Policies in Relational Contracts

Daniel Barron
Northwestern University

Michael Powell
Northwestern University

May 2015

Abstract

How should a firm set policies—public decision plans that determine the role of its employees, divisions, and suppliers—to strengthen its relationships? We explore whether and how a principal might bias the decisions she makes to foster relational contracts with her agents. To this end, we examine a flexible dynamic game between a principal and several agents with unrestricted vertical transfers and symmetric information. We show that if relationships are bilateral—each agent observes only his own output and pay—then the principal may optimally make decisions in a systematically backward-looking, history-dependent way in order to credibly reward agents who performed well in the past. We first show that these backward-looking policies are prevalent in a broad class of settings. Then we show by example how such policies might affect firm performance: for example, hiring might lag increases in demand or investment might be awarded in a biased tournament. In contrast to the game with bilateral relationships, we show that if monitoring is public, optimal policies never involve biased decisions.

1 Introduction

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

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

In this paper, we argue that backward-looking policies can arise in optimally managed relationships among a principal and her agents. Biased decisions lead to lower continuation surplus. However, a principal who promises to bias future decisions towards an agent is better able to credible promise monetary rewards that motivate that agent today. To make this point, we develop a general framework that builds upon Levin's (2003) 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' efforts 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 show that backward-looking policies arise naturally if relationships are bilateral—that is, if each agent cannot observe the principal's interactions with other agents. In this setting, players cannot coordinate punishments or rewards. A decision that makes an agent more integral to the principal ensures that the principal and that agent have more to lose if they do not uphold their promises to one another. In particular, the principal can promise larger rewards to an agent who is expected to produce more future surplus. Decisions biased toward an individual therefore complement more generous reward schemes for that individual but also negatively affect the firm's overall 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 to reward those employees either immediately or in the future. But these promises are only credible if early employees know that they will remain an important part of the firm in the future. One way to ensure that early employees remain valued would be 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 costless, 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 Section 5.

A game with bilateral relationships has imperfect private monitoring—agents do not observe one another's output nor pay. In general, such games are difficult to analyze, because standard equilibrium concepts are not recursive. For most of the paper, we consider an equilibrium refinement to ensure that our relational contracts are recursive. This solution concept provides a tractable way to emphasize the forces that lead to biased policies in optimal relational contracts. In Section 6, we demonstrate in the context of a simple class of games that biased policies arise even if we consider the full (non-recursive) set of Perfect Bayesian Equilibria.

The first step of our analysis is to develop a set of intuitive necessary and sufficient conditions for a policy to be part of a self-enforcing relational contract. Using these conditions, we consider a broad class of environments and show that backward-looking policies are typically part of surplus-maximizing relational contracts. Indeed, decisions are biased with positive probability in nearly every period unless agents either already exert first-best effort or exert no effort at all. 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.

Next, we apply our framework to two examples to illustrate how backward-looking policies manifest in stylized 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 and link this distortion to several recent empirical observations about job growth. We also argue that a firm might both delay and distort investments in projects (or promotions) to better motivate their managers (or employees).

Finally, we explore the assumption of bilateral monitoring with three extensions. First, we show that if the game has public monitoring, then biased decisions are never surplus-maximizing. Unlike the setting with bilateral monitoring, agents can coordinate to jointly punish the principal if she does not uphold her promises to one of them if monitoring is public. Biased decisions decrease total continuation surplus and weaken the principal's incentives to uphold her promises, so they have no place in a surplus-maximizing relationship. Second, we explore the role of biased policies if agents can coordinate to punish the principal, but only imperfectly. Restricting attention to a simple example and a stylized model of imperfect coordination, we show that biased policies may play a role in an optimal relational contract so long as the principal's deviation in one relationship does not become publicly observed with probability 1. Finally, we show that our central intuition is not driven by our restriction to recursive equilibria. For a simple class of games, we prove that similarly backward-looking policies arise if we consider the full set of Perfect Bayesian Equilibria.

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) identifies conditions under which a long-term formal contract is sequentially efficient. We focus on two of these conditions that have been extensively explored in the literature. First, the principal must be able to punish the agent without simultaneously harming herself. Second, players must have symmetric information about future payoffs. If either of these conditions fail, then the optimal formal contract might entail dynamic inefficiencies.

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 respond to past outcomes. For instance, Fong and Li (2012) show that the principal might inefficiently suspend production to punish poor performance if the agent has limited liability. Li, Matouschek, and Powell (2015) show that if transfers are limited but the principal can reward and punish the agent with future control rights, she may permanently alter the firm's organization away from what maximizes continuation surplus. Board (2011) considers a setting in which a principal chooses to trade with a single agent who is liquidity constrained in each period. Because the principal optimally backloads incentive payments, she distorts this allocation decision to favor agents with whom she has traded in the past. Halac (2012), Malcomson (2014), and others study how relational contracts evolve if the players have asymmetric information about the future. Relational concerns influence dynamics in these papers. However, FHM's discussion suggests that the optimal formal contract in these settings might also entail history-dependent inefficiencies.

This paper takes a different approach. We focus on an environment that satisfies the conditions of FHM, so that optimal formal contracts would not exhibit any history-dependent inefficiencies. Even though history-dependent inefficiencies would not arise in an optimal formal contract, we show that they 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 related to Andrews and Barron (2014), who analyze optimal allocation dynamics in a supply chain, and Calzolari and Spagnolo (2011), who consider procurement auctions. The goal of our analysis is to extend the basic intuition of these papers and provide a general framework for analyzing backward-looking policies in relational contracts.

Our dynamic game has imperfect private monitoring: we assume that one agent cannot observe the principal's interactions with 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 a set of recursive equilibria (related to but slightly weaker than the belief free equilibria from Ely, Horner, and Olszewski (2005)), which allow us to highlight the intuition behind biases in surplus-maximizing relationships.

2 Example: Why Do Inefficient Policies Arise?

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

Consider a principal who interacts with two agents in periods $t = 0, 1, ...$ In $t = 0$, the principal and each agent pay one another wages. Players have no liquidity constraints; denote by $w_{i,0} \in \mathbb{R}$ the net wage to agent i. After this payment, each agent i privately chooses a binary effort $e_{i,0} \in \{0, 1\}$ at cost $ce_{i,0}$. Agent i's effort determines his output $y_{i,0} \in \{0, H_i\}$, with $H_1 > H_2 > 0$. The probability that $y_{i,0} = H_i$ equals $pe_{i,0}$. After output is realized, the principal and each agent exchange bonus payments, with the net bonus to agent i denoted $\tau_{i,0} \in \mathbb{R}$. At the start of the second period ($t = 1$), the principal chooses one of the two agents. He repeatedly plays this stage game with the chosen agent, but has no further interactions with the agent who is not chosen. Players share a common discount factor $\delta \in (0, 1)$. The principal and agent i respectively earn $(1 - \delta) \sum_{i=1}^{2} (y_{i,t} - w_{i,t} - \tau_{i,t})$ and $(1 - \delta)(w_{i,t} + \tau_{i,t} - ce_{i,t})$ in period t, with $y_{i,t} = e_{i,t} = 0$ in $t \geq 1$ if agent i is not chosen.

As a benchmark, suppose that monitoring is **public**: all variables except effort are publicly observed, while effort is private. Assume that δ is such that either agent can be motivated to work hard in every period $t \geq 1$ if the principal chooses him. How might the principal motivate both agents to work hard in $t = 0$? Agent i can be motivated by either the expectation of a bonus or fine today ($\tau_{i,0}$) or a continuation payoff in period 1 onwards ($U_{i,1}$). So agent i's **total reward** for producing output $y_{i,0}$ equals

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

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

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

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

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

In Section 4, we show that this **dynamic enforcement** constraint is the only constraint imposed by the relational contract. In particular, we can construct an equilibrium in which the principal implements whatever policy (q_1, q_2) we choose and the agents earn any B_i that satisfies this constraint.

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

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

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

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

3 The Model

A single principal (player 0, 'she') and N agents (players $i \in \{1, ..., N\}$, each 'he') interact repeatedly in a dynamic game. Time is discrete and denoted by $t \in \{0, 1, \ldots\}$. Players are risk-neutral and share a common discount factor $\delta \in (0, 1)$. In each period, the principal makes a **decision** d_t from a set D_t. The decision determines how each agent i's effort $e_{i,t} \in \mathbb{R}_+$ determines his **outcome** $y_{i,t} \in \mathbb{R}_+$. Agent i incurs cost $c(e_{i,t})$, while the principal earns revenue equal to the sum of outcomes, $\sum_{i=1}^{N} y_{i,t}$. There are two rounds of transfers between the principal and each agent. The (net) ex-ante transfer to agent i is denoted $w_{i,t} \in \mathbb{R}$ and is paid before the agent accepts or rejects production, while the (net) ex-post transfer $\tau_{i,t} \in \mathbb{R}$ is paid after output is realized. We sometimes refer to these transfers, respectively, as wage and bonus payments, and we denote the vectors of wages and bonuses by w_t and τ_t. The principal sends a message $m_{i,t}$ to each agent i along with the wage payment $w_{i,t}$. Denote the vector of messages by m_t.

Technology In period t, a set of **available decisions** $D_t \subseteq \mathcal{D}$ and **state of the world** $\theta_t \in \Theta$ 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 $d_t \in D_t$ together with the state of the world θ_t and agent i's effort $e_{i,t}$ determine the marginal distribution over agent i's output: $P_i(y_{i,t}|\theta_t, d_t, e_{i,t})$. Note that outcomes are independent across agents conditional on the decision and the state of the world.

Timing The stage game has eight rounds.

1. 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 wage payments in $\mathbb{R}_+$ to send to one another. Define $w_{i,t} \in \mathbb{R}$ to be the (net) wage paid to agent i.

4. For each agent i, the principal chooses a message $m_{i,t} \in M$ to send to agent i, where M is a large message space.[0]

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 $\bar{u}_i(d_t, \theta_t) \geq 0$.

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

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

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

It is worth pausing to comment briefly on the timing. In our game, agents decide whether or not to take their outside options after they pay or receive the ex-ante transfer w_t. This assumption ensures that agent i can punish the principal by rejecting production following an off-path wage payment, which simplifies our equilibrium construction. A third round of transfers after accept/reject decisions but before efforts would not change our results. We could also allow the principal to send messages to each agent in every stage of the stage game without affecting any results.

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 e_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.

Histories and Strategies A history at the beginning of period t is $h_0^t = \{\theta_{t'}, D_{t'}, d_{t'}, w_{t'}, m_{t'}, p_{t'}, e_{t'}, y_{t'}, \tau_{t'}\}_{t'=0}^{t-1}$, from set $\mathcal{H}_0^t$. Let $h_x^t \in \mathcal{H}_x^t$ denote the within-period history immediately following the realization of variable x, so for example, $h_w^t = h_0^t \cup \{\theta_t, D_t, d_t, w_t\}$. For every agent i, let $\phi_i(h_x^t)$ denote agent i's private

[0]Formally, we assume that M's cardinality is at least as large as the Cartesian product of the set of all histories and the set of all functions $f : \mathbb{R} \to \mathbb{R}$. In practice, we can typically make do with a much smaller message space.

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 $\sigma_i|\phi_i(h^t)$. 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

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

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

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

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 Each player in a game with private information must form beliefs about the true history given his private information. 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, our main results restrict attention to **recursive equilibria (RE),** which are a recursive and hence relatively tractable refinement of **Perfect Bayesian Equilibrium**.

