to several concrete examples.

# 3 The Model

A single principal (player 0, 'she') and $N$ agents (players $i \in \{1, ..., N\}$, each 'he') interact repeatedly in a dynamic game. Time is discrete and denoted by $t \in \{0, 1, \dots\}$. Players are risk-neutral and share a common discount factor $\delta \in (0, 1)$. In each period, the principal makes a **decision** $d$ from a set $D$. The decision determines how each agent $i$'s effort $e_i \in \mathbb{R}_+$ determines his **outcome** $y_i \in \mathbb{R}_+$. Agent $i$ incurs cost $c(e_i)$, while the principal earns revenue equal to the sum of outcomes, $\sum_{i=1}^{N} y_i$. There are two rounds of (vertical) transfers between the principal and each agent. The (net) ex-ante transfer to agent $i$ is denoted by $w_i$, and the (net) ex-post transfer to agent $i$ is denoted by $\tau_i$. We sometimes refer to these transfers, respectively, as wage and bonus payments, and we denote the vectors of wages and bonuses by $w$ and $\tau$. The principal sends a message $m_i$ to each agent $i$ along with the wage payment $w_i$. Denote the vector of messages by $m$.

**Technology** In period $t$, a set of **available decisions** $D \subseteq \mathcal{D}$ and **state of the world** $\theta \in \Theta$ are realized according to distribution $F\left(D, \theta | \{d_{t'}, D_{t'}, \theta_{t'}\}_{t'=0}^{t-1}\right)$, which depends on the history of decisions made by the principal as well as the history of available decisions and realized states. The decision $d \in D$ together with the state of the world $\theta$ and agent $i$'s effort $e_i$ determine the marginal distribution over agent $i$'s output: $P_i(y_i|\theta, d, e_i)$. Note that outcomes are independent across agents conditional on the decision and the state of the world.

**Timing** The stage game has eight rounds.

1. $\theta_t$ and $D_t$ are publicly realized according to $F(D_t, \theta_t| \{d_{t'}, D_{t'}, \theta_{t'}\}_{t'=0}^{t-1})$.

2. The principal makes a public decision $d_t \in D_t$.

3. For each agent $i$, the principal and agent $i$ simultaneously choose wage payments in $\mathbb{R}_+$ to send to one another. Define $w_{i,t} \in \mathbb{R}$ to be the (net) wage paid to agent $i$.

4. For each agent $i$, the principal chooses a message $m_{i,t} \in M$ to send to agent $i$, where $M$ is a large message space.

5. Each agent $i$ chooses whether to participate $(a_{i,t} = 1)$ or not $(a_{i,t} = 0)$. If agent $i$ does not participate, $y_{i,t} = 0$ and $i$ receives payoff $\bar{u}_i\left(d_t, \theta_t\right) \geq 0$.

6. If $a_{i,t} = 1$, then agent $i$ chooses effort $e_i \in \mathbb{R}_+$.

7. The outcome $y_t = (y_{1,t}, \ldots, y_{N,t})$ is realized, where $y_{i,t} \sim P_i\left(\cdot| \theta_t, d_t, e_{i,t}\right)$.

8. For each agent $i$, the principal and agent $i$ simultaneously choose bonus payments in $\mathbb{R}_+$ to send one another. Define $\tau_{i,t} \in \mathbb{R}$ as the net bonus to agent $i$.

It is worth pausing to comment briefly on the timing. In our game, agents decide whether or not to take their outside options after they pay or receive the ex-ante transfer $w_t$. This assumption is inconsequential under public monitoring. Under private monitoring, our equilibrium construction requires that agent $i$ is able to punish the principal by rejecting production following an off-path wage payment. We could allow a third round of transfers after accept/reject decisions but before efforts without any change in our results.

**Information** All players observe the state of the world $\theta$, the set of available decisions $D$, and the principal's decision $d$. The principal observes all transfers $w$ and $\tau$, accept/reject decisions $a$, messages $m$, and outcomes $y$, but she does not observe agent's efforts. Agent $i$ observes his own effort $e_i$, accept/reject decision $a_i$, wage $w_i$, message $m_i$, outcome $y_i$, and bonus $\tau_i$. He does not observe these variables for any other agent.

**Histories and Strategies** A history at the beginning of period $t$ is $h_0^t = \{\theta_{t'}, D_{t'}, d_{t'}, w_{t'}, m_{t'}, p_{t'}, e_{t'}, y_{t'}, \tau_{t'}\}_{t'=0}^{t-1}$, from set $\mathcal{H}_0^t$. Let $h_x^t \in \mathcal{H}_x^t$ denote the within-period history immediately following the realization of variable $x$, so for example, $h_w^t = h_0^t \cup \{\theta_t, D_t, d_t, w_t\}$. For every agent $i$, let $\phi_i\left(h_x^t\right)$ denote agent $i$'s private history at $h_x^t$ and $\phi_i(\mathcal{H}_x^t)$ the set of such histories. Likewise, $\phi_0(h_x^t)$ is the principal's private history and $\phi_0(\mathcal{H}_x^t)$ is the set of these histories. Recall that $\phi_0(h^t)$ includes all variables except effort, while $\phi_i(h^t)$ includes $\theta_t, D_t, d_t$, and those variables with subscript $i$. A **relational contract** is a strategy profile $\sigma = \sigma_0 \times ... \times \sigma_N$, where $\sigma_i$ maps $\phi_i\left(\mathcal{H}^t\right)$ to feasible actions at those private histories. Continuation play at $\phi_i(h^t)$ is denoted $\sigma_i|\phi_i(h^t)$. We refer to a history-contingent plan of decisions as a **policy**.

**Payoffsi** In period $t$, agent $i$'s and the principal's payoffs are

$$u_{i,t} = w_{i,t} + \tau_{i,t} - a_{i,t}c(e_{i,t}) + (1 - a_{i,t})\,\bar{u}_i\left(d_t, \theta_t\right),$$
$$\pi_t = \textstyle\sum_{i=1}^N \left(y_{i,t} - \tau_{i,t} - w_{i,t}\right),$$

respectively. Given a relational contract $\sigma$ and a history $h_x^t$, agent $i$'s continuation payoff is

$$U_i\left(\sigma, h_x^t\right) = E_\sigma\left[\sum_{t'=0}^{\infty} \delta^{t'}\left(1-\delta\right)u_{i,t+t'}\middle|\, h_x^t\right].$$

The principal's continuation payoff, $\Pi(\sigma, h_x^t)$, is defined analogously.

We define the **punishment payoff** for a player as the lowest individually-rational payoff for that player. The principal's punishment payoff is 0. Agent $i$'s punishment payoff is

$$\bar{U}_i(h_x^t) = \min_\sigma E_\sigma\left[\sum_{t'=0}^{\infty} \delta^{t'}(1-\delta)\bar{u}_i(d_{t+t'}, \theta_{t+t'})|h_x^t\right].$$

**Equilibrium** A problem that arises in games with imperfect private monitoring is that each player, given his private information, has to form beliefs about the private information of other players. If players condition their continuation play on their beliefs, then information - and hence play - grows increasingly complicated as the game progresses. To avoid these difficulties, we look for equilibria in which each player chooses the same continuation strategy no matter what private information he believes others possess.

DEFINITION 1. A relational contract $\sigma^*$ is a **Belief-Free Equilibrium (BFE)** if it satisfies the following two conditions. First, if private history $\phi_i(h_x^t)$ is off the equilibrium path,[0] then for any $i$ and any history $\tilde{h}_x^t$ satisfying $\phi_i(h_x^t) = \phi_i\left(\tilde{h}_x^t\right)$, $\sigma_i^*|\, \phi_i(h_x^t)$ is a best response to $\sigma_{-i}^*\big|\, \tilde{h}_x^t$. Second, if $\phi_i(h_x^t)$ is on-path, then for any $i$ and any on-path history $\tilde{h}_x^t$ such that $\phi_i(h_x^t) = \phi_i\left(\tilde{h}_x^t\right)$, $\sigma_i^*|\, \phi_i(h_x^t)$ is a best response to $\sigma_{-i}^*\big|\, \tilde{h}_x^t$.

We say a relational contract is self-enforcing if it is a BFE. Intuitively, a self-enforcing relational contract satisfies two conditions. At a history off the equilibrium path, each player $i$'s continuation strategy must be a best response to other players' strategies at any history that is consistent with what player $i$ has observed. On the equilibrium path, player $i$'s continuation strategy must be a best response to other players' strategies at any history that (i) is consistent with what player $i$ has observed, and (ii) is also on the equilibrium path.

Belief-Free Equilibrium is more restrictive than most standard solution concepts such as Perfect Bayesian Equilibrium. A player rules out any histories that are inconsistent with what he observes. Unless he has observed a deviation, he also rules out histories in which a deviation has occurred. The player must choose a continuation strategy that is simultaneously a best response to the other players at any history that has not been ruled out. In general games, this solution concept is quite limiting.[1]

In our framework, however, belief-free equilibrium lead to intuitive constraints on the relational contract and realistic history-dependent biases. Two features of the game facilitate this analysis. First, it is without loss of generality to focus on BFE in which agents do not condition on their past effort choices.[2] Second, the principal observes the true history and can send messages to the agents. So she can control how much information each agent has on the equilibrium path, which substantially simplifies the equilibrium conditions.

[0] Formally, if $\phi_i(h_x^t)$ is not in the support of the distribution over $\phi_i(\mathcal{H}_x^t)$ induced by $\sigma^*$.

[1] See Ely, Horner, and Olszewski (2005) for more details. Our solution concept is somewhat weaker then the one used in that paper because some variables are perfectly observed by multiple players in our setting.

[2] The proof of this fact may be found in Andrews and Barron (2014).

We focus on surplus-maximizing relational contracts in this paper. A self-enforcing relational contract $\sigma^*$ is **surplus-maximizing** if it yields higher ex ante total expected surplus than any other BFE. It is **sequentially surplus-maximizing** if at every on-path history $h_0^t \in \mathcal{H}_0^t$, continuation play $\sigma^*|h_0^t$ is surplus-maximizing in the continuation game beginning at $h_0^t$.[3] If $\sigma^*$ is not sequentially surplus-maximizing, then decisions are **biased** and the policy is **backward-looking**.

# 4 Public Relationships and Sequential Efficiency

As a benchmark, this section considers surplus-maximizing relationships if all variables (except effort $c_t$) are publicly observed. We show that surplus-maximizing relational contracts are always sequentially surplus-maximizing. That is, while relational contracts may entail different policies than formal contracts, the principal's decisions remain unbiased as her relationships evolve. This discussion throws into sharp relief the role of backward-looking policies in bilateral relationships.

The **game with public relationships** is similar to Section 3, with the following two differences. First, all variables except $\{e_{i,t}\}$ are observed by every player. Efforts remain private. In particular, all players know if someone has deviated from the relational contract. Therefore, a player can be held at his punishment payoff if he does not follow the relational contract. Second, the players are assumed to have access to a public randomization device in each round of the stage game. This assumption is for convenience - we could add a randomization device to the game with bilateral relationships without affecting our results.

Backward-looking policies reduce the total continuation surplus produced in a relational contract. This reduction in total surplus directly lowers *ex ante* total surplus. In the game with public relationships, the total surplus produced by all agents is lost if the principal reneges on any one relationship. Backward-looking policies decrease total continuation surplus, which potentially decreases the principal's payoff from following through on her promises. Thus, biased decisions also decrease total ex ante surplus by weakening the incentives that can be provided in equilibrium. Rather than implementing backward-looking policies, the principal can reward or punish an agent using monetary bonuses. Players can "settle up" without any need for biased decisions.

For these reasons, surplus-maximizing relational contracts are always sequentially surplus-

[3] An immediate implication of Lemma 1 is that without loss of generality, continuation play $\sigma^*|h_0^t$ forms a BFE of the continuation game. So sequential surplus-maximization is well-defined.

maximizing if relationships are public. Past choices may affect the state of the world or the decisions available to the principal, but they have no other impact on an optimal policy commitment.

PROPOSITION 1. *Consider the game with public relationships. Any surplus-maximizing relational contract is sequentially surplus-maximizing.*

**Proof:** See Appendix A.

Proposition 1 says that surplus-maximizing relational contracts need not condition on any past choices, except insofar as those choices affect the state of the world or the decisions available in the continuation game. The proof of this result adapts techniques developed in Levin (2003), Kranz (2014), and others.

