Policy Commitments in Relational Contracts

Daniel Barron
Northwestern University/Yale University

Michael Powell
Northwestern University

April 2014

Preliminary and Incomplete

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 a variety of simple examples. For instance, hiring may lag demand following a recession, investments may be delayed and awarded inefficiently, 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 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. 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 decrease continuation surplus and weaken the principal's incentives to uphold her promises, and therefore have no place in a surplus-maximizing relationship. The principal "settles up" with her agents at the end of each period, while future decisions are made to maximize continuation surplus.

In contrast, backward-looking policies arise naturally if relationships are bilateral–that is, if each agent cannot observe the actions of 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. Future decisions biased toward an individual therefore complement more generous reward schemes for that individual but also negatively affect the firm's overall future performance.

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 the next section.

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 smooth repeated games 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 delay employee-specific investments and bias those investments towards workers who performed well in the past. 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. 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. Similarly, Malcomson (2014), Halac (2012), 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 Board (2011) and 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 Hiring Decisions

We now informally introduce the key ideas of our model in the context of a hiring decision, deferring a formal analysis of this example until Section 6.

A firm anticipates a permanent increase in the demand for its products at some point in the future. This firm currently employs a single worker, whom it has to motivate to exert effort. Hiring a second worker after demand increases would maximize total continuation surplus. Therefore, if a firm could write a formal bonus contract on effort, it would ask for high effort in each period and immediately hire a second worker once demand increases.

Suppose that effort and output are not contractible and the firm and its workers engage in a relational contract. We will show the following two results.

1. If agents observe every relationship, the firm will employ two workers as soon as demand increases.

2. If each agent observes only his own relationship, the firm may continue to employ only one worker after demand increases. However, a firm that starts with high demand immediately employs two workers.

The first result highlights the intuition that if relationships are public, then the firm's hiring decisions are never biased. In contrast, the second result shows that the optimal hiring policy might be backward-looking if relationships are bilateral. In particular, the firm's hiring decisions exhibit history-dependent biases: a firm that has suffered from weak demand is slower to hire workers than one that begins with strong demand, even if both firms currently face identical environments.

More formally, suppose a risk-neutral firm interacts repeatedly with two risk-neutral workers. In each period, the firm can choose to hire worker 1 or to hire both workers 1 and 2. Workers can accept the firm's offer or reject it in favor of an outside option that yields utility $\bar{u} > 0$. If worker i accepts the job, he chooses whether or not to exert observable effort at cost 1. Worker i generates a sale with a probability that depends on three factors: his effort choice, the demand for the firm's product, and the number of workers the firm employs in that period. Worker i generates a sale with probability 1 if alone or with probability $\alpha < 1$ if there are two workers. A sale generates revenue W if demand is weak and R if demand is robust, with $R > W$. Demand is initially weak, but in each period it becomes permanently robust with positive probability. Sales by each worker are independent. We assume $\alpha W > 1$, so that it is always efficient for workers to exert effort.

If the firm can write formal bonus contracts, then it can motivate each worker by paying 1 if that worker exerts effort. The optimal formal contract maximizes total surplus and uses wage payments to hold agents at their outside options. So long as $R > \frac{\bar{u}+1}{2\alpha-1} > W > \bar{u} + 1$, the firm optimally hires one worker if demand is weak and two workers if demand is robust.

If the firm cannot write formal contracts, then any payments the firm promises have to be self-enforcing. If relationships are public, both workers can take their outside options if the firm reneges on a promised payment to any individual worker. Therefore, the firm stands to lose its entire continuation surplus if it reneges on a relationship. As in Levin (2003), a firm can credibly promise a payment only if is smaller than that the future value of its relationships. Because the firm finds it easier to credibly promise larger rewards if total future profits are larger, the firm's hiring policy maximizes total future surplus. If the firm can motivate both workers to work hard when demand is robust, then it mimics the formal contract: two workers are hired if demand is good and one is hired if demand is bad. Note that this hiring policy is independent of past play.

This result depends on the assumption that all workers punish a deviation by the firm. This severe punishment requires that they be able to coordinate, but it seems implausible that individuals not currently employed by the firm would be able to coordinate with current workers to punish the firm. Such coordination is not possible if relationships are bilateral.

The purpose of this paper is to show how backward-looking policies arise when relationships are bilateral. In the context of this hiring example, such a policy might lead to lagged hiring following an increase in demand.

To see why this might be the case, suppose first that the firm pursues a policy of hiring two workers as soon as demand increases. In Section 5, we show that each worker can never receive a reward larger than his expected future productivity. If demand is robust, then individual productivity might be so high that both agents can be motivated to work hard (which we show requires $\frac{\delta}{1-\delta}(\alpha R - 1) - \bar{u} \geq 1$). However, the firm might not be able to motivate worker 1 to work hard while demand is weak, since weak demand is persistent and leads to low per-worker productivity.

Suppose that the firm cannot motivate worker 1 if demand is weak. While total continuation surplus is highest if the firm hires two workers after demand increases, the surplus produced by worker 1 is larger if the firm does not hire a second worker. As a result, the firm can credibly offer a larger reward to worker 1 while demand is weak if it promises not to hire a second worker after demand increases. The resulting policy is backward-looking: it sacrifices total continuation surplus to better motivate one of the two workers.

Further, if demand started off in the high state in period 1, then the firm would have no past promises to fulfill. It would immediately hire two workers and motivate them to work hard in every period. A surplus-maximizing relational contract may therefore entail history-dependent inefficiencies. Over time, the principal makes biased decisions in favor of those who performed well in the past, even though such decisions do not maximize total continuation surplus. In the next section, we consider a general model of relational contracts to show how this intuition can lead to backward-looking policies and biased decisions in many different settings.

3 The Model

A single principal (she) and N agents (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 a vector of **efforts** $c = (c_1, \dots, c_N)$, chosen by the agents and parameterized by cost, affects the distribution over a vector of **outcomes** $y = (y_1, \dots, y_N)$. Agent i incurs cost c_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 τ_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 τ. 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** $D \subseteq D$ 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 denotes a distribution over outcomes, $d(y|c, \theta)$, which depends on the effort vector c and the state of the world θ. Agent i chooses effort from a set that depends on the decision and the state of the world, $C_i(d, \theta) \subseteq R_+$. Outcomes are independent across agents conditional on the state of the world: $d(y, \theta) = \prod_{i=1}^{N} d_i(y_i|c_i, \theta)$.