DEFINITION 1. A Perfect Bayesian Equilibrium σ^* is a **recursive equilibrium (RE)** if for any period t and on-path history $h_0^t \in \mathcal{H}_0^t$, $\sigma^*|h_0^t$ is a Perfect Bayesian Equilibrium of the game starting at h_0^t.

We say a relational contract is **self-enforcing** if it is a recursive equilibrium. In a Perfect Bayesian equilibrium, player i's actions at history h_0^t must be a best response to the other players' actions, given i's information about the true history $\phi_i(h_0^t)$. A recursive equilibrium requires that on the equilibrium path, player i's actions are a best-response not just given i's information $\phi_i(h_0^t)$, but given the **true** history h_0^t. This restriction applies only on the equilibrium path: the principal can renege on agent i without revealing that deviation to the other agents, even if subsequent play would not form a PBE. This equilibrium refinement is related to but weaker than belief-free equilibria, which is a standard recursive solution concept for games with private monitoring.[1]

In a recursive equilibrium, players form expectations **within** a period but must act as if they know the true history at the start of each period on the equilibrium path. This refinement implies that an RE is recursive on the equilibrium path, which is a non-trivial restriction on equilibrium play.[2] Our main analysis restricts to recursive equilibria, because they lead to clean and intuitive constraints on the relational contract and realistic history-dependent biases. Section 6.3 has a limited analysis of the full set of Perfect Bayesian Equilibria for a simple set of games and illustrates that our core intuition is not driven by the restriction to RE.

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

4 Sequential Inefficiency and Biased Decisions

If agents cannot observe one another's relationships, then they cannot coordinate to jointly punish a deviation by the principal. This section demonstrates how policies influence the resulting relational contracts. We develop straightforward necessary and sufficient condi-

[1]See Ely, Horner, and Olszewski (2005) for more details. In a belief-free equilibrium, a player's action at **every** history h_x^t must be a best response to $\sigma_{-i}^*|h_x^t$, rather than just histories at the start of each period on the equilibrium path. Because BFE are recursive and a subset of Perfect Bayesian Equilibrium, any BFE is also an RE. In a simultaneous-move repeated game with full support over private signals following any actions, BFE and RE select the same set of equilibria.

[2]Following a deviation, continuation play need not be recursive. We will nevertheless construct a tight bound on punishment payoffs that allow us to tractably analyze the resulting set of RE.

[3]By definition, continuation play $\sigma^*|h_0^t$ on the equilibrium path is an RE of the continuation game. So sequential surplus-maximization is well-defined.

tions for self-enforcing relational contract in the game with bilateral relationships. Then we show that backward-looking policies are an integral feature of surplus-maximizing relational contracts.

In the game with bilateral relationships, each agent observes only his own output and bonuses, and furthermore cannot communicate with his counterparts. While this assumption is stylized, we believe that it captures an important feature of many real-world business relationships: widespread punishments are difficult to coordinate, especially when some of those involved in the punishment were not involved in the original deviation. In our framework, while a betrayed agent can deny the principal surplus by taking his outside option, the *other* agents do not observe the deviation and so may not punish the deviator. We explore this assumption further in Section 6.

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 promise to bias future decisions towards that agent. 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, agent i's **reward scheme** B_i gives his expected payoff following each possible output realization. We consider the constraints that the bilateral relational contract imposes on each agent's reward scheme.

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

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

For each agent i and period $t \geq 0$, define $\xi_{i,t} = (m_{i,t}, w_{i,t}) \in \Xi = M \times \mathbb{R}$. A reward scheme $B_i : \not< \times \mathbb{R}_+ \times \mathcal{H}_d^t \rightarrow \mathbb{R}$ is **credible in** σ *if*

1. *It satisfies agent i's incentive-compatibility constraint: for each $h_d^t, \xi_{i,t}$, and $C_{i,t}$ on the*

equilibrium path,

$$C_{i,t} \in \underset{C_i | d_t, \theta_t}{\operatorname{argmax}} E_\sigma \left[B_i \left(\xi_{i,t}, y_{i,t} | h_d^t \right) \middle| h_d^t, \xi_{i,t}, C_i \right] - (1-\delta) C_i \tag{2}$$

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

$$\delta E_{\sigma^*} \left[\bar{U}_i \left(h_0^{t+1} \right) \middle| h_d^t \right] \leq B_i \left(\xi_{i,t}, y_{i,t} | h_d^t \right) \leq \delta E_{\sigma^*} \left[S_i \left(\sigma^*, h_0^{t+1} \right) \middle| h_d^t, \xi_{i,t}, y_{i,t} \right] \tag{3}$$

A credible reward scheme satisfies two conditions. First, agent i must be willing to exert effort $e_{i,t}$ if he expects to earn $B_i\left(h_y^t\right)$ following history h_y^t. This **effort IC** constraint is given by (2) and implies that $B_i\left(h_y^t\right)$ must vary in output $y_{i,t}$ to motivate effort. The second condition limits how much B_i can vary by bounding it from above and below. Agent i can never earn more surplus than δS_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 can earn no less than his punishment payoff $\bar{U}_i$. This **dynamic enforcement** constraint (3) must hold output-by-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.

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 states of the world, decisions, efforts, and outcomes as σ.*

Proof: See Appendix A.

Consider the moral hazard problem faced by an agent i in period t. The principal can motivate agent i to work hard by varying his contemporaneous bonus payment $\tau_{i,t}$ and his continuation surplus U_i with his output $y_{i,t}$. For a history h_e^t, define agent i's reward scheme $B_i : \phi_0(\mathcal{H}_y^t) \to \mathbb{R}$ as his expected continuation payoff for each possible outcome:

$$B_i(h_y^t) = E\left[(1-\delta)\, \tau_i + \delta U_i \middle| \phi_0(h_e^t), y_t \right].$$

An agent's reward scheme summarizes his incentives to exert effort through (2). 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 $(1-\delta)\,\tau_{i,t} \geq -\delta\left(U_i - \bar{U}_i\right)$. The principal can similarly walk away from his relationship with agent i by not paying him wages or bonuses. Importantly, she can do so without alerting any other agents because the other agents do not observe i's wages, bonuses, or output. So the principal is willing to pay agent i no more than her continuation surplus from her relationship with i. This logic gives an upper bound on $(1-\delta)\,\tau_{i,t}$, which in turn gives the upper bound on B_i given by (3). Together, these arguments prove the first statement of Lemma 1.

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. Importantly, the principal is willing to follow the equilibrium policy because she earns 0 in each period both on and off the equilibrium path. 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 biased decisions play an important role in surplus-maximizing relationships. Intuitively, future decisions determine both total surplus $\sum_{i=1}^{N} S_i$ and how much of that surplus is produced by each agent i. Future decisions that are biased towards agent i increase agent i's dyad-surplus S_i at the cost of decreasing total continuation surplus. Thus, biased decisions have a **direct cost**: they lead to lower continuation (and hence lower ex ante) total surplus. Lemma 1 suggests that biased decisions can also influence effort decisions. Increasing S_i relaxes agent i's dynamic enforcement constraint (3), leading to an

incentive benefit: the principal can credibly promise agent i larger rewards, which might motivate him to exert more effort. Of course, a policy that is biased towards agent i is also biased away from some agent $j \neq i$. So biased policies also have an **incentive cost**: biasing the future policy away from an agent makes it more difficult to motivate that agent.

The direct cost of a biased policy does not depend on past efforts and outputs, but the incentive cost and incentive benefit both do. If the upper bound of agent i's dynamic enforcement constraint binds—for instance, because he produced high output in the past—then relaxing that upper bound by biasing future decisions towards him leads to higher effort and so has a large incentive benefit. If agent j's dynamic enforcement constraint does not bind, then tightening it does not affect effort and so has no incentive costs. A surplus-maximizing relational contract entails biased decisions exactly when the incentive benefits outweigh both the incentive and direct costs of that bias. So surplus-maximizing policies are backward-looking: decisions will be biased towards those agents who have performed well—in the sense of producing output that indicates high effort—at the expense of those who have performed poorly.

Our main result considers these costs and benefits in a particularly tractable set of games. We restrict attention to a class of "smooth" repeated games and show that backward-looking policies are an integral part of surplus-maximizing relational contracts.

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

1. $D_t = \left\{ (d_1, ..., d_N) | \, d_i \in \mathbb{R}_+, \sum_{i=1}^N d_i \leq 1 \right\}$ *in each period. The distribution of* θ_t *depends only on* $\{\theta_{t'}\}_{t'=0}^{t-1}$.

2. *Outside options depend only on* θ_t: $\{\bar{u}_i(\theta_t)\}_{i=1}^N$ *.Effort costs* $c(\cdot)$ *are smooth, strictly increasing, and strictly convex.*

3. P_i *depends only on* d_i, θ, *and* e_i. *For each* $\{d_i, \theta\}$, P_i *is smooth in all arguments with density* p_i, *is strictly MLRP-increasing in* e_i, *has full support, and satisfies CDFC.* $E[y_i | d_i, \theta, e_i]$ *is strictly increasing and strictly concave in* $\{d_i, e_i\}$.

4. *Higher decisions are more informative: for any* $d_i \geq \tilde{d}_i$ *and* θ, *there exists a distribution* $R(\cdot|y)$ *with density* r such that for any $e_i, \bar{y}_i$,
$$\int_0^{\bar{y}_i} p_i(y_i|\theta, \tilde{d}_i, e_i) dy_i = \int_0^{\bar{y}_i} r_i(x_i|y_i) p_i(y_i|\theta, d_i, e_i) dy_i.$$

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

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

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

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

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

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

A critical feature of these games is that d simultaneously affects all agents' output. Maximizing the surplus produced by agent i requires $d_i = 1$, which requires $d_j = 0$ for all other agents. Relaxing one agent's dynamic enforcement constraint necessarily entails decreasing the surplus produced by some other agent. As a result, biased decisions are typically an important part of surplus-maximizing relationships in smooth games.

PROPOSITION 1. *Consider a smooth game with bilateral relationships. In any surplus-maximizing relational contract σ^*,*

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

2. **Backward-looking policies are optimal**: *For any agents i and j, let E_t be a set of histories h_0^{t+1} such that : (i) $e_{i,t} \in (0, e_i^{FB}(d_{i,t}, \theta_t))$, (ii) $y_{i,t} > y_i^*(d_{i,t}, \theta_t, e_{i,t})$, (iii) $y_{j,t'} < y_j^*(d_{j,t'}, \theta_{t'}, e_{j,t'})$ for all $t' \leq t$, and (iv) $d_{i,t+1}^* < 1$ and $d_{j,t+1}^* > 0$ with positive probability. For almost every $h_0^{t+1} \in E_t$, $\sigma^* | h_0^{t+1}$ is not surplus-maximizing.*

Proof: See Appendix A.

The first statement of Proposition 1 holds because larger d_i increases agent i's dyad-surplus, provides a more precise signal of agent i's effort, and relaxes (3) for agent i in previous periods. So any surplus-maximizing relational contract will use the full "budget" of d_i—the only question is what weight is assigned to each agent.