Consider agent $i$'s moral hazard problem. The principal can motivate agent $i$ to work hard by varying his contemporaneous bonus payment $\tau_{i,t}$ and his continuation surplus $U_i$ with his output $y_{i,t}$. For a history $h_e^t$, define agent $i$'s **reward scheme** $b_i : Y_i \to \mathbb{R}$ as his expected continuation payoff for each possible outcome:

$$b_i(y_t) = E\left[(1-\delta)\tau_i + \delta U_i \middle| h_e^t, y_t\right].$$

An agent's reward scheme summarizes his incentives to exert effort. However, $b_i$ is constrained in a self-enforcing relational contract because it must be credible within the ongoing relationship. Our goal, then, is to provide bounds on $b_i$.

What are the *maximum* and *minimum* bonuses $\tau_i$ that can be credibly promised in a self-enforcing relational contract? Suppose agent $i$ is asked to pay more than his entire continuation utility from the relational contract, $\delta\left(U_i - \bar{U}_i\right)$. Then he would rather renege on this agreement and take his punishment payoff. So bonuses are bounded from below by $(1-\delta)\tau_{i,t} \geq -\delta\left(U_i - \bar{U}_i\right)$. Similarly, although the principal pays bonuses to multiple agents, the *sum* of these bonuses can never exceed her continuation utility. In particular, a necessary condition for equilibrium is that $(1-\delta)\tau_{i,t} \leq \delta\Pi$.[4]

These bounds on $\tau_i$ imply that agent $i$'s reward scheme has to satisfy the following **public dynamic-enforcement constraints**:

$$\delta\bar{U}_i \leq b_i \leq \delta(\Pi + U_i).$$

[4]This condition is not sufficient because it does not include bonuses paid to other agents. This necessary condition suffices to convey the intuition for the proof.

Intuitively, the relational contract constraints the *variation* in incentive pay that can be provided to each agent. Notice that $U_i$ enters only agent $i$'s dynamic-enforcement constraint, whereas the principal's continuation surplus $\Pi$ affects the constraint for *every* agent. Since players are risk-neutral and have deep pockets, agent $i$'s continuation surplus can be costlessly transferred to the principal. Doing so weakly relaxes all dynamic-enforcement constraints.

If all agents are held at their punishment payoffs, then the principal earns the total continuation surplus net of punishment payoffs. Consider a relational contract that prescribes inefficient on-path continuation play. Replacing an inefficient continuation with an efficient continuation maximizes total *ex ante* expected surplus. Increasing total surplus also increases the principal's payoff if agents earn their punishment payoffs, which relaxes all dynamic-enforcement constraints. So any sequentially inefficient relational contract is strictly worse than a sequentially surplus-maximizing relational contract, proving Proposition 1.

This argument requires that agents coordinate to jointly punish a betrayal by the principal. For instance, if an employer withholds a bonus from a deserving worker, then she faces sanctions from her entire workforce. If we relax this relatively stringent requirement, a principal might no longer be held to her punishment payoff following a deviation. In that case, the principal's policies determine the punishments each agent can impose on her. Biased decisions make it easier for the favored agent to punish the principal. We explore this intuition further in the next sections.

# 5 Bilateral Relationships and Sequential Inefficiency

If agents cannot jointly punish the principal, then relational concerns fundamentally shape the principal's policies. In this section, we develop straightforward necessary and sufficient conditions for self-enforcing relational contract in the game with bilateral relationships. Then we show that backward-looking policies are an integral feature of surplus-maximizing relational contracts.

In the game with bilateral relationships, each agent observes only his own output and bonuses, and furthermore cannot communicate with his counterparts. While this assumption is stylized, we believe that it captures an important feature of many real-world business relationships: widespread punishments are difficult to coordinate, especially when some of those involved in the punishment were not involved in the original deviation. In our framework, while a betrayed agent can deny the principal surplus by taking his outside option, the *other* agents do not observe the deviation and so may not punish the deviator.

The principal's decisions determine how much surplus is produced by each agent. Suppose a principal follows a backward-looking policy that make one agent's efforts relatively important to future profits. Then that agent can threaten to take his outside option if the principal does not follow through on the relational contract. Because the principal is more willing to reward the agent if she otherwise faces a severe punishment, decisions that are biased towards one agent allow the principal to *credibly* promise that agent a large payoff. Backward-looking policies arise because the principal needs to reward an agent who has performed well, and this reward is only credible if it is accompanied by a "hostage" in the form of a commitment to favor that agent with future decisions. In short, the surplus-maximizing relational contract balances *ex post* efficient policy choices against providing effective *ex ante* effort incentives.

At a history following effort $h_c^t$, recall that agent $i$'s reward scheme $b_i$ gives his expected payoff following each possible output realization. We consider the constraints that the bilateral relational contract imposes on each agent's reward scheme.

DEFINITION 2. *Define agent i's net cost* $C_{i,t} = a_{i,t}c(e_{i,t}) - (1 - a_{i,t})\bar{u}_i(d_t, \theta_t)$. *Given a relational contract* $\sigma$, history $h_x^t$, and any agent $i$, **i-dyad surplus** equals the total surplus produced by agent $i$ :

$$S_i\left(\sigma, h_x^t\right) = E_\sigma\left[\left.\sum_{t'=0}^{\infty}\delta^{t'}\left(1-\delta\right)y_{i,t+t'} - C_{i,t}\right| h_x^t\right]. \tag{1}$$

*A reward scheme* $b_i : \mathcal{H}_y^t \to \mathbb{R}$ *is* **credible in** $\sigma$ *if*

1. *It satisfies agent i's incentive-compatibility constraint: for each* $h_w^t$ *and every* $C_{i,t}$ *on the equilibrium path,*

$$C_{i,t} \in \underset{C_i|d_t,\theta_t}{\operatorname{argmax}}\, E_\sigma\left[\left. b_i\left(h_y^t\right)\right| h_w^t, C_i\right] - \left(1-\delta\right)C_i \tag{2}$$

2. *It satisfies bilateral dynamic enforcement: for each* $h_y^t$,

$$\delta E_{\sigma^*}\left[\left.\bar{U}_i\left(h_0^{t+1}\right)\right| h_y^t\right] \leq b_i\left(\phi_0\left(h_y^t\right)\right) \leq \delta E_{\sigma^*}\left[\left.S_i\left(\sigma^*, h_0^{t+1}\right)\right| h_y^t\right] \tag{3}$$

A credible reward scheme satisfies two conditions. First, agent $i$ has to be willing to exert effort $e_{i,t}$ if his continuation surplus equals $b_i\left(h_y^t\right)$. So $b_i\left(h_y^t\right)$ has to vary sufficiently in $y_{i,t}$ in order to motivate effort. The second condition limits the variation in $b_i$ by bounding it from above and below. Agent $i$ can never earn more surplus than $S_i$, the amount he produces in the continuation game. Since $i$-dyad surplus $S_i$ can potentially vary in realized output, this condition has to hold *for each possible output*. In addition, agent $i$ has to always earn at least his punishment payoff. These constraints hold output-by-output and limit variation in $b_i$.

We show that every self-enforcing relational contract has a corresponding credible reward scheme for each agent $i$. Moreover, if a strategy has a credible reward scheme, there exists a self-enforcing relational contract that implements the same policy and generates the same total surplus as that strategy.

LEMMA 1. *Consider the game with bilateral relationships.*

1. *If $\sigma^*$ is a self-enforcing relational contract, then for each agent $i$ there exists a reward scheme $b_i^*$ that is credible in $\sigma^*$.*
2. *Suppose $\sigma$ is a relational contract with a credible reward scheme $b_i$ for each agent $i$. Then there exists a self-enforcing relational contract $\sigma^*$ that induces the same joint distribution over decision sets, decisions, efforts, and outcomes as $\sigma$.*

**Proof:** See Appendix A.

The first statement of Lemma 1 follows from an argument similar to Proposition 1. Define the reward scheme $b_i\left(h_y^t\right)$ as agent $i$'s total continuation surplus at history $h_y^t$, which includes both the bonus $(1-\delta)\,\tau_{i,t}$ and his continuation payoff $\delta U_i$. This reward scheme has to satisfy agent $i$'s effort IC constraint (2) in any self-enforcing relational contract. Because agent $i$ can "walk away" from the contract by rejecting production in each period, continuation surplus is bounded below by his punishment payoff. The principal can similarly walk away from his relationship with agent $i$ by not paying him wages or bonuses. Importantly, she can do so without alerting any other agents because the other agents do not observe $i$'s wages, bonuses, or output. So the principal is willing to pay agent $i$ no more than her continuation surplus from her relationship with $i$. This logic gives an upper bound on $(1-\delta)\,\tau_{i,t}$, which in turn gives the upper bound on $b_i$ given by (3).

The proof of the second statement is a little more involved. Intuitively, we construct a self-enforcing relational contract from the strategy $\sigma$. In each period of this relational contract, the principal chooses the same decision as in $\sigma$. She then sends a message to each agent specifying the equilibrium effort choice and the reward scheme in that period. This message is accompanied by a wage that ensures that the principal earns 0 in each period. The agent exerts the specified effort and then *repays* the principal according to the specified reward scheme. Any deviation by a player is punished by a breakdown of the corresponding relationship. The principal earns 0 in each period and so is willing to choose the equilibrium decision in each period. Her message is made credible by the accompanying wage: if the principal specifies a steep reward scheme, then she has to also pay a large upfront wage. The agent is willing to exert effort and make the specified payments because the reward scheme is credible.

Lemma 1 implies that backward-looking policies play an important role in bilateral relationships. The principal's decision determines the amount of surplus produced by each agent. Increasing $i$-dyad surplus relaxes the dynamic-enforcement constraint for agent $i$, but potentially decreases $j$-dyad surplus and so tightens this constraint for agent $j$. Future decisions can be made conditional on past outputs, so these trade-offs can be made in a history-dependent way. In particular, future decisions can favor agent $i$ precisely when $i$'s reward scheme is constrained by the upper bound of (3). Backward-looking policies can relax the dynamic-enforcement constraint and allow the principal to better motivate some agents. However, biased decisions potentially decrease *total* continuation surplus, leading to surplus-maximizing relational contracts that are not sequentially surplus-maximizing.

Our final goal of this section is to show that biased decisions are a typical feature of surplus-maximizing relational contracts. To make this argument, we restrict attention to a class of "smooth" repeated games. In these games, the decision set $D_t$ is constant, states of the world $\theta_t$ depend only on past states of the world, and all payoffs are smooth in decisions and efforts.

DEFINITION 3. *A game with bilateral relationships is* ***smooth*** *if:*

1. $D_t = \left\{(d_1, ..., d_N)|\, d_i \in \mathbb{R}_+, \sum_{i=1}^N d_i \leq 1\right\}$ *in each period. The distribution of* $\theta_t$ *depends only on* $\{\theta_{t'}\}_{t'=0}^{t-1}$
2. *Outside options depend only on* $\theta_t$, $\{\bar{u}_i(\theta_t)\}_{i=0}^N$ *.Effort costs* $c(\cdot)$ *are smooth, strictly increasing, and strictly convex.*
3. $P_i$ *depends only on* $d_i, \theta$, *and* $e_i$. *For each* $\{d_i, \theta\}$, $P_i$ *has smooth density* $p_i$, *is strictly*

*MLRP-increasing in $e_i$, and satisfies CDFC. $E\left[y_i|d_i,\theta,e_i\right]$ is smooth, strictly increasing, and strictly concave, with $\lim_{d_i\to 0}\frac{\partial}{\partial d_i}E\left[y_i|d_i,\theta,e_i\right]=\infty$ for all $\{\theta,e_i\}$.*

In a smooth game, a decision specifies a weight $d_{i,t}$ for each agent $i$ in period $t$. Agent $i$'s effort together with this weight determine the outcome $y_i$, where a higher weight $d_{i,t}$ leads to a larger expected $y_{i,t}$. Expected outcomes are smooth in all arguments, and we assume boundary conditions to ensure that optimal weights $d_t$ are strictly positive. The distribution of outcomes satisfies the Mirrlees-Rogerson conditions, which ensures that we can replace the incentive-compatibility constraint (2) with its first-order condition.

Given these assumptions, the first-best level of effort can be defined for each state of the world $\theta$ and decision $d_i$:

$$e_i^{FB}(d_i,\theta)=\arg\max_{e_i} E[y_i|d_i,\theta,e_i]-c(e_i). \tag{4}$$

Since output is strictly MLRP-increasing in effort, there exists a unique $y_i^*(d_i,\theta,e_i)\in\mathbb{R}_+$ that satisfies

$$\frac{\partial p_i/\partial e_i}{p_i}(y_i^*(d_i,\theta,e_i)|d_i,\theta,e_i)=0. \tag{5}$$

Loosely, output $y_i>y_i^*$ statistically suggests that agent $i$ chose an effort no lower than $e_i$.