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

Timing The stage game has eight rounds.

1. θ_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 non-negative wages 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 $c_{i,t} \in C_i\left(d_t, \theta_t\right)$.

7. The outcome $y_t = (y_{1,t}, \ldots, y_{N,t})$ is realized, where $y_{i,t} \sim d_{i,t}\left(y| \, c_{i,t}, \theta_t\right)$.

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, outside options are chosen after ex-ante transfer payments w_t have occurred. 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.

Histories and Strategies A history at the beginning of period t is $h_0^t = \{\theta_{t'}, D_{t'}, d_{t'}, w_{t'}, m_{t'}, p_{t'}, c_{t'}, y_{t'}, \tau_{t'}\}_{t'=0}^{t-1}$, 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 $\phi_i(h_x^t)$ 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 θ_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 σ_i maps $\phi_i(\mathcal{H}^t)$ to feasible actions at those private histories. Continuation play at $\phi_i(h^t)$ is denoted $\sigma|\phi_i(h^t)$. We refer to a history-contingent plan of decisions as a **policy**.

Payoffs In period t, agent i's and the principal's payoffs are

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

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

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

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 σ^* 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.

Relative to other typical solution concepts such as Perfect Bayesian Equilibrium, Belief-Free Equilibrium imposes restrictions on how strategies evolve. 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 restrictive.[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.

We focus on surplus-maximizing relational contracts in this paper. A self-enforcing re-

[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 σ^*.

[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).

lational contract 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_x^t, continuation play $\sigma^*|h_x^t$ is surplus-maximizing.[3] If σ^* 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, the players have access to a public randomization device in each round of the stage game. Second, all variables apart from c_t are observed by every player. Efforts c_t 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.

Backward-looking policies reduce the total continuation surplus produced in a relational contract. This reduction in total surplus has a direct impact on *ex ante* total surplus. In this setting, backward-looking policies also potentially decreases the value to the principal of remaining in the relationship. The principal loses her continuation value if she reneges on a promised bonus payment, so she is less willing to follow through on large bonus payments if her continuation value is lower. Agents can be immediately compensated for output using monetary bonuses without any need for biased decisions.

For these reasons, surplus-maximizing relational contracts are always sequentially surplus-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. *Any surplus-maximizing relational contract is sequentially surplus-maximizing. Moreover, there exists a surplus-maximizing relational contract such that on*

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

the equilibrium path, histories that induce identical continuation games induce identical continuation strategies.

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 may motivate agent i to work hard by promising to reward high output with either a large contemporaneous bonus payment $\tau_{i,t}$, substantial continuation rents U_i, or both. For a history h_c^t immediately following effort, define agent i's **reward scheme** $b_i : Y_i \to \mathbb{R}$ as his expected continuation payoff for each possible outcome:

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

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 τ_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, the bonus to agent i must be bounded from above by $(1-\delta)\,\tau_{i,t} \leq \delta\Pi$.

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

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

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 Π 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.

Suppose that all agents are held at their punishment payoffs, so that the principal earns total surplus (net of punishment payoffs) in each period. 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 all agents immediately 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 may 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. *Given a relational contract* σ, history h_x^t, and any agent i, **i-dyad surplus** equals the total surplus produced by agent i :

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

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

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

$$c_t \in \underset{c_i \in C_i(d_t,\theta_t)}{\text{argmax}}\ E_\sigma\left[b_i\left(h_y^t\right)\middle| h_w^t, c_i\right] - (1-\delta)\, c_i \tag{2}$$

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

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

A credible reward scheme satisfies two conditions. First, agent i has to be willing to exert effort $c_{i,t}$ 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 two constraints limit how much b_i can vary with output.

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 σ^* is a self-enforcing relational contract, then for each agent i there exists a reward scheme b_i^* that is credible in σ^*.*

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

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 δU_i. This reward scheme has to satisfy his effort IC constraint (2) in any self-enforcing relational contract. Agent i's continuation surplus is bounded below by his outside option. The principal is willing to pay no more than *his continuation surplus from the relationship with* i. If the principal were asked to pay more than this, he could pay nothing and never again pay agent i, but otherwise act as if no deviation occurred. The other agents would not be able to detect this deviation and so would be willing to continue working for the principal. 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 σ. In each period of this relational contract, the principal chooses the same decision as in σ. 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. 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 decision 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 in 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 and state of the world θ_t are constant and all payoffs are smooth in decisions and efforts.

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

1. $D_t = \left\{ d(\cdot|c,\omega) |\, \omega \in \mathbb{R}_+^N, \sum_{i=1}^N \omega_i \leq 1 \right\}$ *in each period. Outside options are independent of decisions,* $\{\bar{u}_i\}_{i=0}^N$.

2. *Effort is parametrized* $c_{i,t} = \tilde{c}_i(e_{i,t})$ *for* $e_i \in R_+$*, with* $\tilde{c}_i$ *smooth, strictly increasing, and strictly convex.*

3. *The distribution* $d_i(\cdot|\omega)$ *depends only on* ω_i and is smooth with density $\tilde{d}$. d_i *is strictly MLRP-increasing in* e_i *and satisfies CDFC for every* ω. $E[y_i|\omega_i, e_i]$ *is smooth, strictly increasing, and strictly concave in* ω_i*, with* $\lim_{\omega_i \to 0} \frac{\partial}{\partial \omega_i} E[y_i|\omega_i, e_i] = \infty$ *for all* e_i.

4. *For every* $\omega_i > 0$*, there exists a finite* $e_i^{FB}(\omega_i) = \text{argmax}_{e_i} \{E[y_i|\omega_i, e_i] - \tilde{c}_i(e_i)\}$.

In a smooth game, a decision specifies a weight $\omega_{i,t}$ for each agent i in each period t. Agent i's effort together with this weight determine the outcome y_i, where a higher weight $\omega_{i,t}$ 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 ω_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.