For the second statement, consider a smooth game, let h_0^{t+1} satisfy the conditions of the proposition, and suppose σ^* is sequentially surplus-maximizing. Fixing dyad-surplus for all agents $k \notin \{i, j\}$, define $\bar{S}_i(S_j)$ as agent i's **maximum** dyad-surplus in the continuation game if agent j's dyad-surplus equals S_j. If $\bar{S}_i$ is differentiable, then $\frac{\partial \bar{S}_i}{\partial S_j} = -1$ in any sequentially surplus-maximizing equilibrium. In particular, increasing S_i (and decreasing S_j) by biasing future decisions towards agent i has a second-order direct cost. Decreasing S_j has no incentive cost if agent j has never produced high output because the upper bound of j's dynamic enforcement constraint does not bind. However, increasing S_i has a first-order incentive benefit because it means agent i can be credibly promised a larger reward for high output, which motivates him to work harder. This leads to a first-order incentive benefit because $e_{i,t}^* < e_i^{FB}$. So biasing future decisions has a first-order benefit and a second-order cost.

The central difficulty of the proof is showing that $\bar{S}_i$ exists and is differentiable. We construct a perturbation of σ^* by increasing $d_{i,t+1}^*$, decreasing $d_{j,t+1}^*$, and increasing $e_{i,t}^*$. Changing $d_{i,t+1}^*$ or $e_{i,t}^*$ changes the distribution over output and hence continuation play. In principle, these changes affect the entire vector of efforts that can be sustained in every prior period. We construct a mapping from each player's output to continuation play to ensure that all agents $k \notin \{i, j\}$ face the same incentives as in σ^*. Agent i can be motivated to choose a larger $e_{i,t}^*$ because $d_{i,t+1}^*$ and therefore dyad-surplus in period $t+1$ onward is larger, while $e_{i,t+1}^*$ and the distribution over continuation play from period $t+2$ onward is held constant by construction. We show that this perturbation smoothly impacts both S_i and S_j, holding S_k fixed for all $k \notin \{i, j\}$. It follows that $\bar{S}_i$ is differentiable in S_j, proving Proposition 1.

The (incentive and direct) costs and (incentive) benefits of biased decisions are relatively easy to calculate in smooth games. However, the intuition behind backward-looking policies is not limited to these settings. Lemma 1 applies to many settings in which the principal can affect agents' outside options, outputs, and the precision of those outputs. We explore several examples in the next section.

5 Applications of Biased Policies

Biased relational contracts can arise in a many different settings. In this section, we use two simple examples to illustrate the types of biases that might arise in a relationship. First, we consider hiring decisions and prove that a firm might optimally delay hiring after demand increases. Then we show how a firm might distort irreversible investments or promotions to better motivate its divisions or employees. For simplicity, we will assume $N = 2$, $\bar{u}_i = 0$ and $e_{i,t} \in \{0, 1\}$ with cost $ce_{i,t}$ in both examples.

5.1 Hiring and Firing

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

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. For convenience, we assume that if $d_t = 1$, then agent 1 is hired.*[4]
- *If agent i is not hired, then $y_{i,t} = 0$. Otherwise, $y_{i,t} = \theta_t e_{i,t}$ if $d_t = 1$ and $y_{i,t} = \theta_t \alpha e_{i,t}$ with $\alpha < 1$ if $d_t = 2$.*

The principal is a firm that faces demand 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 increases, it remains robust thereafter. The return to an agent's effort in period t is determined by both demand and the number of agents hired in t. We assume that marginal productivity is decreasing in the number of workers ($\alpha < 1$). The optimal number of employees depends on demand and the effort chosen by each worker. We assume that if agents exert effort, then the firm maximizes myopic profit by hiring two workers if $\theta_t = R$ and one worker if $\theta_t = W$.

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

[4] This restriction is without loss for our result.

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

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

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

If $\theta_0 = W$*, then one surplus-maximizing relational contract satisfies the following:* $e_{i,t} = 1$ *in any period in which agent* i *is hired. If* $\theta_t = R$ *for the first time in period* t*, then* $d_t = 1$ *with probability* γ *and otherwise* $d_t = 2$*. In every* $t' > t$*,* $d_{t'} = d_t$*.*

Proof: See Appendix A.

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

This hiring delay could take many different forms. In Proposition 2, we demonstrate that one surplus-maximizing distortion is for the firm to make a *once-and-for-all* decision whether or not to expand as demand grows. While the particulars of the principal's distorted hiring policy depend on our stylized assumptions, this result illustrates that the optimal relational contract may entail substantial and long-lasting distortions.

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

5.2 Irreversible Investments

Suppose a principal can choose to make a permanent investment in one of her agents which increases that agent's productivity. This investment can be interpreted as training or human capital, or more broadly as a promotion or another organizational decision that increases the returns from one agent's efforts. Some workers may benefit more from this investment than others. Which agent should the principal choose?

In this example, we show that the principal might optimally award the investment in a (potentially biased) tournament among her agents. The agent who performs "best" according to this tournament is chosen, even if he might not have the largest return from investment. This example also illustrates the types of biases that can arise if output is continuous.

DEFINITION 5. *The **irreversible investment game** has* $|\Theta| = 1$ *and the following features:*

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

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

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

Define

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

as the likelihood ratio for output y_i given decision d. Because P_i is MLRP-increasing in e_i, L_i is strictly increasing in y_i. In the irreversible investments game, the principal chooses one of the two agents to receive the investment at the end of the first period. Agents have identical productivities without the investment, but agent 1's productivity benefits more from the investment than agent 2's. The investment is irreversible.

The principal can use the promise of an investment to make a large reward to the chosen agent credible. As a result, the principal can potentially motivate both agents in period 0 by promising to invest in (and monetarily reward) whichever agent produces high output. The result is a tournament in which the less-efficient agent may receive the investments if he performs well in the first period. This tournament will typically be "biased," since the principal wants to maximize the probability that the efficient agent is chosen subject to the constraint that both agents exert effort in the first period.

PROPOSITION 3. *Consider the irreversible investment game with bilateral relationships. There exists* $0 \leq \underline{\delta} < \bar{\delta} < 1$ *and* $\bar{\Delta} > 0$ *such that if* $\underline{\delta} < \delta < \bar{\delta}$ *and* $\Delta < \bar{\Delta}$, *any surplus-maximizing relational contract* σ^* *satisfies:.*

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

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

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

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

Proof: See Appendix A.

If the agents' productivities are not too different, then the principal finds it optimal to choose an agent who produces high output. Because both agents have the opportunity to "win" the investment, both are willing to work hard in the first period. After the principal chooses one agent, that agent's dynamic enforcement constraint is slack and so he is willing to continue working hard. However, the principal can no longer credibly promise strong monetary incentives to the other agent, who becomes "discouraged" and stops exerting effort.

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

6 The Role of Private Monitoring

Bilateral monitoring plays an essential role in Lemma 1 and so is an integral ingredient of our results. This section further explores this monitoring assumption via three extensions. Section 6.1 considers surplus-maximizing relationships if monitoring is public. As suggested in Section 2, decisions are always chosen to maximize continuation surplus in this setting and so surplus-maximizing relationships never entail backward-looking policies. In Section 6.2, we revisit the hiring game and consider a hybrid monitoring structure that allows the agents to imperfectly coordinate punishments. If this coordination is not perfect, we show that biased decisions may continue to play an important role in optimal relationships. Section 6.3 revisits our restriction to Recursive Equilibria and demonstrates that this restriction does not drive our results: backward-looking policies remain important in surplus-maximizing Perfect Bayesian Equilibria.

6.1 Sequential Efficiency Under Public Monitoring

Define the **game with public relationships** as in Section 3, with the following two differences. First, all variables except $\{e_{i,t}\}$ are observed by every player, while efforts remain private. Second, the players are assumed to have access to a public randomization device in each round of the stage game. [5]

Backward-looking policies have a direct cost because they reduce total continuation surplus. Furthermore, all agents can observe a deviation by the principal in the game with public relationships. So the principal stands to lose the future production of **every** agent if she reneges on any **one** relationship. Backward-looking policies potentially make the principal **more** willing to renege on an agent because she has less continuation surplus to lose following a deviation. Rather than biasing her future decisions, the principal can reward or punish all agents using transfers and choose a policy that maximizes total continuation surplus.

As a result, 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 the optimal policy.

PROPOSITION 4. *Consider the game with public relationships. Any surplus-maximizing*

[5] This assumption is for convenience - we could add public randomization devices to the game with bilateral relationships without affecting our results.

relational contract is sequentially surplus-maximizing.

Proof: See Appendix A.

Proposition 4 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 the reward scheme $B_i(h_y^t)$ defined as in Lemma 1. This reward scheme must satisfy (2) to motivate agent i to work hard in period t. As before, the relational contract constrains the *maximum* and *minimum* bonuses rewards that can be promised to an agent. As before, agent i can earn no less than $\delta \bar{U}_i$, since otherwise he would rather renege and be punished than continue the relationship. In contrast to the game with bilateral relationships, however, the principal cannot renege on a single agent without the other agents observing this deviation. Therefore, the principal faces an **aggregate** temptation to deviate: if she reneges on any one relationship, then she optimally betrays every other agent as well. A necessary[6] condition for B_i to be part of an equilibrium is therefore

$$\sum_{i=1}^{N} b_i \leq \delta \sum_{i=1}^{N} S_i.$$

Notice that the right-hand side of this inequality is equal to **total continuation surplus.** Consider a relational contract that prescribes inefficient on-path continuation play. Then there exists some efficient continuation that leads to strictly higher continuation surplus $\sum_{i=1}^{N} S_i$. Holding effort constant, *ex ante* total surplus is higher under this efficient continuation by definition. Moreover, higher $\sum_{i=1}^{N} S_i$ relaxes the relational contract's aggregate dynamic enforcement constraint, which implies that agents can be credibly promised stronger incentives. So any sequentially inefficient relational contract is strictly dominated by a sequentially surplus-maximizing relational contract that both has higher continuation surplus and higher effort. This proves Proposition 4.

Agents can perfectly coordinate to jointly punish the principal in the game with public relationships, so the principal might lose as much as $\sum_{i=1}^{N} S_i$ following a deviation. In the game with bilateral relationships, the principal can deviate in her relationship with agent i without any other agents learning of that deviation. So the principal loses no more than

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

S_i from such a deviation. These different punishment payoffs drive the differences between Propositions 1 and 4. Note that public relationships imply that agents **immediately** and **perfectly** coordinate to punish a deviating principal: for example, if an employer withholds a bonus from a deserving worker, then she faces sanctions from her entire workforce. The next subsection illustrates what might happen if this coordination is imperfect.

6.2 Biased Decisions Under Imperfect Coordination

This section considers the hiring example from Section 5.1 and shows that biased decisions might be optimal even if agents can **imperfectly** coordinate to punish a deviating principal. This result suggests that Proposition 4 strongly relies on public relationships, while the forces driving Proposition 1 might continue to play a role even if agents imperfectly observe one another's relationships.

In the hiring game from Section 5.1, suppose that relationship breakdowns are **ϵ-private**: the **first** time an agent rejects production, **all** agents observe this choice with probability $1-\epsilon$. With probability $\epsilon > 0$, only the principal observes the agent's rejection. Subsequent rejections are observed only by the principal. In any surplus-maximizing equilibrium of this game, agent i rejects production only following a deviation. Therefore, this monitoring structure gives agents a "once and for all" chance to communicate and coordinate their punishments following a deviation. If the principal deviates in her relationship with agent i, then i will reject production. All agents will observe this rejection with probability $1-\epsilon$ and can jointly punish the principal if they do. The rejection remains private with probability $\epsilon > 0$, in which case only agent i punishes the principal.

So long as $\epsilon > 0$, Proposition 5 shows that the principal's optimal hiring policy will be biased for at least some parameter values.