A critical feature of these games is that decisions entail trade-offs among agents. Maximizing the surplus produced by agent $i$ requires $d_i=1$, which requires that all other agents are given no weight and so are not very productive. These trade-offs drive the biased policies that arise in a surplus-maximizing relational contract. Indeed, transfers can be used to costlessly punish low performance by the agents. So surplus-maximizing relationships use biased decisions to *relax dynamic enforcement constraints* for high-performing agents. However, relaxing one agent's dynamic enforcement constraint necessarily entails decreasing the surplus produced by some other agent.

The next result shows that biased decisions are typically a part of surplus-maximizing relationships in smooth games.

PROPOSITION 2. *Consider a smooth game with bilateral relationships. In any surplus-maximizing relational contract $\sigma^*$,*

1. **Money is never burned***: $\sum_{i=1}^{N} d_{i,t}=1$ with ex ante probability 1.*

2. **Sequential surplus maximization entails equal marginal returns**: *$\sigma^*$ is sequentially surplus-maximizing only if for any agents $i$ and $j$ and any period $t$,*

$$\frac{\partial}{\partial d_{i,t}} E_{\sigma^*}\left[ y_{i,t} | \, h_e^t \right] = \frac{\partial}{\partial d_{j,t}} E_{\sigma^*}\left[ y_{j,t} | \, h_e^t \right] \tag{6}$$

*holds with ex-ante probability* 1.

3. **Biased policies are optimal**: *For any agents $i$ and $j$, let $E_t$ be a set of histories $h_e^t$ such that : (i) $e_{i,t} \in (0, e_i^{FB}(d_{i,t}, \theta_t))$, (ii) $y_{i,t'} > y_i^*(d_{i,t'}, \theta_{t'}, e_{i,t'})$ for some $t' < t$, and (iii) $y_{j,t'} < y_j^*(d_{j,t'}, \theta_{t'}, e_{j,t'})$ for all $t' < t$. If $\Pr_{\sigma^*}\{E_t\} > 0$, then (6) fails to hold in $E_t$.*

**Proof:** See Appendix A.

The first statement of Proposition 2 holds because larger $d_i$ both increases total surplus and relaxes (3) for agent $i$. So any surplus-maximizing relational contract will use the full "budget" of $d_i$ - the only question is what weight is assigned to each agent. For the second statement, actions in $t = 0$ do not affect any dynamic-enforcement constraints and are chosen to maximize myopic total surplus. So (6) has to hold in period 0 of any surplus-maximizing relational contract. If $\sigma^*$ is sequentially surplus-maximizing, then $\sigma^* | \, h_0^t$ is surplus-maximizing and so (6) has to hold at $h_0^t$.

For statement 3, suppose that the stated conditions hold. We construct a backward-looking policy that has a second-order cost relative to the sequentially surplus-maximizing policy but leads to a first-order increase in effort. The credible reward scheme that induces maximal effort from agent $i$ in period $t$ is

$$b_i = \begin{cases} \delta \bar{U}_i & y_{i,t} < y_i^* \\ \delta S_i & y_{i,t} \geq y_i^* \end{cases} \tag{7}$$

for some $y_i^*$ that depends on the decision, the state of the world, and $i$'s equilibrium effort. Note that $S_i$ depends on future decisions, which in equilibrium can depend on realized outputs $y_t$. Consider some period $t \geq 1$. Suppose $y_{i,t'} > y_i^*$ for agent $i$ and agent $j$ has never produced output larger than $y_j^*$. Then $b_j = \delta \bar{U}_j$ for agent $j$ in each period $t' < t$, and in particular $j$-dyad surplus is irrelevant for agent $j$'s effort incentives in those periods. Consider increasing $d_{i,t}$ and decreasing $d_{j,t}$ by some small amount. Holding effort fixed, this leads to a second-order decrease in total surplus if (6) holds. This change has no effect on agent $j$'s effort because $b_j = \delta \bar{U}_j$ in every previous period. However, it strictly relaxes agent $i$'s dynamic-enforcement constraint in period $t'$. Since $y_{i,t'} > y_i^*$, agent $i$ is motivated to exert strictly higher effort, which has a first-order impact on surplus because $e_i < e_i^{FB}$. Thus, biased decisions increase both effort provision and total surplus.

# 6 Examples of Biases

Biased relational contracts can arise in a many different settings. In this section, we use three simple examples to illustrate the types of biases that might arise in a relationship. First, we consider hiring decisions and prove that employment can optimally lag demand. Second, we show how a firm might distort irreversible promotions or investments to better motivate its employees or divisions. Finally, we argue that a manager might continue to favor an employee even after learning that he is worse than an alternative worker. For simplicity, we will assume $\bar{u}_i = 0$ and $e_{i,t} \in \{0, 1\}$ with cost $ce_{i,t}$ in all three examples..

## 6.1 Hiring and Firing

Consider a firm who faces persistent demand shocks and decides how many agents to hire in each period. This example illustrates how persistent shocks in demand and diminishing per-worker productivity can lead to hiring that substantially lags demand.

DEFINITION 4. *The* ***hiring game with demand shocks*** *has* $N = 2$ *and the following features:*

- *Demand is* $\Theta = \{W, R\}$ *with* $0 < W < R$. *If* $\theta_t = R$, then $\theta_{t+1} = R$. If $\theta_t = W$, then $\theta_{t+1} = R$ with probability $q < 1$.
- *In each period,* $D_t = \{1, 2\}$. *The principal hires* $d_t \in D_t$ *agents. For convenience, we assume that if* $d_t = 1$, *then agent* 1 *is hired.*
- *If agent* $i$ *is not hired, then* $y_{i,t} = 0$. *Otherwise,* $y_{i,t} = \theta_t e_{i,t}$ *if* $d_t = 1$ *and* $y_{i,t} = \theta_t \alpha e_{i,t}$ *with* $\alpha < 1$ *if* $d_t = 2$..

The principal is a firm that faces demand $\theta_t$ in period $t$. If demand is weak ($\theta_t = W$), then it might either grow (to $\theta_{t+1} = R$) or remain the same in the next period. Once demand increases, it remains robust thereafter. The return to an agent's effort in period $t$ is determined by both demand and the number of agents hired in $t$. We assume that marginal productivity is decreasing in the number of workers ($\alpha < 1$). The optimal number of employees depends on demand and the effort chosen by each worker: so long as agents exert effort, the firm maximizes myopic profit by hiring two workers if $\theta_t = R$ and one worker if $\theta_t = W$.

Suppose that the game has public relationships as in Section 4. Then every relationship is strongest if policies are chosen to maximize *total continuation surplus*. Therefore, as long as the agents exert effort when demand is robust, surplus is maximized if the firm hires two agents when demand is robust and a single agent otherwise.

In contrast, the surplus-maximizing relational contract in the game with bilateral relationships exhibits substantial history-dependent biases. The firm might delay hiring a second worker following an increase in demand in order to credibly reward the existing employee for his hard work during a low-demand period.

PROPOSITION 3. *Consider the hiring game with bilateral relationships. Suppose that* $R > \frac{c}{2\alpha-1} > W > c$ *and* $\alpha R > W$. *Then there exists a range of discount factors* $\left(\underline{\delta}, \bar{\delta}\right) \subset [0,1]$ such that for $\delta \in \left(\underline{\delta}, \bar{\delta}\right)$, *any surplus-maximizing relational contract* $\sigma^*$ *satisfies:*

1. *If* $\theta_0 = R$, *then* $d_t = 2$ *in every period* $t$.

2. *If* $\theta_0 = W$, *then* $d_t = 1$ *whenever* $\theta_t = W$. *Moreover, there exists some period* $t'$ *such that* $\Pr_{\sigma^*}\{d_{t'} = 1, \theta_{t'} = G\} > 0$.

*If* $\theta_0 = W$, *then one surplus-maximizing relational contract satisfies the following.* $e_{i,t} = 1$ *in any period in which agent* $i$ *is hired . If* $\theta_t = R$ *for the first time in period* $t$, *then* $d_t = 1$ *with probability* $\gamma$. *In every subsequent* $t' > t$, $d_{t'} = d_t$.

**Proof:** See Appendix A.

The firm immediately hires two workers if it begins with robust demand. If demand is initially weak, then the firm hires only one worker. Moreover, it may continue to hire only one worker even after demand becomes robust. If players are neither too patient nor too impatient, then the dynamic enforcement constraint (3) is satisfied for $c_{i,t} = 1$ in the high-demand state with $d_t = 2$. Since low demand is persistent, however, it might be impossible to satisfy (3) in the weak-demand state without distorting hiring policies. By not hiring a second worker after demand increases, the principal can ensure that the agent hired in the weak-demand state can be credibly motivated to work hard. That is, the principal promises an inefficient hiring policy in the future to motivate his remaining workers in weak-demand times. The two assumptions required for this result ensure that (i) myopic profit is maximized by hiring two workers in a high-demand state and one worker in the low-demand state, and (ii) *net per-worker productivity* is higher if demand is robust, regardless of the number of workers hired.

This hiring delay could take many different forms. In Proposition 3, we demonstrate that one surplus-maximizing distortion is for the firm to make a *once-and-for-all* decision whether or not to expand as demand grows. While the particulars of this equilibrium rely on the simple setting considered here, this equilibrium illustrates that the optimal relational contract may entail substantial distortions.

This example is consistent with recent empirical work by Ariely, Belenzon, and Tsolmon (2013), who argue that firms that rely on relational contracts tend to expand more slowly than those that rely on formal contracts. Here, hiring remains slow because the firm must fulfill its promises to old employees before expanding. New firms have no promises to fulfill, so they can immediately expand to take advantage of improved productivity. Therefore, following a recession or other period of low demand, this model would suggest that new entry may drive increased employment immediately after a recession or other period of low demand. Consistent with this argument, Haltiwanger, Jarmin, and Miranda (2013) find that young firms tend to drive net job growth in the US from 1976-2004.

## 6.2 Irreversible Promotions

Suppose a principal has to permanently promote one of her agents at some point in the future. The promoted agent takes on responsibilities so that her effort leads to strictly more profits for the firm. Some workers would be more adept at these tasks than others. Which agent should be promoted?

In this example, we show that the principal should optimally run a (potentially biased) tournament among her agents for the promoted position. The agent who performs "best" according to this tournament is promoted, even if he might not be best suited to the promoted position. This example applies equally well to other kinds of irreversible and agent-specific investments, such as firm-specific human capital or division-specific physical plant.

DEFINITION 5. *The* ***irreversible promotions game*** *has* $N = 2, |\Theta| = 1$, *and the follow-*

*ing features:*

- *The set of possible decisions is* $D = \{0, 1, 2\}$. *No promotion is denoted* $d = 0$ *while* $d \in \{1, 2\}$ *indicates agent* $d$ *is promoted.*

- *Investments are delayed and permanent.* $D_0 = \{0\}$ *and* $D_1 = \{1, 2\}$. *For any* $t > 1$, $D_t = \{d_{t-1}\}$.

- *The outcome distribution* $P_i(\cdot|d, e_i)$ *is smooth with density* $p_i$ *and strictly MLRP increasing in* $e_i$. *It is the same for each agent* $i$ *if* $d_t \neq i$, *while* $E[y_1|d_t = 1, e_{1,t}] - E[y_2|d_t = 2, e_{2,t}] \equiv \Delta > 0$. *For each agent* $i$, $E[y_i|d_t = i, e_{i,t}] > E[y_i|d_t \neq i, e_{i,t}]$.

Define

$$L_i(y_i|d) = \frac{p_i(y_i|d, e_i = 1)}{p_i(y_i|d, e_i = 0)}$$

as the likelihood ratio for output $y_i$ given decision $d$. Because $P_i$ is MLRP-increasing in $e_i$, $L_i$ is strictly increasing in $y_i$.

In the irreversible promotions game, the principal chooses one of the two agents to promote at the end of the first period. Agents have identical productivities if they are not promoted, but agent 1's output is higher than agent 2's in the promoted position. Once promoted, an agent cannot be demoted. Proposition 1 implies that the principal should maximize total continuation surplus if relationships are public. Hence, the principal should always promote agent 1.