A critical feature of these games is that decisions entail trade-offs among agents. Maximizing the surplus produced by agent i requires $\omega_i = 1$, 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 typical in surplus-maximizing relationships.

PROPOSITION 2. *In any surplus-maximizing relational contract* σ^**,*

1. **Money is never burned***:* $\sum_{i=1}^{N} \omega_{i,t} = 1$ *in any on-path* h^t*.*

2. **Sequential surplus maximization entails equal marginal returns***:* σ^* *is sequentially surplus-maximizing only if for any on-path* $h_0^t \in \mathcal{H}_0^t$ *and any agents* i *and* j*,*
$$E_{\sigma^*}\left[\left. \frac{\partial}{\partial \omega_i} E\left[y_{i,t}|\omega_{i,t}, e_{i,t}\right] \right| h_0^t \right] = E_{\sigma^*}\left[\left. \frac{\partial}{\partial \omega_j} E\left[y_{j,t}|\omega_{j,t}, e_{j,t}\right] \right| h_0^t \right]. \tag{4}$$

3. **Biased policies are optimal***: Suppose in every surplus-maximizing relational contract, there exists an agent* i *with* $0 < e_{i,0} < e_i^{FB}(d_{i,0})$*. Then no surplus-maximizing relational contract is sequentially surplus-maximizing. For every* $t \geq 1$*, there exist* i *and* j *such that with positive ex-ante probability,* (4) *fails to hold.*

Proof: See Appendix A.

The first statement of Proposition 2 holds because larger ω_i both increases total surplus and relaxes (3) for agent i. So any surplus-maximizing relational contract will use the full "budget" of ω_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 (4) has to hold in period 0 of any surplus-maximizing relational contract. If σ^* is sequentially surplus-maximizing, then $\sigma^*|\, h_0^t$ is surplus-maximizing and (4) has to hold in any period 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

$$b_i = \begin{cases} \delta \bar{U}_i & y_{i,t} < y_i^* \\ \delta S_i & y_{i,t} \geq y_i^* \end{cases} \tag{5}$$

for some y_j^* that depends on the decision and i's effort. Note that S_i depends on future decisions, which potentially depend on realized outputs y_t. Consider some period $t \geq 1$. Suppose $y_{i,0'} > y_i^*$ for agent i in period 0 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 $\omega_{i,t}$ and decreasing $\omega_{j,t}$ by some small amount. Holding effort fixed, this leads to a second-order decrease in total surplus if (4) 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 0. Since $y_{i,0} > y_i^*$, agent i's dynamic enforcement constraint is strictly weaker in $t = 0$, which has a first-order impact on surplus because i is not exerting first-best effort. Thus, biased decisions increase both effort provision and total surplus.

To explore the types of biases that arise, we consider a special class of smooth games.

COROLLARY 1. *Suppose output is given by $y_i = x_i + g_i(\omega_i)$, where x_i is distributed according to $P_i(x|e) = eP_i^H(x) + (1-e)P_i^L(x)$ with P_i^H strictly MLRP-dominating P_i^L. Define x_i^* as the unique solution to $p_i^H(x_i^*) = p_i^L(x_i^*)$ and ω^{FB} as the unique solution to (4). Suppose no surplus-maximizing relational contract is sequentially surplus-maximizing. If σ^* is surplus-maximizing, then*

1. $\omega_{i,t} > \omega_i^{FB}$ *only if* $x_{i,t'} > x_i^*$ *for some* $t' < t$.
2. *Suppose there exist agents i and j such that $x_{j,t'} < x_j^*$ for all $t' < t$ and $x_{i,t'} > x_i^*$ for some $t' < t$. Then either $e_{i,t'} = e_i^{FB}$ or $\omega_{j,t} < \omega_j^{FB}$.*

Proof: See Appendix A.

Corollary 1 follows almost immediately from the proof of Proposition 2. The first result says that a decision is only biased *towards* an agent if that agent has produced high output in the past. That is, biased policies *favor* agents who have already performed well. The second result shows that policies will typically be biased against those who have performed poorly in the past. Loosely, agents "compete" for policies that favor them, and an agent who performs well "wins" future decisions that favor him.

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 occur in a relationship. First, we revisit the hiring setting from Section 2 and prove that employment can optimally lag demand. Second, we show how a firm might both delay and distort irreversible investments to better motivate its divisions or employees. Finally, we argue that a manager might continue to favor an employee even after learning that he is worse than the alternantive.

We make several assumptions that are common to all three examples. There are two agents, each of whom chooses a binary effort $c_{i,t} \in \{0, c\}$ in each period. If agent i's effort is low ($c_{i,t} = 0$), then his output is $y_{i,t} = 0$ regardless of the policy chosen. If his effort is $c_{i,t} = c$, then output is $y_{i,t} = 0$ with probability $1 - p$ and is otherwise $y_{i,t} = y_i(d_t, \theta_t) > 0$. Note that high output potentially depends on both the policy d_t and the state of the world θ_t.[4]

6.1 Hiring and Firing

Consider the following formalization of the hiring model discussed in Section 2. This example illustrates how persistent shocks to the production function can lead to history-dependent biases.

DEFINITION 4. *The* ***hiring game with demand shocks*** *has 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. We assume that if* $d_t = 1$, *then agent* 1 *is hired.*[5] *Agent* i*'s outside option is* $\bar{u}_i > 0$.
- *The probability of high output following high effort is* $p = 1$. $y_1(\theta, 1) = \theta c$ and $y_2(\theta, 1) = 0$ if only worker 1 is hired. For each agent i, $y_i(\theta, 2) = \alpha\theta c$ with $\alpha < 1$ if both workers are hired.

The principal is a firm that faces demand θ_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

[4] This description entails a slight abuse of notation, since d is no longer a distribution per se. Instead, d can be mapped 1-1 to a distribution over outputs.