PROPOSITION 5. *Consider the hiring game with ϵ-private monitoring with $\epsilon > 0$. Then there exists an open set of parameters such that for those parameter values, no surplus-maximizing relational contract is sequentially surplus maximizing.*

Proof: See Appendix A.

The basic intuition for Proposition 5 is fairly straightforward. Suppose that the principal reneges on her relationship with agent i. Then agent i will reject production in all subsequent periods. With probability $(1-\epsilon)$, all players observe this rejection and jointly punish the principal, resulting in a total of $\delta \sum_{i=1}^{N} S_i$ surplus lost during the punishment. Otherwise, only agent i observes a deviation and so only δS_i surplus is lost. Note that agent i's dyad-surplus is always lost, while the other agents' surplus is lost with probability $1 - \epsilon < 1$. So the principal can make stronger incentives to i credible by biasing future decisions towards i—in this case, by refraining from hiring additional agents.

This basic intuition masks considerable complexity that arises from this monitoring structure. Unlike the proof of Lemma 1, we can no longer make the principal indifferent among possible policies, so we must ensure that she is willing to follow the equilibrium policy in each period. While these additional incentive constraints make a general analysis very difficult, the optimal policy in the hiring game depends only on current and past demand and hiring, both of which are publicly observed. All agents can observe when the principal chooses the incorrect decision and so all agents can jointly punish such a deviation. This ensures that the principal is willing to follow the surplus-maximizing policy.

A full exploration of this monitoring structure is very difficult for the reasons outlined above. Proposition 5 illustrates that, at least in our hiring example, our central intuition does not depend on agents being totally unable to coordinate. Backwards-looking policies can play an important role in the surplus-maximizing relational contract even if agents might imperfectly observe one another's relationships.

6.3 Biased Decisions in Surplus-Maximizing PBEs

Recursive Equilibria are relatively tractable but potentially entail a loss of generality. This section considers the full set of Perfect Bayesian Equilibria in the context of a simple class of games. Biased decisions and backwards-looking policies may continue to be surplus-maximizing, which shows that the intuition from Proposition 1 does not depend on the restriction to Recursive Equilibria.

The principal difficulty in extending Proposition 1 is that Perfect Bayesian Equilibria are not recursive in games with private monitoring. Different players observe different variables and so potentially form different beliefs about the true history in each period. Each player responds optimally to others' predicted actions given these beliefs, but unlike a Recursive Equilibrium, a player's strategy need not be a best response at the **true** history. So continuation play at any given history does not necessarily form an equilibrium.

This complication implies that we must modify the definitions of surplus-maximizing and sequentially surplus-maximizing relational contracts in order to extend the analysis to Perfect Bayesian Equilibria. In particular, continuation play at a given history might attain payoffs that are impossible to attain at the start of the game. Our first result considers a class of repeated games and shows that even though continuation payoffs at a **given** history might not be replicable at the start of the game, *ex ante* **expected** continuation payoffs in any period can be replicated by some equilibrium at the start of the game.

LEMMA 2. *Suppose $F(\cdot|\cdot)$ does not depend on history. Then for any $t \geq 0$, there exists a PBE σ^* such that $E_{\sigma^*}\left[\sum_{i=1}^{N} S_i(\sigma^*, h_0^t)|h_0^t \in \mathcal{H}_0^t\right] = \bar{V}$ if and only if there exists a PBE $\tilde{\sigma}^*$ such that $\sum_{i=1}^{N} S_i(\sigma^*, h_0^0) = \bar{V}$.*

Proof: See Appendix A.

The proof of Lemma 2 begins by establishing an appropriate extension of the necessary and sufficient conditions from Lemma 1 for Perfect Bayesian Equilbria. This extension uses a similar construction to the proof of Lemma 1, though care must be taken to ensure that each agent has the correct beliefs to make them willing to exert the equilibrium effort at each history. A statement of the resulting incentive and dynamic enforcement constraints for Perfect Bayesian Equilibria may be found in Appendix A.

Given this result, we turn to the "if" statement of Lemma 2. Given an equilibrium σ^*, consider the following equilibrium $\tilde{\sigma}^*$: at the start of the game, the principal chooses a history h_0^t according to the distribution induced by σ^*. In period 0, the principal sends a message to agent i that consists of $\phi_i(h_0^t)$. Play then continues as in $\sigma^*|h_0^t$. If the principal is willing to perform this initial randomization, then all players have the same information that they have at h_0^t, and so all players are willing to play as in $\sigma^*|h_0^t$. As in the equilibrium constructed in Lemma 1, transfers can be used to set the principal's payoff equal to 0 at every history. So the principal is completely indifferent across histories and hence is willing to perform the desired randomization. This proves the "if" statement. The "only if" statement is straightforward: given $\tilde{\sigma}^*$, define σ^* in which players play $\tilde{\sigma}^*$ after $t-1$ periods of playing the static Nash equilibrium.

Given Lemma 2, we can define surplus-maximizing and sequentially surplus-maximizing Perfect Bayesian Equilibria. let $\bar{V}$ be the **maximum expected total surplus** attainable in a PBE. Then a **surplus-maximizing PBE** σ^* satisfies $\sum_{i=1}^{N} S_i(\sigma^*, h_0^0) = \bar{V}$. A **sequentially surplus-maximizing PBE** satisfies $E_{\sigma^*}\left[\sum_{i=1}^{N} S_i(\sigma^*, h_0^t)|h_0^t \in \mathcal{H}_0^t\right] = \bar{V}$ for every period $t \geq 0$.

Our main result in this section considers a class of games for which sequentially surplus-maximizing PBE are particularly simple and give conditions under which surplus-maximizing PBE are not sequentially surplus-maximizing. Backward-looking policies arise for much the same reason as in Proposition 1: biasing future decisions towards an agent who has performed well in the past strengthens the incentives that can be promised to that agent in previous periods and leads to more effort.

PROPOSITION 6. Consider a smooth game such that $F(\cdot|\cdot)$ does not depend on history. Suppose that for any θ, d_i, e_i, the random variable y_i may be written

$$y_i = x_i + \gamma_i(\theta, d_i)$$

where $\gamma_i : \Theta \times D \to \mathbb{R}$ is smooth, with $\frac{\partial \gamma_i}{\partial d_i} > 0$, $\frac{\partial^2 \gamma_i}{\partial d_i^2} < 0$, and $\lim_{d_i \to 0} \frac{\partial \gamma_i}{\partial d_i} = \infty$, while $x_i \tilde{\ } \tilde{P}_i(\cdot|\theta, e_i)$. Suppose that no sequentially surplus-maximizing RE is surplus-maximizing. Then no sequentially surplus-maximizing PBE is surplus-maximizing.

Proof: See Appendix A.

Proposition 6 restricts attention to a class of games in which the effects of d_i and e_i on output are additively separable: the decision d_i only affects the value of γ_i, while e_i only affects the distribution over x_i. In each period t of a sequentially surplus-maximizing equilibrium, d_t^* is determined by the expression

$$\frac{\partial \gamma_i}{\partial d_i}(\theta_t, d_{i,t}^*) = \frac{\partial \gamma_j}{\partial d_j}(\theta_t, d_{j,t}^*).$$

In particular, d_t^* depends only on θ_t and so cannot change in response to the past performance of the agents. But then each agent's relationship with the principal is stationary, so the maximum effort that that agent can be motivated to choose is also stationary. Agent i's beliefs about the true history are irrelevant for his continuation play in this stationary equilibrium, so any sequentially surplus-maximizing PBE is also a sequentially surplus-maximizing Recursive Equilibrium. So if backward-looking policies are surplus-maximizing in an RE, they are surplus-maximizing in a PBE as well.

This class of games demonstrates the central tension in our model continues to play a role when we consider the full set of PBE. A sequentially surplus-maximizing equilibrium cannot implement a policy that responds to past performance, but such a backward-looking policy can make stronger incentives to the agents credible and so induce higher effort.

7 Discussion and Conclusion

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

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

In practice, the principal might try to make some aspects of her payments public information in order to help her agents coordinate. For example, the principal might implement a fixed, public bonus pool from which all of her agents are paid. If agents are able to observe this bonus pool, then they can jointly punish the principal whenever she withholds some of the pool. For such a scheme to work, the principal must be able to commit both to make the bonus pool public and to refrain from secretly withdrawing funds from it. Such pools might also entail their own dynamic inefficiencies, particularly if the size of the pool depends on the outputs realized in each period.

We justify bilateral punishments as the result of bilateral monitoring between the principal and each agent: other agents do not observe and so cannot punish the principal if she betrays agent i. An alternative approach (taken in Levin (2002)) would be to make a behavioral restriction to equilibria that satisfy a "bilateral punishment" condition. Formally analyzing such a restriction lies outside the scope of this paper. Nevertheless, our results suggest that such a condition might lead to similarly biased decisions and backward-looking policies.[7]

Finally, two critical assumptions in our framework are that (i) each agent's effort affects only his own output and (ii) the principal earns the sum of agent outputs. An important extension would be to consider cases in which agent efforts are either substitutes or complements in profit. The techniques we use in this paper do not directly extend to these settings. In particular, it is substantially harder to ensure that the principal implements the equilibrium policy in such environments. We conjecture that conditions similar to those in Lemma 1 are necessary (though not sufficient) if efforts are substitutes. If efforts are complements, then the relational contract is further complicated because it must deter the principal from simultaneously reneging on multiple relationships at once.

In our setting, firms design policies in part to credibly reward those workers, divisions, and suppliers that have performed well in the past. The nature of the resulting decisions—and hence the momentum of a given firm—depends critically on the history of that firm. Therefore, relational contracts provide a justification for the tremendous heterogeneity among organizations in many markets.

[7] For example, suppose that agents only condition their actions on their own past efforts, outputs, and payments, as well as past states of the world and decisions. Appropriately defined, this restriction would select the set of recursive equilibria from our analysis.

Appendix A: Proofs

Proof of Lemma 1

1 : Suppose σ^* is an RE and define $B_i : \Xi \times \mathbb{R}_+ \times \mathcal{H}_d^t \to \mathbb{R}$ by

$$B_i(\xi_{i,t}, y_{i,t}|h_w^t) = E_{\sigma^*}\left[(1-\delta)\,\tau_{i,t} + \delta U_i|\, h_d^t, \xi_{i,t}, y_{i,t}\right]. \tag{6}$$

Consider an on-path history h_0^t. Then $\sigma^*|h_0^t$ is a Perfect Bayesian Equilibrium of the continuation game. In particular, for each on-path D_t, θ_t, d_t, and ξ_t, agent i is willing to choose $e_{i,t}$ only if

$$e_{i,t} \in \underset{e_i \in \mathbb{R}_+}{\operatorname{argmax}}\, E_{\sigma^*}\left[(1-\delta)\,\tau_{i,t} + \delta U_i|\, h_d^t, \xi_{i,t}, e_{i,t} = e_i\right] - (1-\delta)\,c(e_i). \tag{7}$$

We can rewrite (7) as

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

Conditional on $y_{i,t}$, $e_{i,t}$ does not affect continuation play. So this constraint implies (2).