The surplus-maximizing policy is very different in the game with bilateral relationships. The principal can use a promised promotion to make a large reward to the promoted agent credible. As a result, the principal can potentially motivate both agents by offering to promote (and monetarily reward) whichever agent produces high output, since high output is indicative of high effort. The result is a tournament in which the less-efficient agent may be promoted if he performs well in the early periods of the game. This tournament will typically be "biased," since the principal wants to maximize the probability that the efficient agent is promoted subject to the constraint that both agents exert effort in the first period.

PROPOSITION 4. *Consider the irreversible promotions game with bilateral relationships. There exists* $0 \leq \underline{\delta} < \bar{\delta} < 1$ *and* $\bar{\Delta} > 0$ *such that if* $\underline{\delta} < \delta < \bar{\delta}$ *and* $\Delta < \bar{\Delta}$, *any surplus-maximizing relational contract* $\sigma^*$ *satisfies:.*

1. $e_{1,0} = e_{2,0} = 1$;

2. $d_1 = 2$ *with strictly positive probability. Either* $d_1 = 2$ *with probability 1 or* $d_2 = 2$ *if* $L_2(y_{2,0}|d = 0) > 1$ *and*

$$\frac{1}{L_2(y_{2,0}|d = 0)} < \alpha + \beta \left( \frac{1}{L_1(y_{1,0}|d = 0)} \right)$$

*for some* $\alpha \in R$ *and* $\beta \geq 0$.

**Proof:** See Appendix A.

If the agents' productivities are not too different, then the principal finds it optimal to promote an agent who produces high output. Because both agents have the opportunity to "win" the promotion, both are willing to work hard in the first period. After the principal promotes one agent, that agent's dynamic enforcement constraint is slack and so he is willing to continue working hard. However, the other agent no longer works hard, because he cannot be credibly given strong incentives.

In short, the surplus-maximizing relational contract entails a tournament between the agents. Promotion is used as a "prize" but does not directly compensate the promoted worker. Instead, it is used to make monetary compensation credible within the context of the relational contract.

## 6.3 Learning and Task Allocation

A principal can give an assignment to one of two agents. One of the agents has a known productivity, while the other's ability is unknown and can only be learned by giving the assignment to that worker. How should the principal's task assignment policy evolve as he learns more information?

This example illustrates a cost of experimentation in a relational contract: the principal might "stick with" an inefficient agent, rather than switching to the more efficient alternative.[5]

DEFINITION 6. *The learning assignment game has the following features:*

[5] Strictly speaking, belief-free equilibria does not entail Bayesian updating. We define "learning" as a mechanical feature of the dynamic game in this example.

- *The state space represents (symmetric) beliefs about agent 1's type,* $\Theta = \{L, \alpha L + (1-\alpha) H, H\}$*. Feasible decisions are* $D_t \in \{1,2\}$*, with agent* $d_t$ *assigned to the task.*

- *If 1's type is known or 2 is assigned the task, then no learning occurs:* $\theta_{t+1} = \theta_t$ *if* $\theta_t \in \{L, H\}$ *or* $d_t = 2$*. If* $\theta_t = \alpha L + (1-\alpha)H$ and $d_t = 1$, then $\theta_{t+1} = L$ *with probability* $\alpha$ *and otherwise* $\theta_{t+1} = H$*.*[6]

- *If* $d_t \neq i$*, then* $y_{i,t} = 0$*. If* $d_t = 1$*, then* $y_{1,t} = \theta_t c$ *with probability* $p$ *and otherwise* $y_{1,t} = 0$*. If* $d_t = 2$*, then* $y_{2,t} = Rc$ *with probability* $p$ *and otherwise* $y_{2,t} = 0$*. Assume* $\frac{1}{p} < L < R < qL + (1-q)H$*.*

In each period of the learning assignment game, the principal chooses one of two agents to exert effort. Agent 1's productivity is unknown: with probability $1-\alpha$ it is higher than agent 2's, but otherwise it is lower. To learn 1's productivity, the principal must allocate production to 1. But 1 will shirk unless the principal's policy ensures that he is given production in the future as well. In particular, the principal might have to promise to continue allocating production to agent 1 even if he turns out to have low productivity.

Proposition 5 illustrates this point.

PROPOSITION 5. *Consider the learning assignment game with bilateral relationships. Suppose* $(1-q)(pH-1) < pL - 1$*. There exists a range* $0 < \underline{\delta} < \bar{\delta} < 1$ *such that if* $\underline{\delta} < \delta < \bar{\delta}$*, then* $d_0 = 1$ *in any surplus-maximizing relational contract. There exists some* $t' > 0$ *such that* $\Pr_{\sigma^*}\{\theta_{t'} = L, d_{t'} = 1 | y_{1,0}\} > 0$ *if and only if* $y_{1,0} > 0$*.*

**Proof:** See Appendix A.

The intuition for Proposition 5 is similar to the previous two examples. The principal faces a trade-off between *ex post* surplus-maximizing decisions and motivating agent 1 *ex ante*. If the returns to learning about agent 1 are large, then she initially allocates to him. If he performs well, then she might continue choosing him, even if he ends up having a low productivity.

[6] In this example, all players learn agent 1's type. Nothing would change if we instead assumed that agent 2 did not learn agent 1's type.

## 7 Conclusion

Biased policies are a prominent feature of many long-term relationships. Managers favor high-performing workers, divisions, and suppliers by choosing policies that make those parties integral to the production process. In this paper, we have argued that biased decisions can arise in surplus-maximizing relational contracts, even if the principal may freely reward or fine her agents. By increasing the surplus produced by one agent (at the cost of reducing the surplus produced by others), biased decisions complement and make credible large monetary rewards. As a result, employees are rewarded with both higher compensation and greater responsibilities, divisions are promised both monetary incentives and non-monetary investments, and suppliers are motivated by both contemporaneous fines and the promise of future business.

We have presented a series of simple examples to argue that these biases manifest in intuitive ways. Future research is needed to both expand the scope and enrich the analysis in different settings. For example, our analysis of hiring decisions during recoveries implies that new entrants would be responsible for a substantial share of new hires, since these entrants would not be bound by past promises. Productivity should be higher during a recovery than before the recession. Both of these results are broadly consistent with stylized facts from the 2008 recession. A richer analysis could identify other predictions that might be amenable to empirical analysis.

In our setting, organizations do not adjust to changing circumstances because they are weighed down by relational obligations. The nature of these obligations - and hence the momentum of a given firm - depends critically on the history of that firm. Therefore, relational contracts provide an explanation for the tremendous heterogeneity among organizations in many markets.

# Appendix A: Proofs

## Proof of Proposition 1

We begin the proof with a lemma that gives necessary and sufficient conditions for a strategy profile to be an equilibrium of the game with public monitoring.

### Statement of Lemma A.1

1. If $\sigma^*$ is a BFE, then for any agent $i \in \{1, \ldots, N\}$ there exists a function $b_i : \phi_0\left(\mathcal{H}_y^t\right) \to \mathbb{R}$ satisfying

   (a) **Effort IC:** $b_i$ satisfies (2).

   (b) **Public Dynamic Enforcement:** for any $I \subset \{1, \ldots, N\}$ and $h_y^t$,

$$\delta \sum_{i \in I} E_{\sigma^*}\left[\bar{U}_i\left(h_0^{t+1}\right) \middle| h_y^t\right] \leq \sum_{i \in I} b_i\left(\phi_0\left(h_y^t\right)\right) \leq \delta \sum_{i \in I} E_{\sigma^*}\left[\sum_{i \in I} U_i\left(\sigma^*, h_0^{t+1}\right) + \Pi\left(\sigma^*, h_0^{t+1}\right) \middle| h_y^t\right]. \tag{8}$$

   (c) **Individual Rationality**: for any $h_d^t \in \mathcal{H}_d^t$, $i \in \{1, ..., N\}$, and $I \subseteq \{1, ..., N\}$,

$$\begin{aligned} U_i(\sigma^*, h_d^t) &\geq \bar{U}_i(h_d^t) \\ \Pi\left(\sigma^*, h_d^t\right) &\geq \sum_{i \in I}\left(E_{\sigma^*}\left[b_i\left(\phi_0\left(h_y^t\right)\right) - (1-\delta)\, c_{i,t} \middle| h_d^t\right] - U_i\left(\sigma^*, h_d^t\right)\right) \end{aligned} \tag{9}$$

2. For strategy $\sigma$, suppose there exists $b_i : \phi_0\left(\mathcal{H}_y^t\right) \to \mathbb{R}$ satisfying (2), (8), and (9). Then there exists a BFE $\sigma^*$ that induces the same distribution over $\left\{D_t, d_t, e_t, y_t, \{u_i\}_{i=0}^N\right\}_{t=0}^\infty$ as $\sigma$.

### Proof of Lemma A.1

**1:** Suppose $\sigma^*$ is a BFE. Then at any $h_0^t \in \mathcal{H}_0^t$, agent $i$ can earn at least $\bar{U}_i\left(h_0^t\right)$ by taking his outside option in each period. Similarly, the principal can earn no less than 0.

Define

$$b_i\left(\phi_0\left(h_y^t\right)\right) = E_{\sigma^*}\left[(1-\delta)\,\tau_{i,t} + \delta U_i \middle| \phi_0\left(h_y^t\right)\right]. \tag{10}$$

Agent $i$ chooses $e_{i,t}$ to solve

$$e_{i,t} \in \underset{e_i \in \mathbb{R}_+}{\operatorname{argmax}}\, E_{\sigma^*}\left[(1-\delta)\,\tau_{i,t} + \delta U_i \middle| h_w^t, e_{i,t} = e_i\right] - (1-\delta)\, c(e_i), \tag{11}$$

which implies (2). Suppose $b_i\left(\phi_0\left(h_y^t\right)\right) < \delta E_{\sigma^*}\left[\bar{U}_i\left(h_0^{t+1}\right)\middle|\,\phi_i\left(h_y^t\right)\right]$. Then agent $i$ may profitably deviate by choosing $\tau_{i,t}=0$ and earning no less than $\bar{U}_i\left(h_0^{t+1}\right)$ in the continuation game. Suppose there exists a set $I \subset \{1,\ldots,N\}$ such that

$$\sum_{i\in I} E_{\sigma^*}\left[\tau_{i,t}\middle|\,\phi_0\left(h_y^t\right)\right] > \delta E_{\sigma^*}\left[\Pi(\sigma^*, h_0^{t+1})\middle|\,\phi_0\left(h_y^t\right)\right].$$

Then the principal may profitably deviate by choosing $\tau_{i,t}=0$ for all $i\in I$, earning no less than 0 in the continuation game. Together, these arguments imply (8).

For agent $i$'s per-period payoff at history $h_d^t$ to equal $E_{\sigma^*}\left[u_{i,t}\middle|\,h_d^t\right]$, it must be that

$$E_{\sigma^*}\left[w_{i,t}\middle|\,h_d^t\right] = E_{\sigma^*}\left[u_{i,t}+c(e_{i,t})-\frac{1}{1-\delta}\left(b_i\left(\phi_0\left(h_y^t\right)\right)-\delta U_i\left(\sigma^*,h_0^{t+1}\right)\right)\middle|\,h_d^t\right].$$

If $w_{i,t}<0$, then agent $i$ is only willing to pay if

$$E_{\sigma^*}\left[(1-\delta)\left(w_{i,t}-c(e_{i,t})\right)+b_i\left(\phi_0\left(h_y^t\right)\right)\middle|\,h_d^t\right] = U_i\left(\sigma^*,h_d^t\right) \geq \bar{U}_i\left(h_d^t\right),$$

implying the first line of (9).

Let $I=\{i|\,E_{\sigma^*}\left[w_{i,t}\middle|\,h_d^t\right]\leq 0\}$. Then the principal is only willing to pay $\sum_{i\notin I} w_{i,t}>0$ if

$$E_{\sigma^*}\left[(1-\delta)\left(\sum_{i=1}^N y_{i,t}-\sum_{i\notin I} w_{i,t}\right)-\sum_{i=1}^N\left(b_i\left(\phi_0\left(h_y^t\right)\right)-\delta U_i\left(\sigma^*,h_0^{t+1}\right)\right)+\delta\Pi\left(\sigma^*,h_0^{t+1}\right)\middle|\,h_d^t\right]\geq \bar{u}_0.$$

Plugging in $w_{i,t}$ and noting that $\sum_{i=1}^N y_{i,t}-\sum_{i=1}^N\left(u_{i,t}+c(e_{i,t})\right)=\pi_t$, we may rewrite this expression

$$\Pi\left(\sigma^*,h_d^t\right)\geq\sum_{i\in I}\left(E_{\sigma^*}\left[b_i\left(\phi_0\left(h_y^t\right)\right)-(1-\delta)\,c(e_{i,t})\middle|\,h_d^t\right]-\delta U_i\left(\sigma^*,h_d^t\right)\right)$$

If this expression holds for the crucial set of agents $I$, then a fortiori it holds for any other set of agents, implying the second line of (8).