[5] This assumption simplifies the analysis but is irrelevant for the result.

increases, it remains robust thereafter. The marginal 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. The optimal number of employees depends on demand and the effort chosen by each worker. We assume that parameters are such that 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 high, surplus is maximized if the firm hires two agents when demand is high 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 another 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* $R > \frac{\bar{u}+1}{2\alpha-1} > W > \bar{u}+1$ *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* σ^* *satisfies:*

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

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

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 if $d_t = 2$. Since low demand is persistent, however, it might be impossible to satisfy (3) in the low-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 low-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 low-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 example is consistent with the stylized macroeconomic fact that hiring tends to lag productivity during a recovery. In Proposition 3, dyad-specific surplus–and hence the amount by which each agent can be motivated–increases with demand. Hiring, however, 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, the model suggests that new entry may drive increased employment immediately following a recession.

More broadly, our model suggests that relational considerations can lead to *history-dependent*, *endogenous* "adjustment costs" that prevent (some) firms from adapting quickly to changing demand. By a similar intuition, firms might refrain from firing workers at the beginning of a recession to maintain long-term relationships with them. In an extreme form, this policy would resemble Lincoln Electric's "continuous employment guarantee," which forbids laying off workers. As in this example, Lincoln Electric is also slow to hire new employees during a boom (see Dawson (1999)).

6.2 Irreversible, Agent-Specific Investments

Suppose a principal can permanently invest in one of her agents to increase that agent's productivity. The agent could be an employee or a division within the firm, and the investment is agent-specific but might represent human capital (e.g., additional training), physical plant (e.g., division-specific equipment), or job design (e.g., a permanent promotion). How should the principal invest to strengthen its relationships?

Int his example, we show that the principal might refrain from implementing a permanent decision with a positive return in order to use it as a motivational tool.

DEFINITION 5. *The* ***irreversible investments game*** *has the following features:*

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

- *Investments are permanent, so* $D_{t+1} = \{d_t\}$ *if* $d_t \in \{1, 2\}$*. So long as* $d_t = 0$*, then* $D_{t=1} = \mathcal{D}$*. Outside options are* $\bar{u}_i = 0$*.*

- *If* $d_t \neq i$, *then* $y_i(d_t, \theta_t) = \alpha_0 c$, *with* $\alpha_0 > \frac{1}{p}$. *If* $d_t = i$, *then* $y_i(d_t, \theta_t) = \alpha_i c$, *with* $\alpha_1 \geq \alpha_2 > \alpha_0$.*It is better for both agents to work hard without investment than for one to work hard with investment:* $(2\alpha_0 - \alpha_1)p - 1 > 0$.

In the irreversible investments game, the principal chooses when to make a productivity-enhancing investment. Proposition 1 implies that the principal should maximize total continuation surplus if relationships are public. Hence, the principal would immediately choose to invest in agent 1, who has the higher productivity gain from investment.

The surplus-maximizing policy is very different in the game with bilateral relationships. The principal can use the possibility of future investments as a motivational tool. By offering to invest in whichever agent produces high output, the principal can potentially reward both agents for hard work. Note that this promise entails two inefficiencies relative to the policy in the game with public relationships. First, the principal delays investment to retain the option value rewarding either worker. Second, the principal might choose the less-efficient agent if that agent happens to perform well in the early periods of the game.

PROPOSITION 4. *Consider the irreversible investments game with bilateral relationships. Assume* $(1-p)(p\alpha_2 - 1) > p\alpha_0 - 1$. *For fixed* α_2, *there exists a range* $0 \leq \underline{\delta} < \bar{\delta} < 1$ *such that if* $\underline{\delta} < \delta < \bar{\delta}$ *and* $\alpha_1 - \alpha_2 < \frac{1-\delta}{\delta} \frac{(2\alpha_0 - \alpha_1)p - 1}{p(1-p)}$, *any surplus-maximizing relational contract* σ^* *satisfies:*

1. *In the first period,* $d_0 = 0$ *and* $c_{1,0} = c_{2,0} = c$.

2. *As* $t \to \infty$, $d_t \in \{1, 2\}$ *with probability* 1. *Both* $d_t = 1$ *and* $d_t = 2$ *occur with strictly positive probability. If* $d_t \in \{1, 2\}$, *then* $c_{d_t,t} = c$ *but* $c_{i,t} = 0$ *for* $i \neq d_t$.

Proof: See Appendix A.

If the two agents' productivities are not too different and the benefits of investment not too large, then the principal finds it optimal to award investment to an agent who produces high output. This policy resembles a tournament in which the winner receives an investment, which in turn makes a substantial reward for that agent credible. Because both agents may "win" the investment, both are willing to work hard in the early periods. After the principal invests in 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. Investment is delayed so that it can be used as the "prize," and the policy is biased, because the less productive agent might win.

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?

We present an example that demonstrates the relational costs of experimenting to learn about an unknown state of the world.[6]

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

- *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* α *and otherwise* $\theta_{t+1} = H$*.*[7]

- *If* $d_t \neq i$*, then* $y_i(-i, \theta_t) = 0$*. Otherwise,* $y_1(1, \theta_t) = \theta_t c$ *and* $y_2(1, \theta_t) = Rc$ *with* $\frac{1}{p} < L < R < qL + (1-q)H$*. Outside options are* $\bar{u}_i = 0$*.*

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 strictly higher than agent 2's, but otherwise it is strictly 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 allocated 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.

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

[7] 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.

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.

7 Conclusion

Biased policies are a prominent feature of long-term relationships in many settings. Managers favor high-performing workers, divisions, and suppliers by biasing policies to 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 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 σ^* 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,

$$\delta \sum_{i \in I} E_{\sigma^*}\left[\bar{U}_i\left(h_0^{t+1}\right)\middle| h_y^t\right] \leq \sum_{i \in I} b_i\left(\phi_0\left(h_y^t\right)\right) \leq \delta \sum_{i \in I} E_{\sigma^*}\left[\sum_{i \in I} U_i\left(\sigma^*, h_0^{t+1}\right) + \Pi\left(\sigma^*, h_0^{t+1}\right)\middle| h_y^t\right]. \tag{6}$$

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

$$\begin{aligned} U_i(\sigma^*, h_d^t) &\geq \bar{U}_i(h_d^t) \\ \Pi\left(\sigma^*, h_d^t\right) &\geq \sum_{i \in I}\left(E_{\sigma^*}\left[b_i\left(\phi_0\left(h_y^t\right)\right) - (1-\delta)\, c_{i,t}\middle| h_d^t\right] - U_i\left(\sigma^*, h_d^t\right)\right) \end{aligned} \tag{7}$$

2. For strategy σ, suppose there exists $b_i : \phi_0\left(\mathcal{H}_y^t\right) \to \mathbb{R}$ satisfying (2), (6), and (7). Then there exists a BFE σ^* that induces the same distribution over $\left\{D_t, d_t, c_t, y_t, \{u_i\}_{i=0}^N\right\}_{t=0}^\infty$ as σ.