Suppose $B_i\left(\xi_{i,t}, y_{i,t}|h_d^t\right) < \delta E_{\sigma^*}\left[\bar{U}_i\left(h_0^{t+1}\right)\big|\, h_d^t, \xi_{i,t}, y_{i,t}\right]$ at some on-path history h_y^t. 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. So the left-hand side of (3) must hold in any RE. Suppose that the principal refuses to pay $\tau_{i,t} \geq 0$ following h_y^t. Following this deviation, the refinement to RE has no bite. We claim the principal's continuation payoff after this deviation is bounded from below by

$$E_{\sigma^*}\left[\Pi\left(\sigma^*, h_0^{t+1}\right) - \sum_{t'=0}^{\infty}(1-\delta)\,\delta^{t'}\left(y_{i,t+t'} - w_{i,t+t'} - \tau_{i,t+t'}\right)\Bigg|\, h_y^t\right]. \tag{8}$$

To prove this claim, consider the following strategy for the principal following a deviation in $\tau_{i,t}$. Agent i observes this deviation, but no other agents do. Denote all variables that are observed by at least one agent $j \neq i$ by $\cup_{j\neq i}\phi_j\left(h_0^{t'}\right)$. In each period $h_0^{t'} \in \mathcal{H}_0^{t'}$, the principal plays according to $\sigma^*|\cup_{j\neq i}\phi_i\left(h_0^{t'}\right)$, 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 $j \neq i$. The principal's payoff from this strategy equals (8), so this strategy bounds the principal's payoff from below.

The principal is willing to pay $\tau_{i,t} \geq 0$ only if

$$(1-\delta) E_{\sigma^*} \left[\tau_{i,t} | \, h_y^t \right] \leq E_{\sigma^*} \left[\sum_{t'=1}^{\infty} (1-\delta) \, \delta^{t'} \left(y_{i,t+t'} - w_{i,t+t'} - \tau_{i,t+t'} \right) \middle| \, h_y^t \right]. \tag{9}$$

Adding δU_i to both sides of this expression yields

$$E_{\sigma^*} \left[(1-\delta)\tau_{i,t} + \delta U_i(\sigma^*, h_0^{t+1}) | h_y^t \right] \leq \delta E_{\sigma^*} \left[S_i(\sigma^*, h_0^{t+1}) | h_y^t \right].$$

This inequality must hold a fortiori in expectation. Because $h_d^t, \xi_{i,t}$, and $y_{i,t}$ are elements of h_y^t, iterating expectations and applying the definition of B_i yields the right-hand inequality of (3).

2: We construct a RE σ^* from σ. Recursively define σ^* as follows:

1. If $t = 0$, then $h_0^{t,*} = h_0^t = \emptyset$, the unique null history. Otherwise, begin with $h_0^t, h_0^{t,*} \in \mathcal{H}_0^t$ such that (i) h_0^t is on-path for σ, (ii) $h_0^{t,*}$ is on-path for σ^*, and (iii) h_0^t and $h_0^{t,*}$ induce identical continuation games. Define $\sigma^* | h_0^{t,*}$ as follows, where starred variables represent actions played in σ^* and unstarred variables represent actions played in σ.

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

   $$w_{i,t}^* = E_\sigma \left[y_{i,t} - \frac{1}{1-\delta} \left(B_i \left(\xi_{i,t}, y_{i,t} | h_d^t \right) - \delta S_i \left(\sigma, h_0^{t+1} \right) \right) \middle| \, h_d^t, \xi_{i,t} \right].$$

   Note $w_{i,t}^* \geq 0$ by (3). The principal sends a message to agent i:

   $$m_{i,t}^* = \left\{ h_0^{t,*}, a_{i,t}, e_{i,t}, \left\{ B_i \left(\xi_{i,t}, y_{i,t} | h_d^t \right) - \delta E_\sigma \left[S_i \left(\sigma, h_0^{t+1} \right) | h_d^t, \xi_{i,t}, y_{i,t} \right] \right\}_{y_{i,t} \geq 0} \right\}.$$

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

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

   $$(1-\delta)\tau_{i,t}^* = B_i \left(\xi_{i,t}, y_{i,t}^* | h_d^t \right) - \delta E_\sigma \left[S_i \left(\sigma, h_0^{t+1} \right) | h_d^t, \xi_{i,t}, y_{i,t}^* \right]$$

   Note that $\tau_{i,t}^* \leq 0$ by (3). Given $m_{i,t}^*$ and the realized $y_{i,t}^*$, agent i can perfectly infer $\tau_{i,t}^*$.

(d) Let $h_0^{t+1,*}$ be the realized history at the end of period t. The principal draws $h_0^{t+1} \in \mathcal{H}_0^{t+1}$ from $\sigma|\{h_e^t, y_t\}$. Continuation play $\sigma^*|h_0^{t+1,*}$ is constructed by repeating steps (a)-(d) using h_0^{t+1} and $h_0^{t+1,*}$.

2. If agent i observes a deviation, then he takes his outside option and pays no transfers in this and every subsequent period. If the principal observes a deviation, then $m_{j,t'} = w_{j,t'} = \tau_{j,t'} = 0$ for each agent $j \in \{1, \ldots, N\}$ in each future period. If agent i deviates, the principal chooses d_t to min-max agent i. Otherwise, d_t is chosen uniformly at random following a deviation.

First, note that past effort choices do not affect current play. Moreover, agents know the true history at the start of a period $h_0^{t,*}$ whenever they take actions in that period. Therefore, if we can show that no player has a profitable deviation from σ^*, it immediately follows that $\sigma^*|h_0^{t,*}$ is a PBE for every on-path $h_0^{t,*}$ and hence σ^* is a RE. Furthermore, both total continuation surplus and i-dyad surplus for every $i \in \{1, ..., N\}$ are identical in $\sigma^*|h_0^{t,*}$ and $\sigma|h_0^t$ by construction.

First, consider the principal. For any on-path $h_d^{t,*}$ and each agent $i \in \{1, \ldots, N\}$, the distribution over $y_{i,t}^*$ is the same as in $\sigma|h_d^t$. So

$$E_{\sigma^*}\left[y_{i,t}^* - w_{i,t}^* - \tau_{i,t}^* \middle| h_d^{t,*}\right] = 0$$

and $\Pi\left(\sigma^*, h_d^{t,*}\right) = 0$. If the principal deviates in d_t^*, $w_{i,t}^*$, or $m_{i,t}^*$, then each agent i either knows that this action is a deviation or not. If agent i knows these actions are a deviation, then the principal earns 0 from agent i because that agent rejects production in this and all future periods. If agent i does not know these actions are a deviation, then the principal must announce some on-path history $\tilde{h}_d^t$ such that $\phi_i(\tilde{h}_d^t) = \phi_i(h_d^{t,*})$. But then $E_{\sigma^*}[y_{i,t'}^* - w_{i,t'}^* - \tau_{i,t'}^*|\tilde{h}_d^t] = 0$ in every $t' \geq t$ because $\tilde{h}_d^t$ is on-path. So the principal cannot profitably deviate in $d_t^*, w_{i,t}^*$, or $m_{i,t}^*$. The principal likewise has no profitable deviation from $\tau_{i,t}^*$ because $\tau_{i,t}^* \leq 0$. So the principal has no profitable deviation on the equilibrium path.

Suppose the principal deviates off the equilibrium path. As with the argument in the previous period, if this deviation is detected by agent i then the principal earns 0 from agent i. If this deviation is not detected by agent i, then it must be in $d_t^*, w_{i,t}^*$, or $m_{i,t}^*$, since $\tau_{i,t}^* > 0$ is always detected as a deviation. But the principal earns payoff 0 following a deviation $d_t^*, w_{i,t}^*$, or $m_{i,t}^*$ by the argument above. So she has no profitable deviation off the equilibrium path.

Consider agent i. At each on-path history $h_0^{t,*}$, $E_{\sigma^*}\left[u_{i,t}\middle| h_0^{t,*}\right] = E_{\sigma^*}\left[y_{i,t} - c(e_{i,t})\middle| h_0^{t,*}\right]$. So $U_i\left(\sigma^*, h_0^{t,*}\right) = S_i\left(\sigma^*, h_0^{t,*}\right)$. By construction of σ^*, $S_i\left(\sigma^*, h_0^{t,*}\right) = S_i\left(\sigma, h_0^t\right)$. Since $w_{i,t}^* \geq 0$, agent i has no profitable deviation from $w_{i,t}^*$. After agent i observes an on-path $\left(w_{i,t}^*, m_{i,t}^*\right)$ and chooses $e_{i,t}^*$, he earns

$$E_{\sigma^*}\left[(1-\delta)\tau_{i,t} + \delta S_i(\sigma^*, h_0^{t+1,*})\middle| h_d^{t,*}, w_{i,t}^*, m_{i,t}^*, e_{i,t}^*\right] - c(e_{i,t}^*),$$

since he can perfectly infer $h_d^{t,*}$ from D_t^*, θ_t^*, d_t^*, and $m_{i,t}^*$. Plugging in the definition of $\tau_{i,t}$ yields

$$E_{\sigma^*}\left[B_i\left(\xi_{i,t}, y_{i,t}^*|h_d^t\right) - \delta E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right)|h_d^t, w_{i,t}, y_{i,t}^*\right] + \delta S_i(\sigma^*, \tilde{h}_0^{t+1})\middle| h_d^{t,*}, w_{i,t}^*, m_{i,t}^*, e_{i,t}^*\right] - c(e_{i,t}^*).$$

Now, $E_{\sigma^*}\left[B_i(\xi_{i,t}, y_{i,t}^*|h_d^t)|h_d^{t,*}, w_{i,t}^*, m_{i,t}^*, e_{i,t}^*\right] = E_\sigma\left[B_i(\xi_{i,t}, y_{i,t}|h_d^t)|h_d^t, e_{i,t}^*\right]$ because θ_t^*, d_t^* are the same in h_d^t and $h_d^{t,*}$. Similarly, $E_\sigma\left[S_i(\sigma, h_0^{t+1})|h_d^t, \xi_{i,t}, y_{i,t}^*\right] = E_{\sigma^*}\left[S_i(\sigma^*, h_0^{t+1,*})|h_d^{t,*}, w_{i,t}^*, m_{i,t}^*, y_{i,t}^*\right]$ by construction. Therefore, the agent is willing to choose $e_{i,t}^*$ so long as

$$e_{i,t}^* \in \arg\max_{e_i} E_\sigma\left[B_i\left(\xi_{i,t}, y_{i,t}|h_d^t\right)|h_d^t, e_i\right] - c(e_i).$$

Effort $e_{i,t}^*$ satisfies this constraint because (2) holds. Off the equilibrium path, continuation play is independent of $e_{i,t}$ and so the agent optimally chooses $a_{i,t} = 0$.