2 : Define $\zeta\left(h^t\right)=\{D_{t'},d_{t'},e_{t'},y_{t'}\}_{t'=0}^t$. Given history $h_0^t\in\mathcal{H}_0^t$, consider a history $\tilde{h}_0^t\in\mathcal{H}_0^t$ such that $h_0^t$ and $\tilde{h}_0^t$ induce the same continuation games. We recursively construct $\sigma^*$ so that $U_i\left(\sigma^*,\tilde{h}_0^t\right)=U_i\left(\sigma,h_0^t\right)$ for all agents $i\in\{1,\ldots,N\}$ and $\Pi(\sigma^*,\tilde{h}_0^t)=\Pi(\sigma,h_0^t)$.

1. If $\tilde{h}_0^t$ is on-path for $\sigma^*$, then $\sigma^*$ specifies

    (a) For $D_t$, the public randomization device chooses $h_d^t\in\mathcal{H}_d^t$ according to $\sigma|\,\{h^t,D_t\}$.

(b) The principal chooses $d_t \in D_t$ as in $h_d^t$.

(c) Agent $i$'s wage equals $w_{i,t} = E_\sigma \left[ u_{i,t} + c(e_{i,t}) - \frac{1}{1-\delta} \left( b_i \left( \phi_0 \left( h_y^t \right) \right) - \delta U_i \left( \sigma, h_0^{t+1} \right) \right) \middle| \, h_d^t \right]$.

(d) The public randomization device chooses $h_c^t \in \mathcal{H}_c^t$ according to $\sigma | \, h_d^t$.

(e) Agent $i$ chooses $c(e_{i,t})$ as in $h_e^t$.

(f) Following realization of output $y_t$, agent $i$'s bonus equals

$$\tau_{i,t} = \frac{1}{1-\delta} E_\sigma \left[ b_i \left( \phi_0 \left( h_y^t \right) \right) - \delta U_i \left( \sigma, h_0^{t+1} \right) \middle| \, h_e^t, y_t \right].$$

(g) If no player deviates in period $t$, then $\left\{ \Pi(\sigma^*, \tilde{h}_0^{t+1}), \left\{ U_i \left( \sigma^*, \tilde{h}_0^{t+1} \right) \right\}_{i=1}^N \right\}$ is chosen according to $\sigma | \left\{ h_e^t, y_t \right\}$.

2. Following a publicly observed unilateral deviation by agent $i$, the principal chooses all future $d_{t'}$ to hold agent $i$ at $\bar{U}_i \left( h_0^t \right)$. Each agent $j$ chooses $a_{j,t} = 0$ and $w_{j,t} = \tau_{j,t} = 0$. Following a unilateral deviation by the principal, play as if agent 1 deviated. Following a simultaneous deviation by multiple players, play as if agent 1 deviated.

We claim $\sigma^*$ is a BFE. Consider an off-path history $\tilde{h}^t$. Agent $j$ earns no more than 0 if $a_{j,t} = 1$, which is not profitable because $\bar{U}_i \geq 0$. $\tau_{j,t} = w_{j,t} = 0$ is clearly optimal for each player. The principal is willing to choose the specified $d$, because her payoff is 0 regardless of the policy chosen. These punishments are therefore a BFE in which the principal and agent $i$ earn 0 and $\bar{U}_i \left( h_0^t \right)$ respectively.

Suppose $\tilde{h}_0^t$ is on-path. We want to show $(i)$ players earn $U_i \left( \sigma, h_0^t \right)$ by conforming to $\sigma^*$, and $(ii)$ players have no profitable one-shot deviation. For $(i)$, agent $i$'s payoff is

$$\begin{aligned} &(1-\delta) E_{\sigma^*} \left[ E_\sigma \left[ u_{i,t} + c(e_{i,t}) - \frac{1}{1-\delta} \left( b_i \left( \phi_0 \left( h_y^t \right) \right) - \delta U_i \left( \sigma, h_0^{t+1} \right) \right) - c_{i,t} \middle| \, h_d^t \right] \middle| \, \tilde{h}_0^t \right] \\ &+ E_{\sigma^*} \left[ E_\sigma \left[ b_i \left( \phi_0 \left( h_y^t \right) \right) - \delta U_i \left( \sigma, h_0^{t+1} \right) \middle| \, h_e^t \right] \middle| \, \tilde{h}_0^t \right]. \end{aligned}$$

Recall $h_d^t$ and $h_e^t$ have distributions $\sigma | \, h_0^t$ and $\sigma | \, h_d^t$, respectively. Moreover, $\sigma^* | \tilde{h}_e^t$ and $\sigma | h_e^t$ induce identical distributions over $y_t$. Applying Iterated Expectations, agent $i$'s payoff equals

$$(1-\delta) E_\sigma \left[ u_{i,t} \middle| \, h_0^t \right] + \delta E_\sigma \left[ U_i \left( \sigma, h_0^{t+1} \right) \middle| \, h_0^t \right] = U_i \left( \sigma, h_0^t \right),$$

as desired. Since $\sigma | \, h_0^t$ and $\sigma^* | \, \tilde{h}_0^t$ generate the same total surplus, the principal's continuation surplus must likewise equal $\Pi \left( \sigma, h_0^t \right)$.

Consider potential deviations by the players. The only variable that is not commonly observed is $e_t$. Players do not condition on past effort choices, so it suffices to check that there are no proftiable deviations at each public history. If $\tilde{h}_d^t$ is on-path for $\sigma^*|\,\tilde{h}_0^t$, then by an argument similar to above $U_i\left(\sigma, h_d^t\right) = U_i\left(\sigma^*, \tilde{h}_d^t\right)$ for all agents $i \in \{1, \ldots, N\}$ and $\Pi(\sigma, h_d^t) = \Pi(\sigma^*, \tilde{h}_d^t)$. Agents $i \in \{1, \ldots, N\}$ have no profitable deviation in $a_{i,t}$ because $U_i\left(\sigma, h_d^t\right) \geq \bar{U}_i(h_d^t)$ by (9). Similarly, the principal has no profitable deviation: setting $I = \emptyset$ in (2) implies $\Pi\left(\sigma, h_d^t\right) \geq 0$.

Consider deviations in the wage $w_{i,t}$. If $w_{i,t} < 0$, then agent $i$ earns $\bar{U}_i\left(\tilde{h}_d^t\right)$ following a deviation. But $\bar{U}_i\left(h_d^t\right) = \bar{U}_i\left(\tilde{h}_d^t\right)$ by construction. So agent $i$ has no profitable deviation, because $U_i\left(\sigma, h_d^t\right) \geq \bar{U}_i\left(h_d^t\right)$. Let $I = \{i \in \{1, \ldots, N\} |\, w_{i,t} \leq 0\}$. If the principal has any profitable deviation, then she has a profitable deviation in which $w_{i,t} = 0$ for all $i \notin I$. But this deviation is not profitable by an argument essentially identical to the argument in statement 1.

Agent $i$ chooses effort to maximize

$$e_{i,t} \in \underset{e_i \in \mathbb{R}_+}{\operatorname{argmax}}\, E_{\sigma^*}\left[(1-\delta)\left(\tau_{i,t} - c(e_i)\right) + \delta U_i\left(\sigma^*, h_0^{t+1}\right) \Big|\, \tilde{h}_w^t, e_{i,t} = e_i\right].$$

Applying the Law of Iterated Expectations and the definition of $\tau_{i,t}$ shows that this condition reduces to (2). So agents do not deviate from the specified effort.

Finally, consider deviations in $\{\tau_{i,t}\}_{i=1}^N$. If $\tau_{i,t} < 0$, agent $i$ has no profitable deviation by the first inequality in (8). Let $J = \{i \in \{1, \ldots, N\} |\, \tau_{i,t} \leq 0\}$. The principal has no profitable deviations as long as

$$-(1-\delta)\sum_{i \notin J} \tau_{i,t} + \delta E_{\sigma^*}\left[\Pi\left(\sigma^*, \tilde{h}_0^{t+1}\right) \Big|\, \tilde{h}_y^t\right] \geq \delta \bar{u}_0.$$

By construction, $E_{\sigma^*}\left[\Pi\left(\sigma^*, \tilde{h}_0^{t+1}\right) \Big|\, \tilde{h}_y^t\right] = E_{\sigma}\left[\Pi\left(\sigma, h_0^{t+1}\right) |\, h_y^t\right]$. So the second inequality in (8) implies that the principal has no profitable deviation.

## Completing Proof of Proposition 1

Towards contradiction: suppose a surplus-maximizing BFE $\sigma^*$ is not sequentially surplus-maximizing. We first define a strategy $\tilde{\sigma}$ that induces the same distribution over $\{D_t, d_t, e_t, y_t\}_{t=0}^{\infty}$ as $\sigma^*$, but with $U_i\left(\tilde{\sigma}, h^t\right) = \bar{U}_i\left(h^t\right)$ for all agents $i$ and $h^t \in \mathcal{H}_0^t$. Define $\tilde{\sigma}$ from $\sigma^*$ as in the recursive construction from Lemma A.1, with the sole exception that $\tau_{i,t} = 0$ in each period, anad

$$w_{i,t} = E_{\tilde{\sigma}}\left[c(e_{i,t}) + \frac{1}{1-\delta}\left(\bar{U}_i\left(h_d^t\right) - \delta \bar{U}_i\left(h_0^{t+1}\right)\right) \Big|\, h_d^t\right].$$

Then agent $i$'s continuation surplus equals $\bar{U}_i\left(h_d^t\right)$ at each on-path $h_d^t$.

Let $b_i^*$ be the reward scheme that satisfies (2), (8), and (9) for $\sigma^*$. Then $b_i^*$ satisfies these constraints for $\tilde{\sigma}$. In particular, Lemma A.1 applies and there exists a BFE $\tilde{\sigma}^*$ that induces the same distribution over $\{D_t, d_t, e_t, y_t,\}_{t=0}^{\infty}$ as $\tilde{\sigma}$. Recall that $\tilde{\sigma}$ and $\sigma^*$ generate the same total surplus, so $\tilde{\sigma}^*$ is a surplus-maximizing BFE. Because $\sigma^*$ is not sequentially surplus-maximizing, there exists some on-path history $h_0^t \in \mathcal{H}_0^t$ such that $\tilde{\sigma}^*|\, h_0^t$ is not surplus-maximizing.

Finally, consider a strategy profile $\bar{\sigma}$ that is identical to $\tilde{\sigma}^*$, except that the continuation strategy $\bar{\sigma}|\, h_0^t$ is surplus-maximizing. Because $h_0^t$ is reached on the equilibrium path, $\bar{\sigma}$ generates strictly higher total ex-ante expected surplus than $\tilde{\sigma}^*$. At any history inconsistent with or following $h_0^t$, $\bar{\sigma}$ clearly satisfies Lemma 1. If $h_0^{t'}$ is a predecessor to $h_0^t$, consider the reward scheme $\bar{b}_i = b_i^*$. This scheme immediately satisfies (2). All agents are held at their outside options in $\bar{\sigma}$, so the principal's payoff equals total expected continuation surplus minus agents' outside options. Agents' outside options at $h_0^{t'}$ are identical under $\tilde{\sigma}^*$ and $\bar{\sigma}$. Increasing the principal's payoff relaxes (8) and (9). Since the principal's continuation payoff is higher under $\bar{\sigma}$ than under $\tilde{\sigma}^*$, we conclude that $\bar{\sigma}$ satisfies Lemma 1. So $\sigma^*$ cannot be surplus-maximizing, which is a contradiction.

## Proof of Lemma 1

1 : Suppose $\sigma^*$ is a BFE and define $b_i : \phi_0\left(\mathcal{H}_y^t\right) \to \mathbb{R}$ as in (10). Fix a history $h_w^t \in \mathcal{H}_w^t$. Then (11) is a necessary condition for a BFE. Therefore, $b_i$ satisfies (2).