Proof of Lemma A.1

1: Suppose σ^* 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

$$b_i\left(\phi_0\left(h_y^t\right)\right) = E_{\sigma^*}\left[(1-\delta)\,\tau_{i,t} + \delta U_i\middle| \phi_0\left(h_y^t\right)\right]. \tag{8}$$

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

$$c_{i,t} \in \underset{c_i \in C_i(d_t, \theta_t)}{\operatorname{argmax}}\ E_{\sigma^*}\left[(1-\delta)\,\tau_{i,t} + \delta U_i\middle| h_w^t, c_{i,t} = c_i\right] - (1-\delta)\, c_i, \tag{9}$$

which implies (2). Suppose $b_i\left(\phi_0\left(h_y^t\right)\right) < \delta E_{\sigma^*}\left[\left.\bar{U}_i\left(h_0^{t+1}\right)\right|\phi_i\left(h_y^t\right)\right]$. 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

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

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 (6).

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

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

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

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

implying the first line of (7).

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

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

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

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

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 (6).

2 : Define $\zeta\left(h^t\right)=\{D_{t'},d_{t'},c_{t'},y_{t'}\}_{t'=0}^t$. 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 σ^* 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 σ^*, then σ^* 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 $w_{i,t} = E_\sigma \left[u_{i,t} + c_{i,t} - \frac{1}{1-\delta} \left(b_i \left(\phi_0 \left(h_y^t \right) \right) - \delta U_i \left(\sigma, h_0^{t+1} \right) \right) \middle| h_d^t \right]$.

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

(e) Agent i chooses $c_{i,t}$ as in h_c^t.

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

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

(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 $\sigma| \{h_c^t, y_t\}$.

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 σ^* 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 σ^*, and (ii) players have no profitable one-shot deviation. For (i), agent i's payoff is

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

Recall h_d^t and h_c^t have distributions $\sigma| h_0^t$ and $\sigma| h_d^t$, respectively. Moreover, $\sigma^*|\tilde{h}_c^t$ and $\sigma|h_c^t$ 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 (7). 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

$$c_{i,t} \in \underset{c_i \in C_i(d_t, \theta_t)}{\operatorname{argmax}}\ E_{\sigma^*}\left[(1-\delta)\left(\tau_{i,t} - c_i\right) + \delta U_i\left(\sigma^*, h_0^{t+1}\right) \middle|\, \tilde{h}_w^t, c_{i,t} = c_i\right].$$

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 (6). Let $J = \{i \in \{1, \ldots, N\}|\, \tau_{i,t} \leq 0\}$. The principal has no profitable deviations as long as

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

By construction, $E_{\sigma^*}\left[\Pi\left(\sigma^*, \tilde{h}_0^{t+1}\right) \middle|\, \tilde{h}_y^t\right] = E_\sigma\left[\Pi\left(\sigma, h_0^{t+1}\right) \middle|\, h_y^t\right]$. So the second inequality in (6) implies that the principal has no profitable deviation.

Completing Proof of Proposition 1

Towards contradiction: suppose a surplus-maximizing BFE σ^* is not sequentially surplus-maximizing. We first define a strategy $\tilde{\sigma}$ that induces the same distribution over $\{D_t, d_t, c_t, y_t\}_{t=0}^\infty$ as σ^*, 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 σ^* as in the recursive construction from Lemma A.1, with the sole exception that $\tau_{i,t} = 0$ in each period, anad

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

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), (6), and (7) for σ^*. 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 $\left\{D_t,d_t,c_t,y_t,\pi\left\{u_i\right\}_{i=1}^N\right\}_{t=0}^{\infty}$ as $\tilde{\sigma}$. Recall that $\tilde{\sigma}$ and σ^* generate the same total surplus, so $\tilde{\sigma}^*$ is a surplus-maximizing BFE. Because σ^* 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 (6) and (7). Since the principal's continuation payoff is higher under $\bar{\sigma}$ than under $\tilde{\sigma}^*$, we conclude that $\bar{\sigma}$ satisfies Lemma 1. So σ^* cannot be surplus-maximizing, which is a contradiction.

Proof of Lemma 1

1 : Suppose σ^* is a BFE and define $b_i:\phi_0\left(\mathcal{H}_y^t\right)\to\mathbb{R}$ as in (8). Fix a history $h_w^t\in\mathcal{H}_w^t$. Then (9) 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

$$E_{\sigma^*}\left[\Pi\left(\sigma^*,h_0^{t+1}\right)-\sum_{t'=0}^{\infty}\sum_{i\in I}\left(1-\delta\right)\delta^{t'}\left(y_{i,t+t'}-w_{i,t+t'}-\tau_{i,t+t'}\right)\middle|\,h_y^t\right].\qquad(10)$$

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 σ^* except for transfer payments. Transfer payments do not affect the continuation game, so this strategy is feasible. Moreover, this strategy and σ^* are indistinguishable for every agent $i \notin I$. Therefore, the principal's payoff from this strategy is no less than (10).

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

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

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

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

   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_c^t, and (ii) the reward scheme minus dyad-specific surplus for each h_y^t that might follow h_c^t:

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

   (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 $i \in \{1, \ldots, N\}$,

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

   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 $\sigma | \{h_c^t, y_t\}$.

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 σ^*. 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 $\tilde{h}_0^t$, $E_{\sigma^*}\left[u_{i,t} | \tilde{h}_0^t \right] = E_{\sigma^*}\left[y_{i,t} - c_{i,t} | \tilde{h}_0^t \right]$. 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 σ^*, $S_i\left(\sigma^*, \tilde{h}_0^t\right) = S_i\left(\sigma, h_0^t\right)$.