Following any deviation in $\tau_{i,t}^* < 0$, agent i earns continuation surplus $\bar{U}_i\left(h_0^{t+1,*}\right)$. Agent i observes or infers $h_d^{t,*}, w_{i,t}^*, m_{i,t}^*, y_{i,t}^*$ in σ^*. So agent i is willing to pay $\tau_{i,t}^* < 0$ if

$$-(1-\delta)\tau_{i,t}^* \leq \delta E_{\sigma^*}\left[U_i\left(\sigma^*, h_0^{t+1,*}\right) - \bar{U}_i\left(h_0^{t+1,*}\right)\middle| h_d^{t,*}, w_{i,t}^*, m_{i,t}^*, y_{i,t}^*\right]$$

Recall that $U_i(\sigma^*, h_0^{t+1,*}) = S_i(\sigma^*, h_0^{t+1,*})$ by construction. Moreover,

$$E_{\sigma^*}\left[S_i\left(\sigma^*, h_0^{t+1,*}\right)|h_d^{t,*}, w_{i,t}^*, m_{i,t}^*, y_{i,t}^*\right] = E_\sigma\left[S_i(\sigma, h_0^{t+1})|h_d^t, \xi_{i,t}, y_{i,t}^*\right]$$

because continuation dyad-surplus in σ^* following $h_d^{t,*}, w_{i,t}^*, m_{i,t}^*, y_{i,t}^*$ is drawn is drawn according to $\sigma|(h_e^t, y_t^*)$. Furthermore, $E_{\sigma^*}\left[\bar{U}_i\left(h_0^{t+1,*}\right)|h_d^{t,*}, w_{i,t}^*, m_{i,t}^*, y_{i,t}^*\right] = E_\sigma\left[\bar{U}_i(h_0^{t+1})|h_d^t\right]$ because h_d^t and $h_d^{t,*}$ induce identical continuation games in periods $t+1$ onwards, and all other variables do not affect the continuation game. Therefore, agent i is willing to pay so long as

$$-(1-\delta)\tau_{i,t}^* \leq \delta E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right)\middle| h_d^t, \xi_{i,t}, y_{i,t}^*\right] - E_{\sigma^*}\left[\bar{U}_i\left(h_0^{t+1}\right)\middle| h_d^t\right].$$

$\tau_{i,t}^*$ depends only on variables that agent i has observed or can perfectly infer from $m_{i,t}^*$. Plugging in $\tau_{i,t}^*$, agent i is willing to pay so long as

$$-\left(B_i\left(\xi_{i,t}, y_{i,t}^*|h_d^t\right) - \delta E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right)\middle| h_d^t, \xi_{i,t}, y_{i,t}^*\right]\right) \leq \delta E_\sigma\left[S_i\left(\sigma, h_0^{t+1}\right)\middle| h_d^t, \xi_{i,t}, y_{i,t}^*\right] - E_\sigma\left[\bar{U}_i\left(h_0^{t+1}\right)\middle| h_d^t\right]$$

or $B_i\left(\xi_{i,t}, y^*_{i,t}|h^t_d\right) \geq E_\sigma\left[\bar{U}_i\left(h^{t+1}_0\right)\middle| h^t_d\right]$. This inequality holds 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 an RE with the desired properties.

Proof of Proposition 1

7.0.1 Proof of Statement 2

Definition 1. Define the transformation

$$G_i\left(y_i|\,\theta, d_i, \tilde{d}_i, e_i, \tilde{e}_i\right) = F_i^{-1}\left(F_i\left(y_i|\,\theta, \tilde{e}_i, \tilde{d}_i\right)\middle|\,\theta, e_i, d_i\right).$$

The distribution over outcomes y_i has full support, so F_i is strictly increasing and hence F_i^{-1} is a function. F_i^{-1} is continuously differentiable, because F_i is continuously differentiable.

Claim 1. The distribution over $G_i\left(y_i|\,\theta, d_i, \tilde{d}_i, e_i, \tilde{e}_i\right)$ induced by $\left(\theta, \tilde{d}_i, \tilde{e}_i\right)$ is identical to the distribution over y_i induced by (θ, d_i, e_i): $G_i(y_i)|\,\theta, \tilde{d}_i, \tilde{e}_i \overset{d}{=} y_i|\,\theta, d_i, e_i$.

Proof of Claim 1. To prove the claim, it suffices to show that for every y_i,

$$F\left(y_i|\,\theta, \tilde{d}_i, \tilde{e}_i\right) = F_i\left(G_i\left(y_i|\,\theta, d_i, \tilde{d}_i, e_i, \tilde{e}_i\right)\middle|\,\theta, d_i, e_i\right).$$

This is true by definition of G_i.■

Definition 2. Fix a distribution ψ over i-dyad surplus. Define

$$\bar{e}_i(\theta, d_i, \psi) = \underset{e_i}{\operatorname{argmax}}\, E[y_i|\,\theta, d_i, e_i] - c(e_i)$$

subject to: there exists a mapping $S_i : [0,\infty) \to \mathbb{R}$ and a reward scheme $B_i : [0,\infty) \to \mathbb{R}$ satisfying

1. Effort IC: $\bar{e}_i \in \operatorname{argmax}_{e_i}\{E[B_i(y_i)|\,\theta, d_i, e_i] - c(e_i)\}$
2. Dynamic enforcement: for all $y_i \in [0,\infty)$, $\delta\bar{U}_i(\theta) \leq B_i(y_i) \leq \delta S_i(y_i)$
3. Distribution-matching: $S_i|\,\theta, d_i, e_i \overset{d}{=} \psi$.

Definition 3. For monotonically increasing $S_i : [0,\infty) \to \mathbb{R}$, define $\hat{e}_i\left(\theta, d_i, \tilde{d}_i, e_i\middle|\,S_i\right)$ implicitly by

$$\begin{aligned} c'(\hat{e}_i) \;=\; & \int_0^{y_i^*(\theta,d_i,e_i)} \bar{U}_i(\theta)\frac{\partial f_i}{\partial e_i}\left(y_i|\,\theta, \tilde{d}_i, \hat{e}_i\right)dy_i \\ & + \int_{y_i^*(\theta,d_i,e_i)}^{\infty} S_i\left(G_i\left(y_i|\,\theta, d_i, \tilde{d}_i, e_i, \hat{e}_i\right)\right)\frac{\partial f_i}{\partial e_i}\left(y_i|\,\theta, \tilde{d}_i, \hat{e}_i\right)dy_i. \end{aligned} \tag{1}$$

Claim 1. Suppose $\left(\theta, d_i, \tilde{d}_i, e_i\right)$ are such that $d_i = \tilde{d}_i$ and $\hat{e}_i\left(\theta, d_i, d_i, e_i \middle| S_i\right) = e_i$. Then $\hat{e}_i$ is differentiable on a neighborhood about that point.

Proof of Claim 1. Take $S_i : [0, \infty) \rightarrow \mathbb{R}$ to be a monotonically increasing function. The equation (1) may be rewritten as $H = 0$, where

$$\begin{aligned} H \equiv & \int_0^{y_i^*(\theta,d_i,e_i)} \bar{U}_i(\theta) \frac{\partial f_i}{\partial e_i}\left(y_i \middle| \theta, \tilde{d}_i, \hat{e}_i\right) dy_i \\ & + \int_{G_i\left(y_i^*(\theta,d_i,e_i)\right)}^{\infty} S_i(y_i) \frac{\partial G_i^{-1}}{\partial y_i} \frac{\partial f_i}{\partial e_i}\left(G_i^{-1}(y_i) \middle| \theta, \tilde{d}_i, \hat{e}_i\right) dy_i - c'(\hat{e}_i). \end{aligned}$$

This expression is continuously differentiable in both $\tilde{d}_i$ and $\hat{e}_i$. Therefore, by the Implicit Function Theorem, $\frac{\partial \hat{e}_i}{\partial \tilde{d}_i}$ exists on a neighborhood about $\left(\theta, d_i, \tilde{d}_i, e_i\right)$ as long as $\frac{\partial H}{\partial \hat{e}_i} \neq 0$.

To show that this is the case, we will bound H from above by a function $\bar{H}$ that is differentiable in $\hat{e}_i$ and strictly decreasing in $\hat{e}_i$ on a neighborhood about $\left(\theta, d_i, \tilde{d}_i, e_i\right)$, and coincides with H at $\hat{e}_i = e_i$. For $\varepsilon > 0$, let

$$\begin{aligned} \bar{H} \equiv & \int_0^{y_i^*(\theta,d_i,e_i)} \bar{U}_i(\theta) \frac{\partial f_i}{\partial e_i}\left(y_i \middle| \theta, d_i, \hat{e}_i\right) dy_i \\ & + \int_{y_i^*(\theta,d_i,e_i)}^{y_i^*(\theta,d_i,e_i)+\varepsilon} S_i\left(G_i\left(y_i \middle| \theta, d_i, d_i, e_i, \hat{e}_i\right)\right) \frac{\partial f_i}{\partial e_i}\left(y_i \middle| \theta, d_i, \hat{e}_i\right) dy_i \\ & + \int_{y_i^*(\theta,d_i,e_i)+\varepsilon}^{\infty} S_i(y_i) \frac{\partial f_i}{\partial e_i}\left(y_i \middle| \theta, d_i, \hat{e}_i\right) - c'(\hat{e}_i). \end{aligned}$$

At $\hat{e}_i = e_i$, $G_i(y_i) = y_i$, so $\bar{H} = H$. For $\hat{e}_i > e_i$ sufficiently close, we claim that $\bar{H} \geq H$. Note that $G_i\left(y_i \middle| \theta, d_i, d_i, e_i, \hat{e}_i\right) \leq y_i$ if $\hat{e}_i \geq e_i$, because F_i is FOSD increasing in e_i. Therefore, $S_i\left(G_i\left(y_i \middle| \theta, d_i, d_i, e_i, \hat{e}_i\right)\right) \leq S_i(y_i)$, because S_i is monotonically increasing. Further, for $\hat{e}_i$ sufficiently close to e_i, $\frac{\partial f_i}{\partial e_i}\left(y_i \middle| \theta, d_i, \hat{e}_i\right) \geq 0$ for $y_i \geq y_i^*(\theta, d_i, e_i) + \varepsilon$, because $\frac{\partial f_i}{\partial e_i}$ is strictly monotonically increasing in y_i and equals 0 at $y_i^*(\theta, d_i, e_i)$ for decision d_i and effort e_i. This proves that $\bar{H} \geq H$.

If $\varepsilon = 0$, then $\frac{\partial \bar{H}}{\partial \hat{e}_i} < 0$ by CDFC. $\frac{\partial \bar{H}}{\partial \hat{e}_i}$ is continuous in ε, because G_i is continuous in y_i, so $\frac{\partial \bar{H}}{\partial \hat{e}_i} < 0$ for ε sufficiently close to 0. So $\bar{H}$ is such that $\bar{H} = H$ at $e_i = \hat{e}_i$, $\bar{H} \geq H$ for $\hat{e}_i > e_i$ sufficiently close, and $\frac{\partial \bar{H}}{\partial \hat{e}_i} < 0$. We conclude that $\frac{\partial H}{\partial \hat{e}_i} < 0$.■

Claim 2. Let σ^* be a surplus-maximizing equilibrium. Then for any on-path history $h_d^t \in \mathcal{H}_d^t$, and agent i, let $\psi_i^*\left(\cdot \middle| h_d^t\right)$ be the distribution $S_i\left(h_0^{t+1}\right) \middle| h_d^t$ induced by σ^*. Then $e_{i,t}^* = \bar{e}_i\left(\theta_t, d_{i,t}^*, \psi_i^*\left(\cdot \middle| h_d^t\right)\right)$.

Proof of Claim 2. Suppose $e_{i,t}^* > \bar{e}_i\left(\theta_t, d_{i,t}^*, \psi_i^*\left(\cdot \middle| h_d^t\right)\right)$. It is easy to see that (IC) is a necessary condition for a credible reward scheme to exist. So $e_{i,t}^*$ satisfies (2) and (3) by

Lemma 1 and induces a distribution over continuation dyad-surplus equal to $\psi_i^*(\cdot|\, h_d^t)$. So it must be that $e_{i,t}^* > e_i^{FB}(\theta_t, d_{i,t})$ by definition of $\bar{e}_i$. By Definition 2, there exists some mapping $\bar{G}(y_{i,t})$ such that the distribution $S_i\left(\bar{G}(y_{i,t})\right)\big|\, h_d^t, \bar{e}_{i,t}\left(\theta_t, d_{i,t}^*, \psi_i^*(\cdot|\, h_d^t)\right)$ equals $\psi_i^*(\cdot|\, h_d^t)$.