As in the proof of Lemma A.1, for any $h_y^t \in \mathcal{H}_y^t$,

$$b_i\left(\phi_0\left(h_y^t\right)\right) \geq E_{\sigma^*}\left[\bar{U}_i\left(h_0^{t+1}\right)\middle|\, h_y^t\right],$$

which implies the left-hand side of (3). Fix a set of agents $I \subset \{1, \ldots, N\}$ such that $\tau_{i,t} \geq 0$ for all $i \in I$. Suppose that following $h_y^t$, the principal does not pay $\{\tau_{i,t}\}_{i \in I}$. We claim the principal's continuation payoff after this deviation is bounded below by

$$E_{\sigma^*}\left[\Pi\left(\sigma^*, h_0^{t+1}\right) - \sum_{t'=0}^{\infty}\sum_{i\in I}(1-\delta)\,\delta^{t'}\left(y_{i,t+t'} - w_{i,t+t'} - \tau_{i,t+t'}\right)\middle|\, h_y^t\right]. \tag{12}$$

To prove this claim, consider the following strategy for the principal following a deviation observed by agents in $I$. Denote all variables that are observed by at least one agent $i \notin I$ by $\cup_{i\notin I}\phi_i\left(h_0^{t'}\right)$. In each period $h_0^{t'} \in \mathcal{H}_0^{t'}$, the principal plays according to $\sigma^*|\cup_{i\notin I}\phi_i\left(h_0^{t'}\right)$, with the sole exception that $w_{i,t'} = \tau_{i,t'} = 0$ for all $i \in I$. This strategy is identical to $\sigma^*$ except for transfer payments. Transfer payments do not affect the continuation game, so this strategy is feasible. Moreover, this strategy and $\sigma^*$ are indistinguishable for every agent $i \notin I$. Therefore, the principal's payoff from this strategy is no less than (12).

Hence, the principal is willing to pay $\{\tau_{i,t}\}_{i \in I}$ only if

$$(1-\delta)\sum_{i\in I} E_{\sigma^*}\left[\tau_{i,t} \middle| h_y^t\right] \leq E_{\sigma^*}\left[\sum_{t'=1}^{\infty}\sum_{i\in I}(1-\delta)\delta^{t'}\left(y_{i,t+t'} - w_{i,t+t'} - \tau_{i,t+t'}\right)\middle| h_y^t\right].$$

This expression holds for all subsets $I$ if and only if it holds for each agent $i$. Rearranging this expression yields the right-hand side of (3).

**2:** We construct a BFE $\sigma^*$ from $\sigma$. Define $\sigma^*$ as follows:

1. If $t = 0$, then $\tilde{h}_0^t = h_0^t = \emptyset$, the unique null history. Otherwise, begin with $h_0^t, \tilde{h}_0^t \in \mathcal{H}_0^t$ that induce identical continuation games.

2. If $\tilde{h}_0^t$ is on-path for $\sigma^*$ :

   (a) Following the realization of $D_t$, the principal chooses history $h_e^t \in \mathcal{H}_e^t$ using distribution $\sigma | \{h_0^t, D_t\}$. The principal chooses $d_t$ as in $h_e^t$. For each agent $i \in \{1, \ldots, N\}$, the principal pays

   $$w_{i,t} = E_\sigma\left[y_{i,t} - \frac{1}{1-\delta}\left(b_i\left(\phi_0\left(h_y^t\right)\right) - \delta S_i\left(\sigma, h_0^{t+1}\right)\right)\middle| h_e^t\right].$$

   Note $w_{i,t} \geq 0$ by (3). The principal sends a message to agent $i$ consisting of $(i)$ agent $i$'s effort in $h_e^t$, and $(ii)$ the reward scheme minus dyad-specific surplus for each $h_y^t$ that might follow $h_e^t$:

   $$m_{i,t} = \left\{a_{i,t}, e_{i,t}, \left\{b_i\left(\phi_0\left(h_y^t\right)\right) - \delta E_\sigma\left[\delta S_i\left(\sigma, h_0^{t+1}\right)\middle| h_y^t\right]\right\}_{h_y^t \in \text{supp}(\sigma | h_c^t)}\right\}.$$

   (b) Agent $i$ chooses $a_{i,t}, c_{i,t}$ as in $m_{i,t}$.

   (c) If output is $y_t$, then for each agent $i \in \{1, \ldots, N\}$,

   $$\tau_{i,t} = \frac{1}{1-\delta}\left(b_i\left(\phi_0\left(h_y^t\right)\right) - \delta E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right)\middle| h_e^t, y_t\right]\right)$$

   Note that $\tau_{i,t} \leq 0$ by (3), and $\tau_{i,t}$ can be perfectly inferred from $m_{i,t}$.

(d) Let $\tilde{h}_0^{t+1}$ be the realized history at the end of period $t$. The principal draws $h_0^{t+1} \in \mathcal{H}_0^{t+1}$ from $\sigma | \{h_e^t, y_t\}$.

3. If a deviation occurs in $\{w_{i,t}, m_{i,t}\}$, then agent $i$ takes his outside option in this and every subsequent period. If the principal observes a deviation by agent $i$, then $m_{j,t'} = w_{j,t} = 0$ for each agent $j \in \{1, \ldots, N\}$ in each future period. Agent $j$ chooses $a_{j,t'} = w_{j,t'} = \tau_{j,t'} = 0$, upon observing $m_{j,t'} = 0$. The principal chooses $d_t$ to min-max agent $i$.

We claim that no player has a profitable deviation from $\sigma^*$. Suppose $\tilde{h}_0^t$ is on-path. By construction, total continuation surplus is identical in $\sigma^* | \tilde{h}_0^t$ and $\sigma | h_0^t$.

First, consider the principal. For any on-path $\tilde{h}_d^t$ and each agent $i \in \{1, \ldots, N\}$,

$$E_{\sigma^*}\left[y_{i,t} - w_{i,t} - \tau_{i,t} | \tilde{h}_d^t\right] = 0.$$

Therefore, $\Pi\left(\sigma^*, \tilde{h}_d^t\right) = 0$. Following a deviation in $d_t$, the principal's payoff is also 0. So the principal has no profitable deviation in $d_t$.

At $\tilde{h}_d^t$, suppose the principal follows the equilibrium wage-message combination $(w_{i,t}, m_{i,t})$. Then she earns 0 continuation profit. Suppose she deviates to some other $(\hat{w}_{i,t}, \hat{m}_{i,t})$. Then either $(i)$ there exists some history $\hat{h}_d^t$ such that $(\hat{w}_{i,t}, \hat{m}_{i,t})$ is on-path at $\hat{h}_d^t$ and $\phi_i\left(\hat{h}_d^t\right) = \phi_i\left(\tilde{h}_d^t\right)$, or $(ii)$ not.

If $(i)$, then by construction $E_{\sigma^*}\left[y_{i,t} - \hat{w}_{i,t} - \tau_{i,t} | \hat{h}_w^t\right] = 0$. Moreover, neither messages nor wages affect the continuation game, so the principal's continuation surplus remains 0. Therefore, this deviation is not profitable. If $(ii)$, then agent $i$ chooses $a_{i,t} = 0$. The principal earns 0 from that agent in this period and 0 in future periods. Again, this deviation is not profitable. So no profitable deviation from $(w_{i,t}, m_{i,t})$ exists. $\tau_{i,t} \leq 0$ so the principal has no profitable deviation from $\tau_{i,t}$.

Consider agent $i$. At each on-path history $\tilde{h}_0^t$, $E_{\sigma^*}\left[u_{i,t} | \tilde{h}_0^t\right] = E_{\sigma^*}\left[y_{i,t} - c(e_{i,t}) | \tilde{h}_0^t\right]$. Therefore, agent $i$'s continuation payoff is $U_i\left(\sigma^*, \tilde{h}_0^t\right) = S_i\left(\sigma^*, \tilde{h}_0^t\right)$. By construction of $\sigma^*$, $S_i\left(\sigma^*, \tilde{h}_0^t\right) = S_i\left(\sigma, h_0^t\right)$.

Since $w_{i,t} \geq 0$, agent $i$ has no profitable deviation from $w_{i,t}$. At any on-path history $\tilde{h}_c^t$ consistent with $(w_{i,t}, m_{i,t})$, agent $i$ earns

$$E_{\sigma^*}\left[(1-\delta)\tau_{i,t} + \delta S_i(\sigma^*, \tilde{h}_0^{t+1}) \middle| \tilde{h}_c^t\right] = E_{\sigma^*}\left[b_i\left(\phi_0\left(h_y^t\right)\right) | \tilde{h}_c^t\right].$$

The equality follows by construction of $\sigma^*$: at any on-path $\tilde{h}_e^t$, $U_i\left(\sigma^*, \tilde{h}_0^{t+1}\right) = S_i\left(\sigma^*, h_0^{t+1}\right)$ and $E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right) \middle| h_e^t, y_t\right] = E_{\sigma^*}\left[S_i\left(\sigma^*, \tilde{h}_0^{t+1}\right) \middle| \tilde{h}_e^t, y_t\right]$. But $b_i$ satisfies (2) at every on-path history consistent with $m_{i,t}$, so agent $i$ has no profitable deviation from his accept/reject decision and effort regardless of his beliefs about the history $\tilde{h}_e^t$. Off the equilibrium path, continuation play is independent of history and so $a_{i,t} = 0$ is optimal regardless of agent $i$'s beliefs.

Following any deviation in $\tau_{i,t} < 0$, agent $i$ earns continuation surplus $\bar{U}_i\left(\tilde{h}_0^{t+1}\right)$. If agent $i$ believes the on-path history is $\tilde{h}_y^t$, then he is willing to pay $\tau_{i,t}$ if

$$-(1-\delta)\tau_{i,t} \leq \delta E_{\sigma^*}\left[U_i\left(\sigma^*, \tilde{h}_0^{t+1}\right) - \bar{U}_i\left(\tilde{h}_0^{t+1}\right) \middle| \tilde{h}_y^t\right] = \delta E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right) - \bar{U}_i\left(h_0^{t+1}\right) \middle| h_y^t\right].$$

The equality follows because $\sigma^*| \tilde{h}_y^t$ and $\sigma| h_y^t$ induce identical distributions over $S_i$, and $U_i\left(\sigma^*, \tilde{h}_0^{t+1}\right) = S_i\left(\sigma^*, \tilde{h}_0^{t+1}\right)$. But at any such history $\tilde{h}_y^t$, agent $i$ can infer $\tau_{i,t}$ from $m_{i,t}$. Plugging in $\tau_{i,t}$, this inequality holds as long as

$$-\left(b_i\left(h_y^t\right) - \delta E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right) \middle| h_y^t\right]\right) \leq \delta E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right)\right] - \delta \bar{U}_i\left(\tilde{h}_0^{t+1}\right)$$

or $b_i\left(h_y^t\right) \geq \bar{U}_i\left(\tilde{h}_0^{t+1}\right)$. But $\bar{U}_i\left(\tilde{h}_0^{t+1}\right) = \bar{U}_i\left(h_0^{t+1}\right)$ because $\tilde{h}_0^{t+1}$ and $h_0^{t+1}$ induce the same continuation game. So this inequality is implied by (3). Off the equilibrium path, agent $i$'s payoff is independent of $\tau_{i,t}$ and so he chooses $\tau_{i,t} = 0$. So agent $i$ has no profitable deviation from $\tau_{i,t}$, regardless of his beliefs about the true history.

We conclude that $\sigma^*$ is a BFE with the desired properties.

## Proof of Proposition 2

(1) : If $\sum_{i=1}^N d_{i,t} < 1$ in some period $t$, consider the alternative decision $\tilde{d}_t$ with $\sum_{i=1}^N \tilde{d}_{i,t} = 1$ and $\tilde{d}_{i,t} \geq d_{i,t}$ for all $i$. This alternative generates strictly higher surplus holding efforts fixed. It also weakly relaxes all dynamic-enforcement constraints.

(2) : By Lemma 1, the decision $d_0$ in the first period does not affect any incentive-compatibility or dynamic-enforcement constraints (2) or (3). Suppose there exist $i, j$ for which (6) fails for $t = 0$. Because $E\left[y_i | e_{i,0}, \omega_{i,0}\right]$ is strictly concave in $d_i$, total surplus is strictly higher under an alternative decision that satisfies (6). Such an alternative exists by the Intermediate Value Theorem. So (6) holds for $t = 0$ in any surplus-maximizing relational contract. But $\sigma^*| h_0^t$ is surplus-maximizing for any on-path $h_0^t \in \mathcal{H}_0^t$ if $\sigma^*$ is sequentially surplus-maximizing. Therefore, (6) holds at every on-path $h_e^t$.