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 σ^*: at any on-path $\tilde{h}_c^t$, $U_i\left(\sigma^*, \tilde{h}_0^{t+1}\right) = S_i\left(\sigma^*, h_0^{t+1}\right)$ and $E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right) \middle| h_c^t, y_t\right] = E_{\sigma^*}\left[S_i\left(\sigma^*, \tilde{h}_0^{t+1}\right) \middle| \tilde{h}_c^t, y_t\right]$. But b_i satisfies (2), so agent i has no profitable deviation from effort regardless of his beliefs about the true (on-path) history $\tilde{h}_c^t$. 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 σ^* is a BFE with the desired properties.

Proof of Proposition 2

(1) : If $\sum_{i=1}^N \omega_{i,t} < 1$ in some period t, consider the alternative decision $\tilde{\omega}_t$ with $\sum_{i=1}^N \tilde{\omega}_{i,t} = 1$ and $\tilde{\omega}_{i,t} \geq \omega_{i,t}$ 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 (4) fails. 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 (4). Such an alternative exists by the Intermediate Value Theorem. So (4) holds for $t = 0$ in any surplus-maximizing relational contract. But $\sigma^*|\,h^t$ is surplus-maximizing for any on-path $h_0^t \in \mathcal{H}_0^t$ if σ^* is sequentially surplus-maximizing. Therefore, (4) holds at every on-path h_0^t.

(3) : Let σ^* be a surplus-maximizing relational contract. Suppose towards contradiction that σ^* is sequentially surplus-maximizing. By assumption, there exists some agent i for whom $0 < e_{i,0} < e_i^{FB}(\omega_{i,0})$.

Because F_i satisfies MLRP and CDFC, we can replace the IC constraint (2) with its first-order condition:

$$(1-\delta) c'(e_{i,t}) \geq \int_0^\infty E_{\sigma^*}\left[b_i | h_c^t, y_{i,t}\right] \frac{\partial \tilde{d}_i}{\partial e_{i,t}} (y_{i,t} | e_{i,t}) \, dy_{i,t} \tag{11}$$

with equality if $e_{i,t} > 0$. There exists a unique $y_i^*(\omega_i, e_i)$ for which $\frac{\partial \tilde{d}_i}{\partial e_{i,t}}(y_i^*(\omega_i, e_i) | \omega_i, e_i) = 0$ because d_i is strictly MLRP-increasing in $e_{i,t}$. Together with the dynamic-enforcement constraint (3), (11) implies that the optimal reward scheme $b_i(y_t)$ is the "step function,"

$$b_i = \begin{cases} \delta \bar{U}_i(h_0^{t+1}) & y_{i,t} < y_i^*(\omega_{i,t}, e_{i,t}) \\ \delta S_i(\sigma^*, h_0^{t+1}) & y_{i,t} \geq y_i^*(\omega_{i,t}, e_{i,t}) \end{cases} \tag{12}$$

Suppose $e_{i,t} < e_i^{FB}(\omega_{i,t})$. Then we claim that increasing the upper bound of the dynamic-enforcement constraint (3) by $\varepsilon > 0$ for all $y_{i,t} > y_i^*(\omega_{i,t}, e_{i,t})$ leads to strictly higher effort. Because $e_{i,t} > 0$, (11) may be written

$$\frac{1-\delta}{\delta} c'(e_{i,t}) = \int_0^{y_i^*(\omega_{i,t}, e_{i,t})} \bar{u}_i \frac{\partial \tilde{d}_i}{\partial e_{i,t}} (y_i | \omega_{i,t}, e_{i,t}) \, dy_i + \int_{y_i^*(\omega_{i,t}, e_{i,t})}^\infty S_i \frac{\partial \tilde{d}_i}{\partial e_{i,t}} (y_i | \omega_{i,t}, e_{i,t}) \, dy_i.$$

Increasing S_i implies that the left-hand side is strictly smaller than the right-hand side in this expression. By assumption, $c'(e)$ and $\frac{\partial \tilde{d}_i}{\partial e}$ are differentiable in e. Similarly, $y_i^*(\omega_{i,t}, e_{i,t})$ is differentiable in e by the Implicit Function Theorem. Therefore, either agent i is willing to choose e_i^{FB} under the new contract, or there exists some $e' > e_{i,t}$ which satisfies the first-order IC constraint. Define $e_i(S)$ 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 $e_i(S)$ is differentiable for all $e_i(S) < e_i^{FB}$.

Consider the set of histories for which agent j has never produced output $y_{j,t}$ larger than $y_j^*(\omega_{j,t}, e_{j,t})$: $L_{j,t} = \left\{h^t \in \mathcal{H}_0^t | \, y_{j,t'} < y_j^*(\omega_{j,t'}, e_{j,t'}) \text{ for all } t' < t\right\}$. Also define the set of histories for which agent i produces $y_{i,0} > y_i^*(\omega_{i,0}, e_{i,0})$: $H = \{h^t \in \mathcal{H}_0^t | \, y_{i,0} > y_i^*(\omega_{i,0}, e_{i,0})\}$. Because f is smooth, for any $h_0^t \in L_{j,t}$, $\Pr_{\sigma^*}\{L_{j,t+1} | h_0^t\} > 0$. Define $E_t = \{h^t \in \mathcal{H}_0^t | \, h^t \in H \bigcap \{\cup_{j \neq i} L_{j,t}\}\}$. Then it follows that $\Pr_{\sigma^*}\{E_t\} > 0$ for each $t > 0$.

At all histories $h^t \in E_t$, consider the alternative decision $\tilde{\omega}_{i,t} = \omega_{i,t} + \varepsilon$, $\tilde{\omega}_{j,t} = \omega_{j,t} - \varepsilon$, and $\tilde{\omega}_{k,t} = \omega_{k,t}$ for all $k \neq i, j$. This alternative is feasible for small $\varepsilon > 0$, because

$\lim_{\omega_i \to 0} \frac{\partial}{\partial \omega_i} E[y_i|\omega_i, e_i] = \infty$ and hence ω_i^{FB} and ω_j^{FB} are interior. 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 (12) and the fact that $y_{j,t} < y_j^* (\omega_{j,t'}, e_{j,t'})$ in all $t' < t$. This alternative strictly relaxes agent i's dynamic-enforcement constraint (3) in period 0 because $\Pr_{\sigma^*} \{E_t\} > 0$.