Define $\tilde{\sigma}$ as a strategy that is identical to σ^* except following history h_d^t. At history h_d^t, agents $j \neq i$ play as in σ^*. Agent i chooses effort $\bar{e}_i\left(\theta_t, d_{i,t}^*, \psi_i^*(\cdot|\, h_d^t)\right)$. Following the realization of output y_t, agent i's output $y_{i,t}$ is treated as output $\bar{G}(y_{i,t})$, but otherwise continuation play is identical to σ^*. Following h_0^{t+1}, $\tilde{\sigma}$ has a credible reward scheme, because σ^* does. In period t, there exists a credible reward scheme that induces agent i to choose $\bar{e}_i\left(\theta_t, d_{i,t}^*, \psi_i^*(\cdot|\, h_d^t)\right)$ by Definition 2. Agents $j \neq i$ face marginal distributions over continuation payoffs that are identical to σ^*, so they are willing to choose $e_{j,t}^*$ as in σ^*. At histories $h_0^{t'}$ for $t' < t$, $S_j\left(\tilde{\sigma}, h_0^{t'}\right) = S_j\left(\sigma^*, h_0^{t'}\right)$ for $j \neq i$, while $S_i\left(\tilde{\sigma}, h_0^{t'}\right) \geq S_i\left(\sigma^*, h_0^{t'}\right)$, because $\bar{e}_i\left(\theta_t, d_{i,t}^*, \psi_i^*(\cdot|\, h_d^t)\right)$ leads to strictly higher surplus at history h_d^t than $e_{i,t}^*$. So $\tilde{\sigma}$ has a credible reward scheme and so is payoff-equivalent to an equilibrium, which contradicts σ^* being surplus-maximizing.

If $e_{i,t}^* < \bar{e}_i\left(\theta_t, d_{i,t}^*, \psi_i^*(\cdot|\, h_d^t)\right)$, then $e_{i,t}^* < e_i^{FB}(\theta_t, d_{i,t})$ as well. Then the alternative equilibrium that is identical to σ^* except that agent i chooses $\bar{e}_i\left(\theta_t, d_{i,t}^*, \psi_i^*(\cdot|\, h_d^t)\right)$ leads to a strictly higher total surplus, which is a contradiction. Therefore $e_{i,t}^* = \bar{e}_i\left(\theta_t, d_{i,t}^*, \psi_i^*(\cdot|\, h_d^t)\right)$.■

Claim 3. $\bar{e}_i(\theta_t, d_{i,t}, \psi_i^*(\cdot|\, h_d^t))$ is weakly increasing in $d_{i,t}$.

Proof of Claim 3. Fix d_i^* and $d_i > d_i^*$. By assumption, for any $d_i \geq d_i^*$, there exists a conditional distribution $R_i\left(x_i|\, \theta_t, d_{i,t}^*, d_{i,t}, y_i\right)$ with density r_i such that

$$\int_0^y \int_0^\infty r_i\left(x_i|\, \theta_t, d_{i,t}^*, d_{i,t}, y_i\right) f_i(y_i|\, \theta, d_i, e_i)\, dx_i dy_i = \int_0^y f_i(y_i|\, \theta, d_i^*, e_i)\, dy_i.$$

Suppose e_i satisfies the conditions of Definition 2 under $(\theta, d_i^*, \psi_i^*)$. Then there exists some $B_i, S_i : [0,\infty) \to \mathbb{R}$ that satisfy the conditions of Definition 2 and induce effort e_i. For $d_i \geq d_i^*$, consider implementing the scheme $\tilde{B}_i, \tilde{S}_i$ derived by first transforming the realized output y_i using the conditional distribution $r_i\left(x_i|\, \theta_t, d_{i,t}^*, d_{i,t}, y_{i,t}\right)$, then applying $\tilde{B}_i, \tilde{S}_i$ to the transformed output. If the agent exerts effort e_i, then this results in an identical distribution over continuation play ψ_i^*. Furthermore, e_i satisfies (IC) and $\tilde{B}_i$ satisfies (DE), because both of these expressions were satisfied under the original B_i, S_i. So for any $e_i \leq \bar{e}_i(\theta, d_i^*, \psi_i^*(\cdot|\, h_d^t))$, $e_i \leq \bar{e}_i(\theta, d_i, \psi_i^*(\cdot|\, h_d^t))$ as well. ■

Claim 4. Let σ^* be a surplus-maximizing equilibrium with on-path history h_d^t. Define $S_i(y_{i,t}) = E_{\sigma^*}\left[S_i\left(\sigma^*, h_0^{t+1}\right)\big|\, h_d^t, y_{i,t}\right]$. Without loss of generality, $S_i(y_{i,t})$ is monotonically increasing. Moreover, if $\hat{e}_i\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\big|\, S_i\right) \leq e_i^{FB}\left(\theta_t, d_{i,t}^*\right)$, then $e_{i,t}^* = \hat{e}_i\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\big|\, S_i\right)$.

Proof of Claim 4. We first argue that $S_i(y_{i,t})$ is monotonically increasing. Suppose not. Then there exists $y_i < \tilde{y}_i$ such that $S_i(y_i) > S_i(\tilde{y}_i)$. Consider the following alternative: with probability $\varepsilon > 0$, outcome $\tilde{y}_i$ is treated as outcome y_i. With probability $\frac{f_i\left(\tilde{y}_i|\theta_t,d_{i,t},e^*_{i,t}\right)}{f_i\left(y_i|\theta_t,d_{i,t},e^*_{i,t}\right)}\varepsilon$, outcome y_i is treated as outcome $\tilde{y}_i$. Agents $j \neq i$ face identical distributions over continuation play and so are willing to exert the same effort in every period. Agent i's incentive constraint can be written

$$c'(e_i) \leq \int_0^{y_i^*(\theta_t,d_{i,t},e_{i,t})} \bar{U}_i(\theta_t)\frac{\partial f_i}{\partial e_i}dy_i + \int_{y_i^*(\theta_t,d_{i,t},e_{i,t})}^{\infty} S_i(y_i)\frac{\partial f_i/\partial e_i}{f_i} f_i\left(y_i|\,\theta_t,d_{i,t},e_t\right)dy_i.$$

Thus, it suffices to show that

$$\varepsilon\left[S_i(y_i)-S_i(\tilde{y}_i)\right]\frac{(\partial f_i/\partial e_i)(\tilde{y}_i)}{f_i(\tilde{y}_i)}f_i(\tilde{y}_i)+\varepsilon\left[\frac{f_i(\tilde{y}_i)}{f_i(y_i)}S_i(\tilde{y}_i)-\frac{f_i(\tilde{y}_i)}{f_i(y_i)}S_i(y_i)\right]\frac{(\partial f_i/\partial e_i)(y_i)}{f_i(y_i)}f_i(y_i)\geq 0$$

or

$$\varepsilon\left[S_i(y_i)-S_i(\tilde{y}_i)\right]\left[\frac{(\partial f_i/\partial e_i)(\tilde{y}_i)}{f_i(\tilde{y}_i)}-\frac{(\partial f_i/\partial e_i)(y_i)}{f_i(y_i)}\right]\geq 0.$$

By assumption, $S_i(y_i) - S_i(\tilde{y}_i) > 0$, so the first term is positive. Similarly, $\frac{(\partial f_i/\partial e_i)(\tilde{y}_i)}{f_i(\tilde{y}_i)} > \frac{(\partial f_i/\partial e_i)(y_i)}{f_i(y_i)}$, because $\tilde{y}_i > y_i$ and f_i satisfies strictly MLRP. Thus, this inequality holds strictly, and so the proposed perturbation strictly relaxes agent i's (IC) constraint without affecting the (IC) constraints for agents $j \neq i$. So we can assume $S_i(y_i)$ is monotonically increasing without loss.

We already know that $e^*_{i,t} = \bar{e}_i$, and it is easy to show that $\bar{e}_i \geq \hat{e}_i\left(\theta_t, d^*_{i,t}, d^*_{i,t}, e^*_{i,t}\right)$ if $\hat{e}_i \leq e_i^{FB}$. Therefore, it suffices to show that $\bar{e}_i = \hat{e}_i\left(\theta_t, d^*_{i,t}, d^*_{i,t}, e^*_{i,t}\right)$. We know that $\bar{e}_i$ satisfies the first-order condition

$$c'(\bar{e}_i) = \int_0^{\infty} B_i(y_i)\frac{\partial f_i}{\partial e_i}\left(y_i|\,\theta_t, d^*_{i,t}, \bar{e}_i\right).$$

Since B_i must satisfy the (DE) constraint, this first-order condition implies that

$$c'(\bar{e}_i) \leq \int_0^{y_i^*\left(\theta_t,d_i^*,\bar{e}_i\right)} \bar{U}_i(\theta_t)\frac{\partial f_i}{\partial e_i}\left(y_i|\,\theta_t,d^*_{i,t},\bar{e}_i\right) + \int_{y_i^*\left(\theta_t,d^*_{i,t},\bar{e}_i\right)}^{\infty} S_i(y_i)\frac{\partial f_i}{\partial e_i}\left(y_i|\,\theta_t,d^*_{i,t},\bar{e}_i\right).$$

If this expression holds with equality, then $\bar{e}_i = \hat{e}_i\left(\theta_t, d^*_{i,t}, d^*_{i,t}, \bar{e}_i\right)$ and we are done. Suppose that this expression does not hold with equality. Then either $\hat{e}_i\left(\theta_t, d^*_{i,t}, d^*_{i,t}, \bar{e}_i\right) > \bar{e}_i$ or $\hat{e}_i\left(\theta_t, d^*_{i,t}, d^*_{i,t}, \bar{e}_i\right)$ does not exist. The former contradicts $\bar{e}_i \geq \hat{e}_i\left(\theta_t, d^*_{i,t}, d^*_{i,t}, \bar{e}_i\right)$, while the latter implies that (IC) is always satisfied and so $\hat{e}_i\left(\theta_t, d^*_{i,t}, d^*_{i,t}, e^*_{i,t}\right) \leq e_i^{FB}\left(\theta_t, d^*_{i,t}\right)$ cannot hold. ■

Claim 5. Define

$$s_i\left(\theta_t, d_{i,t}, \tilde{d}_{i,t}, e_{i,t}\right) = E\left[y_i|\,\theta_t, \tilde{d}_{i,t}, \hat{e}_{i,t}\left(\theta_t, d_{i,t}, \tilde{d}_{i,t}, e_{i,t}\right)\right] - c\left(\hat{e}_{i,t}\left(\theta_t, d_{i,t}, \tilde{d}_{i,t}, e_{i,t}\right)\right)$$

and

$$s_i^*\left(\theta, \tilde{d}_{i,t}, e_{i,t}^*\right) = E\left[y_i|\,\theta_t, \tilde{d}_{i,t}, e_{i,t}^*\right] - c\left(e_{i,t}^*\right).$$