$(3)$ : Let $\sigma^*$ be a surplus-maximizing relational contract. Suppose towards contradiction that $\sigma^*$ is sequentially surplus-maximizing.

Consider period $t' < t$ such that $0 < e_{i,t'} < e_i^{FB}(d_{i,t'}, \theta_{t'})$. Because $F_i$ satisfies MLRP and CDFC, we can replace the IC constraint (2) in this period with its first-order condition:

$$(1-\delta)\, c'\left(e_{i,t'}\right) = \int_0^{\infty} E_{\sigma^*}\left[b_i | h_e^t, y_{i,t'}\right] \frac{\partial p_i}{\partial e_{i,t'}}\left(y_{i,t'} \middle| d_{i,t'}, \theta_{t'}, e_{i,t'}\right) dy_{i,t'}. \tag{13}$$

Together with the dynamic-enforcement constraint (3), (13) implies that the optimal reward scheme $b_i\left(y_t\right)$ is the "step function,"

$$b_i = \begin{cases} \delta \bar{U}_i(h_0^{t'+1}) & y_{i,t'} < y_i^*(d_{i,t'}, \theta_{t'}, e_{i,t'}) \\ \delta S_i(\sigma^*, h_0^{t'+1}) & y_{i,t'} \geq y_i^*(d_{i,t'}, \theta_{t'}, e_{i,t'}) \end{cases} \tag{14}$$

Note that $\bar{U}_i(h_0^{t'+1})$ depends only on $\{\theta_{t''}\}_{t''=0}^{t'}$, and in particular is independent of the actions of any player. We claim that increasing the upper bound of the dynamic-enforcement constraint (3) by $\varepsilon > 0$ for all $y_{i,t'} > y_i^*\left(d_{i,t'}, e_{i,t'}\right)$ leads to strictly higher effort. The IC constraint (13) may be written

$$\frac{1-\delta}{\delta} c'\left(e_{i,t}\right) = \int_0^{y_i^*(d_{i,t}, e_{i,t})} \bar{U}_i(h_0^{t'+1}) \frac{\partial p_i}{\partial e_{i,t}}\left(y_i \middle| d_{i,t}, e_{i,t}\right) dy_i + \int_{y_i^*(\omega_{i,t}, e_{i,t})}^{\infty} S_i \frac{\partial p_i}{\partial e_{i,t}}\left(y_i \middle| d_{i,t}, e_{i,t}\right) dy_i.$$

Increasing $S_i$ implies that the left-hand side is strictly smaller than the right-hand side in this expression. By assumption, $c'\left(e\right)$ and $\frac{\partial p_i}{\partial e}$ are differentiable in $e$. Similarly, $y_i^*\left(d_{i,t'}, \theta_{t'}, e_{i,t'}\right)$ is differentiable in $e_{i,t'}$ by the Implicit Function Theorem. Therefore, there exists some $e' > e_{i,t'}$ which satisfies the first-order IC constraint. Define $e_i\left(S\right)$ as the minimum of $e_i^{FB}$ and the solution to the first-order IC constraint given maximal reward $S$. By a combination of Leibniz Rule and the Implicit Function Theorem, it can be shown that $e_i\left(S\right)$ is differentiable for all $e_i\left(S\right) < e_i^{FB}$.

Now, suppose there exists a set of period-$t$ histories $E_t$ satisfying the conditions of the Proposition such that $\Pr_{\sigma^*}\{E_t\} > 0$ and (6) holds. At all of these histories, consider the alternative decision $\tilde{d}_{i,t} = d_{i,t} + \varepsilon$, $\tilde{d}_{j,t} = d_{j,t} - \varepsilon$, and $\tilde{d}_{k,t} = d_{k,t}$ for all $k \neq i, j$. This alternative is feasible for small $\varepsilon > 0$, because $\lim_{d_i \to 0} \frac{\partial}{\partial \omega_i} E[y_i | d_i, \theta_t, e_i] = \infty$ and hence $d_{i,t}$ and $d_{j,t}$ must be interior to satisfy (6). Moreover, this alternative leads to the same optimal reward scheme in each period for any $k \neq i$ : for $k \neq j$, this is obvious, and for $k = j$, it follows from the optimal reward scheme (14) and the fact that $y_{j,t} < y_j^*\left(d_{j,t'}, \theta_{t'}, e_{j,t'}\right)$ in all $t' < t$. This alternative strictly relaxes agent $i$'s dynamic-enforcement constraint (3) in period $t'$ because $h_e^{t'}$ precedes $h_0^t$ and hence $\Pr_{\sigma^*}\left\{E_t | h_e^{t'}\right\} > 0$.

This perturbed decision weakly relaxes all dynamic-enforcement constraints. From the perspective of period 0, the cost equals the loss from perturbing the decision in period $t$ :

$$K(\varepsilon) = \begin{array}{c} \\ \delta^t (1-\delta) \left\{ E[y_{i,t} + y_{j,t} | E_t, d_t, \theta_t, e_t] - E[y_{i,t} + y_{j,t} | E_t, \tilde{d}_t, \theta_t, e_t] \right\} \Pr_{\sigma^*} \{E_t\} \end{array} > 0.$$

Agents $i$ and $j$ have independent output, so this perturbation changes agent $i$'s optimal reward scheme (14) in period $t'$ to

$$\tilde{b}_i(y_{i,t'}) = \begin{cases} \delta \bar{U}_i & y_{i,t'} \leq y_i^* (d_{i,t'}, \theta_{t'}, e_{i,t'}) \\ S_i + \xi(\epsilon) & \text{otherwise.} \end{cases}$$

where $\xi(\epsilon) = (1-\delta)\,\delta^t \{E[y_{i,t}|E_t, \tilde{d}_{i,t}, \theta_t, e_{i,t}] - E[y_{i,t}|E_t, d_{i,t}, \theta_t, e_{i,t}]\} \Pr_{\sigma^*}\left[E_t | h_e^{t'}, y_{t'} \geq y_i^*(d_{i,t'}, \theta_{t'}, e_{i,t'})\right]$. Note that $\xi(\varepsilon) > 0$ is a smooth function of $\epsilon$.

As shown above, increasing $S_i$ by $\xi(\varepsilon) > 0$ strictly and smoothly increases effort to $e_i(S + \xi(\varepsilon)) > e_i(S)$. The value of this increase in effort equals

$$B(\varepsilon) = [E[y_i | d_{i,t'}, \theta_{t'}, e_i(S+\xi)] - c(e_i(S+\xi))] - [E[y_i | d_{i,t'}, \theta_{t'}, e_i(S)] - c(e_i(S))].$$

Since $e_i(S) < e_i(S+\xi) < e_i^{FB}$ for $\xi$ sufficiently small, this expression is strictly positive. The benefits of this perturbed policy exceed the costs as long as

$$\begin{aligned} & \frac{1}{\varepsilon} \{[E[y_i | d_{i,t'}, \theta_{t'}, e_i(S + \xi(\varepsilon))] - c(e_i(S + \xi(\varepsilon)))] - [E[y_i | \omega_{i,t'}, \theta_{t'}, e_i(S)] - c(e_i(S))]\} \\ > \; & \frac{1}{\varepsilon} \delta^t (1-\delta) \{E[y_{i,t} + y_{j,t} | E_t, d_t, \theta_t, e_t] - E[y_{i,t} + y_{j,t} | E_t, \tilde{d}_t, \theta_t, e_t]\}. \end{aligned}$$

As $\varepsilon \to 0$, the costs converge to 0 by (6). The benefits converge to

$$\frac{\partial \left( E[y_i | d_{i,t'}, e_i(S)] - c(e_i(S)) \right)}{\partial e} \frac{\partial e_i}{\partial S_i} \frac{\partial \xi}{\partial \varepsilon} > 0,$$

where the first term is strictly positive because $e_i(S + \xi) < e^{FB}$ and the second and third terms are strictly positive by the argument above. So the perturbed decision leads to a strictly higher ex ante total surplus.

## Proof of Proposition 3

Let $S^{R2} = \alpha R - c$, $S^{R1} = R - c$, and $S^{Wj} = (1-\delta)(W - c) + \delta(\rho S^{Rj} + (1-\rho) S^{Wj})$ for $j \in \{1, 2\}$. Note that $S^{W2} < S^{W1} < S^{R2} < S^{R1}$ by assumption.

We claim that $d_t = 1$ in any period with $\theta_t = W$. Let $T$ be the first period in which $d_T = 2$ if $\theta_t = W$. We proceed by induction on $T$. If $T = 0$, then either at least one worker works hard or neither do. If neither do, then the payoff is identical if $d_0 = 1$. If at least one agent works hard, then the payoff is strictly higher if $d_0 = 1$ by assumption. So $d_0 = 1$. Similarly, if $d_T = 2$ for the first time in period $T > 0$ and neither agent works hard, then both total and dyad-surplus is the same if $d_T = 1$. If at least one agent works hard, consider setting $d_T = 1$. This change increases total surplus. It also relaxes agent 1's dynamic enforcement constraints. But $d_{t'} = 1$ for all $t' < T$, so only agent 1 worked hard in previous periods. So this change strictly increases total surplus and relaxes all relevant dynamic enforcement constraints.

Now, define $\bar{\delta}$ as the solution to

$$c = \frac{\bar{\delta}}{1-\bar{\delta}} S^{W2}.$$

For $\delta < \bar{\delta}$, agent 1 cannot be motivated to work hard if $d_t = 2$ whenever $\theta_t = R$. One option is $e_{1,t} = 0$ whenever $\theta_t = W$. Define $\underline{\delta}$ as the solution to

$$c = \frac{\underline{\delta}}{1-\underline{\delta}} S^{R2}.$$

Because $S^{R2} > S^{W2}$, $\underline{\delta} < \bar{\delta}$. For the rest of the proof, consider $\delta \in (\underline{\delta}, \bar{\delta})$

Suppose that an equilibrium in which $e_{1,t} = 0$ whenever $\theta_t = W$ is not efficient. Then $d_t = 1$ in some period such that $\theta_t = R$. Consider a history $h_0^t$ such that (i) $\theta_t = R$ for the first time in period $t$, and (ii) $d_{t'} = 1$ with positive probability in some $t'$ following $h_0^t$. Define $\chi_{t'} = \Pr\{d_{t'} = 1 | h_0^t\}$ for all $t' \geq t$. Then total continuation surplus following history $h_0^t$ is bounded above by

$$\sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)(\chi_{t'} S^{R1} + 2(1-\chi_{t'})S^{R2}),$$

while 1-dyad surplus is bounded above by

$$\sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)(\chi_{t'} S^{R1} + (1-\chi_{t'})S^{R2}).$$

This bound on total surplus may be rewritten

$$(1-\delta)S^{R1} \sum_{t'=t}^{\infty} \delta^{t'-t} \chi_{t'} + 2(1-\delta)S^{R2} \sum_{t'=t}^{\infty} \delta^{t'-t}(1-\chi_{t'})$$

with a similar expression for the bound on 1-dyad surplus.

Note that

$$\sum_{t'=t}^{\infty}\delta^{t'-t}\chi_{t'}+\sum_{t'=t}^{\infty}\delta^{t'-t}(1-\chi_{t'})=\frac{1}{1-\delta}.$$

Therefore, consider the alternative continuation equilibrium: with probability $\chi\equiv(1-\delta)\sum_{t'=t}^{\infty}\delta^{t'-t}\chi_{t'}$, $d_{t'}=1$ in every $t'\geq t$. Otherwise, $d_{t'}=2$ in every $t'\geq t$. Because $\delta>\underline{\delta}$, both agents are willing to work hard in every $t'\geq t$ in this alternative equilibrium. So this alternative attains the upper bound on both total and 1-dyad surplus. Given that $d_t=1$ whenever $\theta_t=1$, it suffices to consider continuation equilibria of this kind once demand becomes robust.

Finally, we argue that for $\delta$ sufficiently near $\bar{\delta}$, the following equilibrium is surplus-maximizing:

- If $\theta_t=W$, then $d_t=1$ and $e_{1,t}=1$.
- In the first period $t$ such that $\theta_t=R$, $d_t=1$ with probability $\chi\in(0,1]$.
- In every subsequent period $t'\geq t$, $d_{t'}=d_t$.

Given the previous arguments, it suffices to show that (i) if $d_t=1$ when $\theta_t=R$ is surplus-maximizing, then $e_{1,t}=1$ and $S_1$ is the same in each period with $\theta_t=W$, and (ii) $d_t=1$ when $\theta_t=R$ is surplus-maximizing.