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 :

$$K(\varepsilon) = \delta^t (1-\delta) E_{\sigma^*} \left[\left(g_i(e_{i,t}, \omega_{i,t}) + g_j(e_{j,t}, \omega_{j,t}) \right) - \left(g_i(e_{i,t}, \omega_{i,t} + \varepsilon) + g_j(e_{j,t}, \omega_{j,t} - \varepsilon) \right) \middle| E_{i,t} \right] \Pr_{\sigma^*} \{E_t\} > 0.$$

Agents i and j have independent output, so this perturbation changes agent i's optimal reward scheme (12) in period 0 to

$$\tilde{b}_i(x_{i,0}) = \begin{cases} \delta \bar{u}_i & y_{i,0} \leq y_i^* (\omega_{i,0}, e_{i,0}) \\ S_i + \xi(\epsilon) & \text{otherwise.} \end{cases}$$

where $\xi(\epsilon) = (1-\delta)\, \delta^t E_{\sigma^*} \left[g_i(\omega_{i,t} + \varepsilon) - g_i(\omega_{i,t}) \,|E_t \right] \Pr_{\sigma^*} \left[\bigcup_{j \neq i} L_{j,t} | H \right]$. Note that $\xi(\varepsilon) > 0$ and is a smooth function of ϵ.

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

$$B(\varepsilon) = \left[E\left[y_i \middle| \omega_{i,0}, e_i(S+\xi) \right] - c\left(e_i(S+\xi) \right) \right] - \left[E\left[y_i \middle| \omega_{i,0}, e_i(S) \right] - c\left(e_i(S) \right) \right].$$

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

$$\begin{aligned} & \frac{1}{\varepsilon} \left\{ \left[E\left[y_i \middle| \omega_{i,0}, e_i(S+\xi(\varepsilon)) \right] - c\left(e_i(S+\xi(\varepsilon)) \right) \right] - \left[E\left[y_i \middle| \omega_{i,0}, e_i(S) \right] - c\left(e_i(S) \right) \right] \right\} \\ > \;& \frac{1}{\varepsilon} \delta^t (1-\delta) \left[\left(g_i(\omega_{i,t}) + g_j(\omega_{j,t}) \right) - \left(g_i(\omega_{i,t} + \varepsilon) + g_j(\omega_{j,t} - \varepsilon) \right) \right]. \end{aligned}$$

As $\varepsilon \to 0$, the costs converge to 0 by (4). The benefits converge to

$$\frac{\partial \left(E\left[y_i \middle| \omega_{i,0}, e_i(S) \right] - c\left(e_i(S) \right) \right)}{\partial e} \frac{\partial e_i}{\partial S_i} \frac{\partial \xi}{\partial \varepsilon} > 0,$$

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

We conclude that surplus-maximizing relational contracts are not sequentially surplus-maximizing. Because $e_i^{FB}(\omega_{i,0}) > e_{i,0} > 0$ for some agent i in every surplus-maximizing relational contract, this same argument proves that (4) fails with positive ex ante probability in every period $t \geq 1$.

Proof of Corollary 1

As before, we can replace the incentive constraint (2) with its first-order condition:

$$c'\left(e_{i,t}\right) \geq \int_0^\infty E_{\sigma^*}\left[b_i | h_c^t, x_{i,t}\right]\left(f_i^H\left(x_{i,t}\right) - f_i^L\left(x_{i,t}\right)\right) dx_{i,t}$$

with equality if $e_{i,t} > 0$. The optimal reward scheme is then

$$b_i = \begin{cases} \delta \bar{u}_i & x_i < x_i^* \\ \delta S_i & x_i > x_i^*. \end{cases}$$

If $\omega_{i,t} > \omega_i^{FB}$, then there exists some agent j for whom $\omega_{j,t} < \omega_j^{FB}$. If $x_{i,t'} < x^*$ for all $t' < t$, then decreasing $\omega_{i,t}$ and increasing $\omega_{j,t}$ by the same amount increases total surplus holding efforts fixed. It also leads to weakly stronger incentives for agent j without weakning incentives for any other agent. So it must be that $x_{i,t'} > x^*$ for some $t' > t$.

If $x_{j,t'} < x^*$ for all $t' < t$ and there exists some agent i and some $t' < t$ such that $x_{i,t'} > x^*$, then the argument from Proposition 2 applies. Hence, if $\omega_{j,t} \geq \omega_j^{FB}$, then there exists a strictly better perturbed policy.

Proof of Proposition 3

1 : Suppose $\theta_0 = R$. Then $\theta_t = R$ in all periods t. Fix δ_1 such that

$$c = \frac{\delta_1}{1-\delta_1}\left(\left(\alpha R - 1\right) c - \bar{u}\right)$$

Then for $\delta \geq \delta_1$, both agents can be motivated to choose $c_{i,t} = c$ if $d_t = 2$. This relational contract attains the maximum feasible surplus and so is optimal.

2 : Suppose $\theta_0 = W$. If $\theta_t = W$ and $c_{1,t} = 0$, then d_t is irrelevant. If $\theta_t = W$ and $c_{1,t} = c$, then $d_t = 1$ maximizes both total and dyad-specific surplus. So without loss of generality we can assume that $d_t = 1$ whenever $\theta_t = W$. Because $\delta \geq \delta_1$, if $\theta_t = R$, then all working agents choose $c_{i,t} = c$.

Define

$$S_{R,1} = (R-1)\,c$$
$$S_{R,2} = (\alpha R - 1)c$$

$S_{R,d}$ is the dyad-surplus for agent 1 in period t if (i) $\theta_t = R$, and (ii) $d_t = d$ in all future periods. Note that $S_{R,1} > S_{R,2}$ because $\alpha < 1$.

Suppose that $d_t = d$ whenever $\theta_t = R$. Suppose that agent 1 chooses $c_{i,t} = c$ whenever $\theta_t = W$. Then agent 1's net dyad-surplus if $\theta_t = W$ equals

$$S_{W,d} = qS_{R,d} + (1-q)\left((1-\delta)(W-1)\,c + \delta S_{W,d}\right).$$

Hence,

$$S_{W,d} = \frac{qS_{R,d} + (1-q)(1-\delta)(W-1)\,c}{1-\delta(1-q)}.$$

Note that $S_{W,1} > S_{W,2}$, because $S_{R,1} > S_{R,2}$. Moreover, $S_{W,2} < S_{R,2}$ because $\alpha R > W$. Defining δ_2 by

$$c = \frac{\delta_2}{1-\delta_2}\left(S_{W,2} - \bar{u}\right),$$

we immediately conclude $\delta_2 > \delta_1$. If $\delta \in [\delta_1, \delta_2)$, then in any sequentially surplus-maximizing equilibrium $c_{1,t} = 0$ whenever $\theta_t = W$.

Consider the following alternative policy: in the first period t such that $\theta_t = R$, the principal chooses $d_t = 2$ with probability ξ and otherwise chooses $d_t = 1$, where ξ is chosen so that

$$c = \frac{\delta}{1-\delta}\left(\xi S_{W,2} + (1-\xi)\,S_{W,1} - \bar{u}\right).$$

Because $S_{W,1} > S_{W,2}$, there exists some δ_3 such that if $\delta \geq \delta_3$, $\xi \in [0,1]$. As $\delta \to \delta_2$, $\xi \to 0$. The gains from this alternative policy are at least $(1-\delta)(W-1)\,c > 0$. Hence, there exists some $\delta_2 > \delta_4 \geq \delta_3$ such that if $\delta \geq \max\{\delta_1, \delta_4\}$, then this alternative policy generates strictly higher surplus than any relational contract in which $d_t = 2$ whenever $\theta_t = R$. This proves the claim.

Proof of Proposition 4

Define $\bar{\delta} < 1$ by

$$\frac{c}{p} = \frac{\bar{\delta}}{1-\bar{\delta}}\left(p\alpha_0 - 1\right)c.$$

Then for any $\delta < \bar{\delta}$, agent i is unwilling to work hard if $d_t \neq i$ in every period with probability 1. So if $d_t = -i$, then $c_{i,t'} = 0$ for all future periods $t' \geq t$. Define $\underline{\delta} > 0$ by

$$\frac{c}{p} = \frac{\underline{\delta}}{1-\underline{\delta}}\left(1-p\right)\left(p\alpha_2 - 1\right)c.$$

By assumption, $\bar{\delta} > \underline{\delta}$.

1 : Suppose $d_0 \neq 0$. Then $c_{d_t,t} = c$ in every period t, while $c_{i,t} = 0$ for $i \neq d_t$. Consider the following alternative policy. In period $t = 0$, $d_0 = 0$. If $y_{2,0} > 0$ but $y_{1,0} = 0$, then $d_1 = 2$. Otherwise, $d_1 = 1$. Because $\delta > \underline{\delta}$, $c_{i,t} = c$ in any period with $d_t = i$. Therefore, both agents are willing to choose $c_{i,t} = c$ in period 0. This alternative dominates the policy with $d_0 \neq 0$ so long as

$$(\alpha_1 - 1) c < (1-\delta) 2 (p\alpha_0 - 1) c + \delta q (1-q) (p\alpha_2 - 1) c + \delta (1 - q(1-q)) (p\alpha_1 - 1) c$$

or

$$(1-\delta)((2\alpha_0 - \alpha_1) p - 1) > \delta q (1-q) p (\alpha_1 - \alpha_2),$$

which holds by assumption. Therefore, any equilibrium in which $d_0 \neq 0$ is strictly worse than an equilibrium in which $d_0 = 0$.

2 : By the previous argument, the optimal equilibrium has $d_0 = 0$. Fix any on-path history h^t such that $d_{t-1} = 0$ on the equilibrium path. If $d_t = 0$ in all future periods, then $c_{i,t} = 0$ for both agents i, because $\delta < \bar{\delta}$. But then the alternative equilibrium outlined above generates strictly higher total and dyad-specific surplus. So there exists some $t' > t$ for which $\Pr[d_{t'} \neq 0 | h^t] > 0$. Moreover, for any $\delta < \bar{\delta}$, there exists some period t' such that this probability is bounded below by some strictly positive number. If $d_{t'} = 0$, then this argument can be repeated with the same strictly positive lower bound. So $\lim_{t\to\infty} \Pr[d_t = 0 | h^t] = 0$, as desired.

Now, suppose that $\lim_{t\to\infty} \Pr[d_t = i | h^t] = 1$ for some $i \in \{1,2\}$. Then agent $-i$ can never be motivated to exert effort. Then this equilibrium is dominated by an equilibrium with $d_0 = 1$. If $d_t \in \{1,2\}$, then $c_{i,t} = 0$ for $i \neq d_t$ follows immediately from $\delta < \bar{\delta}$, while $c_{d_t,t} = c$ follows from $\delta \geq \underline{\delta}$.

Proof of Proposition 5

Suppose $d_0 = 2$ in σ^*. Let $t > 0$ be the first period in which $d_t = 1$ with positive probability. Consider replacing the continuation equilibrium with σ^*. This generates weakly larger total surplus, because σ^* 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 $c_{i,t} = c$, regardless of θ_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 ξ, $d_t = 1$ 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. ξ is chosen to solve

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

Since $\delta \geq \delta_1$, $\xi \in (0,1)$. Under this alternative strategy, agent 1 is willing to work hard in period $t = 0$. Therefore, total surplus under this strategy equals

$$\begin{aligned}&(1-\delta)\,(p\,(qL + (1-q)\,H - 1)\,c)\\&+\delta\,[p\,((1-q)\,(pH-1) + q\xi\,(pL-1) + q\,(1-\xi)\,(pR-1)) + (1-p)\,((1-q)\,pH - \dots)]\end{aligned}$$

Since $qL + (1-q)\,H \geq R$, this equilibrium dominates an equilibrium with $d_0 = 2$ if

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

$\xi \rightarrow 0$ as $\delta \rightarrow \delta_2$ 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 $\Pr_{\sigma^*}[\theta_{t'} = B, d_{t'} = 1] > 0$ 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).

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).

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).