For (i), relax the problem so that agent 1's dynamic enforcement constraint must only hold the first time he chooses $e_{1,t-1}=1$. Then $i$-dyad surplus from $t$ onwards may be written

$$\sum_{t'=t}^{\infty}\delta^{t'-t}\left((1-\rho)^{t'-t}\left[(1-\rho)(1-\delta)(W-c)+\rho\delta(\gamma_{t'}S^{R1}+(1-\gamma_{t'})S^{R2})\right]\right)$$

or

$$\frac{(1-\rho)(1-\delta)}{1-\delta(1-\rho)}(W-c)+\delta\rho S^{R1}\sum_{t'=t}^{\infty}\delta^{t'-t}(1-\rho)^{t'-t}\gamma_{t'}+\delta\rho S^{R2}\sum\delta^{t'-t}(1-\rho)^{t'-t}(1-\gamma_{t'}).$$

Now, define

$$\gamma\equiv(1-\delta(1-\rho))\sum_{t'=t}^{\infty}\delta^{t'-t}(1-\rho)^{t'-t}\gamma_{t'}.$$

Note that $\gamma\in[0,1]$. Consider the equilibrium in which $\gamma_{t'}=\gamma$ in every period $t'\geq t$. This alternative equilibrium satisfies agent 1's dynamic enforcement constraint in $t$, as well as in every $t' \geq t$. Therefore, this alternative generates at least as much total surplus as the original equilibrium.

As $\delta \to \bar{\delta}$, $\gamma \to 0$ satisfies agent 1's dynamic enforcement constraint. Therefore, for $\delta$ sufficiently close to $\bar{\delta}$, $e_{1,t} = 1$ while $\theta_t = W$. This proves the claim.

## Proof of Proposition 4

In period $t$ and given reward scheme $b_i$, agent $i$ chooses $e_{i,t} = 1$ if

$$\int_0^\infty (b_i(x)[p_i(x|e = 1, d_t) - p_i(x|e = 0, d_t)] - (1 - \delta)c)\, dx \geq 0.$$

Let $S_i(y)$ equal $i$-dyad surplus following output $y$. Then by Lemma 1, $b_i$ must satisfy $0 \leq b_i(y) \leq S_i(y)$ in equilibrium. Define $y_i^*$ as the unique output such that $L_i(y_i^*|d \neq i) = 1$. Then if $d_t \neq i$, agent $i$'s incentive constraint is satisfied only if

$$\frac{\delta}{1-\delta} \int_0^\infty \int_{y_i^*}^\infty \delta S_i(x)[p_i(x_i|e = 1, d_t) - p_i(x_i|e = 0, d_t)]dx_i dx_{-i} \geq c.$$

Define $S_i^P = \int_0^\infty xp_i(x|e_i = 1, d = i)dx - c$, $S_i^1 = \int_0^\infty xp_i(x|e_i = 1, d \neq i)dx - c$, and $S^0 = \int_0^\infty xp_i(x|e = 0, d \neq i)dx$. Define $\bar{\delta}$ as the largest discount factor for which agent 2's IC constraint holds for $S_2 = S_2^1$. For $\delta < \bar{\delta}$, agent 2 is only willing to work hard if he expects to be promoted with positive probability. Therefore, if $d_2 = 1$ with probability 1, then agent 2 shirks in each period on the equilibrium path.

How can agent 2 be motivated if $\delta < \bar{\delta}$? agent 2 chooses $e_{2,t} = 0$ in $t \geq 1$ if $d_1 = 1$. Suppose that agent 1 is willing to chooses $e_{1,t} = 1$ for $t \geq 1$ even if $d_1 = 2$. If $\xi > 0$, then there exists an open interval of discount factors $\delta < \bar{\delta}$ that satisfy this condition. Then agent 2 should be promoted with probability 1 if

$$S_1^1 + S_2^P > S_1^P + S_2^0$$

This inequality holds if $\xi$ is not too large.

Now, suppose $e_{1,t} = 0$ for $t \geq 1$ if $d_1 = 2$. Let $\rho(y) \in [0, 1]$ be the probability that $d_1 = 1$ following outcome $y = (y_1, y_2)$. Total surplus for $t \geq 1$ is increasing in $\rho(y)$. Therefore, the surplus-maximizing relational contract solves

$$\begin{aligned} \max_{\rho:\mathbb{R}^2 \to [0,1]} \int_0^\infty \int_0^\infty \rho(x) p_1(x_1|e &= 1, d \neq 1) p_2(x_2|e = 1, d \neq 2) dx_1 dx_2 \\ s.t. \text{ IC for each agent given } d_0 &= 0 \end{aligned}$$

If $y_2 < y_2^*$, then $\rho(y) = 1$ is optimal because it maximizes the objective function, relaxes agent 1's dynamic enforcement constraint, and does not affect agent 2's dynamic enforcement constraint.

The Lagrangian for this constrained optimization problem can be solved separately for each $\rho(y)$. Doing so yields the following first-order expression:

$$1 + \lambda_1 \left( \frac{\delta(S_1^P - S_1^0)}{1-\delta} \left[ 1 - \frac{1}{L_1(y_1 | d \neq 1)} \right] \right) + \lambda_2 \left( \frac{\delta(S_2^0 - S_2^P)}{1-\delta} \left[ 1 - \frac{1}{L_2(y_2 | d \neq 2)} \right] \right)$$

where $\lambda_i$ is the multiplier on agent $i$'s dynamic enforcement constraint. This expression is constant in $\rho$. If it is negative, then $\rho(y) = 0$ optimally. If it is positive, then $\rho(y) = 1$ optimally. Rearranging, we have the desired condition for $\rho(y) = 1$.

If $S_1^P - S_2^P < \frac{1-\delta}{\delta} E[y_2 - c | e_2 = 1, d \neq 2]$, the equilibrium in which both agents work hard in period 1 dominates the equilibrium in which only agent 1 works hard. This condition is satisfied if $\Delta$ is not too large.

## Proof of Proposition 5

Suppose $d_0 = 2$ in $\sigma^*$. Let $t > 0$ be the first period in which $d_t = 1$ with positive probability. Consider replacing the continuation equilibrium with $\sigma^*$. This generates weakly larger total surplus, because $\sigma^*$ is surplus-maximizing. It also weakly relaxes all relevant dynamic-enforcement constraints, since $d_{t'} = 2$ for all $t' < t$. Hence, if $d_0 = 2$, then one surplus-maximizing relational contract entails $d_t = 2$ in each period $t$. This relational contract has ex ante total expected surplus $(pR - 1)\, c$.

Define $\delta_1 < 1$ by

$$\frac{c}{p} = \frac{\delta_1}{1-\delta_1} (pL - 1)\, c.$$

For $\delta \geq \delta_1$, if $d_t = i$ in every period then agent $i$ is willing to choose $e_{i,t} = 1$, regardless of $\theta_t$. Define $\delta_2 > 0$ by

$$\frac{c}{p} = \frac{\delta_2}{1-\delta_2} (1-q)\,(pH - 1)\, c.$$

Let $\delta < \delta_2$, and suppose the policy specifies $d_t = 1$ when $\theta_t = qL + (1-q)\, H$ and $d_{t'} = 2$ in any $t' > t$ such that $\theta_{t'} = L$. Then agent 1 is unwilling to work hard if $\theta_t = qL + (1-q)\, H$. Since $(1-q)\,(pH-1) < pL - 1$, $\delta_2 > \delta_1$.

Suppose $\delta_1 \leq \delta < \delta_2$. Consider the following policy: $d_0 = 1$. If $y_{0,1} = \theta_0 c$, then with probability $\xi$, $d_t = 1$ in every future period $t$. With probability $1 - \xi$, $d_t = 1$ if $\theta_t = H$ and otherwise, $d_t = 2$ in each period $t$. If $y_{0,1} = 0$, then $d_t = 1$ if $\theta_t = H$ and $d_t = 2$ otherwise. $\xi$ is chosen to solve

$$\frac{c}{p} = \frac{\delta}{1-\delta}\left(\left(1-q\right)\left(pH-1\right) + q\xi\left(pL-1\right)\right)c.$$

Since $\delta \geq \delta_1$, $\xi \in (0,1)$. Under this alternative strategy, agent 1 is willing to work hard in period $t = 0$. Since $qL + (1-q)H \geq R$, this equilibrium dominates an equilibrium with $d_0 = 2$ if

$$\left(1-q\right)\left(pH-1\right) + pq\xi\left(pL-1\right) > \left(1-pq\left(1-\xi\right)-\left(1-p\right)q\right)\left(pR-1\right)c.$$

$\xi \to 0$ as $\delta \to \delta_2$ and $pH > pR$, so there exists some $\underline{\delta} > \delta_2$ such that this inequality holds if $\delta > \underline{\delta}$. Thus, if $\underline{\delta} < \delta < \delta_2$, then $d_0 = 1$ in any surplus-maximizing relational contract.

The result $\Pr_{\sigma^*}\left[\theta_{t'} = B, d_{t'} = 1\right] > 0$ if $y_{1,0} = \theta_0 c$ follows immediately from $\delta < \delta_2$. Suppose $y_{1,0} = 0$. Then $b_1 = \delta\bar{u}_1 = 0$ in the optimal reward scheme, so dyad surplus is irrelevant in the continuation game. Hence, a surplus-maximizing continuation equilibrium can be chosen. In every such equilibrium, $d_t = 1$ if and only if $\theta_t = H$, as desired.

# 8 References


Andrews, I. and Barron, D., "The Allocation of Future Business: Dynamic Relational Contracts with Multiple Agents," *Working Paper* (2013).

Ariely, D., Belenzon, S., and Tsolmon, U., "Health Insurance and Relational Contracts in Small American Firms." *Working Paper* (2013).

Asanuma, B., "Manufacturer-Supplier Relationships in Japan and the Concept of Relation-Specific Skill," *Journal of the Japanese and International Economies*, vol. 3, no. 1, 1-30 (1989).

Baker, G., Gibbons, R., and Murphy, K., "Subjective Performance Measures in Optimal Incentive Contracts," *The Quarterly Journal of Economics*, Vol. 109, No. 4, 1125-1156 (1994).

Board, S., "Relational Contracts and the Value of Loyalty," *The American Economic Review*, Vol. 101, No. 7, 3349-3367 (2011).

Bull, C., "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, vol. 102, no. 1, 147-159 (1987).

Calzolari, G. and Spagnolo, G., "Relational Contracts and Competitive Screening," *Working Paper* (2011).

Ely, J., Horner, J., and Olszewski, W., "Belief-Free Equilibria in Repeated Games," *Econometrica*, Vol. 73, No. 2, 377-415 (2005).

Fong, Y. and Li, J., "Relational Contracts, Efficiency Wages, and Employment Dynamics," *Working Paper* (2012).

Fudenberg, D., Holmstrom, B., and Milgrom, P., "Short-Term Contracts and Long-Term Agency Relationships," *Journal of Economic Theory*, Vol. 51, no. 1, 1-31 (1990).

Graham, J., Harvey, C., and Puri, M., "Capital Allocation and Delegation of Decision-Making Authority within Firms," *Working Paper* (2013).

Halac, M., "Relational Contracts and the Value of Relationships," *The American Economic Review*, Vol. 102, No. 2, 750-759 (2012).

Haltiwanger, J., Jarmin, R., and Miranda, J., "Who Creates Jobs? Small Versus Large Versus Young." *The Review of Economics and Statistics,* Vol. 95, No. 2, 347-361 (2013).

Kandori, M., "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory*, Vol. 102, No. 1, 1-15 (2002).

Kranz, S., "Discounted Stochastic Games with Voluntary Transfers," *Working Paper* (2011).

Levin, J., "Relational Incentive Contracts," *American Economic Review*, vol. 93, no. 3, 835-857 (2003).

Li, J., Matouschek, N., and Powell, M., "The Burden of Past Promises," *Working Paper* (2014).

MacLeod, B. and Malcomson, J., "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 447-480 (1989).

Malcomson, J., "Relational Contracts with Private Information," *Working Paper* (2014)

Malcomson, J., "Relational Incentive Contracts," in *The Handbook of Organizational Economics*, R. Gibbons and J. Roberts (eds.), Princeton, NJ: Princeton University Press

Peter, L., and Hull, R. *The Peter Principle.* New York: Morrow (1968)

Segal, I., "Contracts with Externalities," *The Quarterly Journal of Economics*, Vol. 114, No. 2, 337-388 (1999).