In any sequentially surplus-maximizing equilibrium σ^*, let $h_0^t \in \mathcal{H}_0^t$ be an on-path history. For two agents i, j, if $\hat{e}_i\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\right) = e_{i,t}^*$ and $\hat{e}_j\left(\theta_t, d_{j,t}^*, d_{j,t}^*, e_{j,t}^*\right) = e_{j,t}^*$, then

$$\frac{\partial s_i}{\partial \tilde{d}_i}\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\right) = \frac{\partial s_j}{\partial \tilde{d}_j}\left(\theta_t, d_{j,t}^*, d_{j,t}^*, e_{j,t}^*\right).$$

If $\hat{e}_i\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\right) = e_{i,t}^*$ but $\hat{e}_j\left(\theta_t, d_{j,t}^*, d_{j,t}^*, e_{j,t}^*\right) \neq e_{j,t}^*$, then

$$\frac{\partial s_j^*}{\partial \tilde{d}_j} = \frac{\partial s_i}{\partial \tilde{d}_i}.$$

If $\hat{e}_i\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\right) \neq e_{i,t}^*$ and $\hat{e}_j\left(\theta_t, d_{j,t}^*, d_{j,t}^*, e_{j,t}^*\right) \neq e_{j,t}^*$, then

$$\frac{\partial s_i^*}{\partial \tilde{d}_i} = \frac{\partial s_j^*}{\partial \tilde{d}_j}.$$

Proof of Claim 5. We consider each case separately. Let $\hat{e}_i = e_{i,t}^*$ and $\hat{e}_j = e_{j,t}^*$. Suppose towards contradiction that

$$\frac{\partial s_i}{\partial \tilde{d}_i}\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\right) > \frac{\partial s_j}{\partial \tilde{d}_j}\left(\theta_t, d_{j,t}^*, d_{j,t}^*, e_{j,t}^*\right).$$

Because σ^* is sequentially surplus-maximizing, $\sigma^*|\,h_0^t$ must be surplus-maximizing. Consider the following perturbed continuation equilibrium starting in period t (with perturbations denoted by $\tilde{}$): $\tilde{d}_{i,t} = d_{i,t} + \varepsilon$, $\tilde{d}_{j,t} = d_{j,t} - \varepsilon$, $\tilde{e}_{i,t} = \hat{e}_{i,t}\left(\theta_t, d_{i,t}, \tilde{d}_{i,t}, e_{i,t}^*\right)$ and $\tilde{e}_{j,t} = e_{j,t}\left(\theta_t, d_{j,t}^*, \tilde{d}_{j,t}, e_{j,t}^*\right)$. For all agents $k \notin \{i, j\}$, $d_{k,t}$ and $e_{k,t}$ remain the same. Continuation play is as in σ^*, except that $y_{i,t}$ is first transformed by $G_i\left(\cdot|\,\theta, d_{i,t}^*, \tilde{d}_{i,t}, e_{i,t}^*, \hat{e}_{i,t}\left(\theta_t, d_{i,t}^*, \tilde{d}_{i,t}, e_{i,t}^*\right)\right)$ and likewise for $y_{j,t}$. For all $k \notin \{i, j\}$, this reward scheme is identical to the reward scheme in σ^*. For i,

$$B_i\left(y_{i,t}\right) = \begin{cases} \bar{U}_i\left(\theta_t\right) & y_{i,t} \leq y_i^*\left(\theta_t, d_{i,t}^*, e_{i,t}^*\right) \\ S_i\left(G_i\left(y_{i,t}\right)\right) & y_{i,t} > y_i^*\left(\theta_t, d_{i,t}^*, e_{i,t}^*\right) \end{cases}$$

and similarly for agent j. Such a perturbation is feasible, because $d_{i,t}$ and $d_{j,t}$ are both interior.

For agents $k \notin \{i, j\}$, the (IC) constraint is

$$c'\left(e_{k,t}^*\right) = \int_0^\infty \cdots \int_0^\infty B_k\left(y_{k,t}, G_i\left(y_{i,t}\right), G_j\left(y_{j,t}\right), y_{-\{i,j,k\},t}\right) \frac{\partial f_k}{\partial e_k}\left(y_{k,t} \middle| \theta_t, d_{k,t}^*, e_{k,t}^*\right) dy_{k,t} \prod_{\ell \neq k} f_\ell\left(y_{\ell,t} \middle| \theta_t, \tilde{d}_{\ell,t}, \tilde{e}_{\ell,t}\right.$$

Fix $y_{-i,t}$. Then changing variables to $G_i\left(y_i\right)$ yields

$$\begin{aligned} & \int_0^\infty B_k\left(y_{k,t}, G_i\left(y_{i,t}\right), G_j\left(y_{j,t}\right), y_{-\{i,j,k\},t}\right) f_i\left(y_{i,t} \middle| \theta_t, \tilde{d}_{i,t}, \tilde{e}_{i,t}\right) dy_{i,t} \\ = & \int_0^\infty B_i\left(y_{k,t}, y_{i,t}, G_j\left(y_{j,t}\right), y_{-\{i,j,k\},t}\right) f_i\left(y_{i,t} \middle| \theta_t, d_{i,t}, e_{i,t}\right) dy_{i,t} \end{aligned}$$

since $f_i\left(G_i^{-1}\left(y_{i,t}\right) \middle| \theta_t, \tilde{d}_{i,t}, \tilde{e}_{i,t}\right) \frac{\partial G_i^{-1}}{\partial y_i} = f_i\left(y_i \middle| \theta_t, d_{i,t}^*, e_{i,t}^*\right)$ by definition of G_i. A similar calculation holds for $y_{j,t}$. For all $\ell \notin \{i, j, k\}$, $\left(\tilde{d}_{\ell,t}, \tilde{e}_{\ell,t}\right) = \left(d_{\ell,t}^*, e_{\ell,t}^*\right)$. Therefore, agent k's (IC) constraint may be written

$$c'\left(e_{k,t}^*\right) = \int_0^\infty \cdots \int_0^\infty B_k\left(y_t\right) \frac{\partial f_k}{\partial e_k}\left(y_{k,t} \middle| \theta_t, d_{k,t}^*, e_{k,t}^*\right) dy_{k,t} \prod_{\ell \neq k} f_\ell\left(y_{\ell,t} \middle| \theta_t, d_{\ell,t}^*, e_{\ell,t}^*\right) dy_{\ell,t}.$$

But this (IC) constraint is identical to the (IC) constraint under σ^* and so must hold. The dynamic enforcement constraint must hold for $k \notin \{i, j\}$ for similar reasons.

Consider agent i. Agent i is willing to choose effort $\hat{e}_{i,t}\left(\theta_t, d_{i,t}, \tilde{d}_{i,t}, e_{i,t}^*\right)$, because this effort satisfies IC-FOC by construction. Similarly for agent j. The dynamic enforcement constraint holds by construction of the reward scheme and continuation play for agents i and j.

This perturbed equilibrium generates identical surplus in periods $t+1$ onwards. Thus, $\sigma^* \middle| h_0^t$ is surplus-maximizing only if

$$s_i\left(\theta_t, d_{i,t}^*, d_{i,t}^* + \varepsilon, e_{i,t}^*\right) + s_j\left(\theta_t, d_{j,t}^*, d_{j,t}^* - \varepsilon, e_{j,t}^*\right) \leq s_i\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\right) + s_j\left(\theta_t, d_{j,t}^*, d_{j,t}^*, e_{j,t}^*\right).$$

The left-hand side of this expression is the total surplus produced by agents i and j in period t of the perturbed equilibrium. By Claim 4, the right-hand side is the total surplus produced by agents i and j in period t of σ^*. Rearranging, dividing by $\varepsilon > 0$, and taking the limit as $\varepsilon \to 0$ yields:

$$\lim_{\varepsilon \to 0} \frac{s_i\left(\theta_t, d_{i,t}^*, d_{i,t}^* + \varepsilon, e_{i,t}^*\right) - s_i\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\right)}{\varepsilon} \leq \lim_{\varepsilon \to 0} \frac{s_j\left(\theta_t, d_{j,t}^*, d_{j,t}^*, e_{j,t}^*\right) - s_j\left(\theta_t, d_{j,t}^*, d_{j,t}^* - \varepsilon, e_{j,t}^*\right)}{\varepsilon}$$

or

$$\frac{\partial s_i}{\partial \tilde{d}_i}\left(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*\right) \leq \frac{\partial s_j}{\partial \tilde{d}_j}\left(\theta_t, d_{j,t}^*, d_{j,t}^*, e_{j,t}\right),$$

which is a contradiction.

Suppose $\hat{e}_i = e^*_{i,t}$ but $\hat{e}_j \neq e^*_{j,t}$. Then $\hat{e}_j > e^*_{j,t}$, since otherwise $e^*_{j,t}$ could not satisfy IC-FOC for any reward scheme that satisfies that (DE) constraint. In particular,

$$\begin{aligned} c'\left(e^*_{j,t}\right) &< \int_0^{y^*_j\left(\theta_t,d^*_{j,t},e^*_{j,t}\right)} \bar{U}_i\left(\theta_t\right)\frac{\partial f}{\partial e_j}\left(y_{j,t}\middle|\,\theta_t,d^*_{j,t},e^*_{j,t}\right)dy_j \\ &\quad + \int_{y^*_j\left(\theta_t,d^*_{j,t},e^*_{j,t}\right)}^{\infty} S_i\left(y_{i,t}\right)\frac{\partial f}{\partial e_j}\left(y_{j,t}\middle|\,\theta_t,d^*_{j,t},e^*_{j,t}\right)dy_j. \end{aligned}$$

For $\varepsilon > 0$ sufficiently small, $e^*_{j,t}$ continues to satisfy IC-FOC with the same reward scheme. So consider an identical perturbation to the previous case, except that agent j continues to choose $e^*_{j,t}$. Derivations similar to the previous case prove that

$$\frac{\partial s^*_j}{\partial \tilde{d}_j} = \frac{\partial s_i}{\partial \tilde{d}_i}$$

as desired.

If $\hat{e}_i > e^{FB}_i$ and $\hat{e}_j > e^{FB}_j$, then IC-FOC binds for neither agent. So a similar perturbation leads to the desired result.■

Claim 6. Let σ^* be a surplus-maximizing equilibrium. At any history h^{t+1}_0 such that for some agent i, (i) $e^*_{i,t} < e^{FB}_i$, (ii) $y_{i,t} \geq y^*_{i,t}$, and (iii) there exists $j \neq i$ such that $y_{j,t'} \leq y^*_{j,t'}$ for all $t' \leq t$, $\sigma^*|\,h^{t+1}_0$ is not surplus-maximizing.

Proof of Claim 6. Suppose towards contradiction that $\sigma^*|\,h^{t+1}_0$ is surplus-maximizing. By Claim 5,

$$\frac{\partial s_i}{\partial \tilde{d}_i}\left(\theta_t,d^*_{i,t},d^*_{i,t},e^*_{i,t}\right) = \frac{\partial s_j}{\partial \tilde{d}_j}\left(\theta_t,d^*_{j,t},d^*_{j,t},e^*_{j,t}\right).$$

Consider the following perturbation in σ^*: in all periods $t' < e$, $\tilde{\sigma}$ is as in σ^*. In period t, $\tilde{e}_{i,t} = e^*_{i,t} + \eta$ for some $\eta > 0$, while $\tilde{e}_{k,t} = e^*_{k,t}$ for all $k \neq i$. At the end of period t, output for each agent is transformed using $G_i\left(y_{i,t}\middle|\,\theta_t,d^*_{i,t},d^*_{i,t},e^*_{i,t},\tilde{e}_{i,t}\right)$, and these transformed outputs are used to calculate the "effective history" h^{t+1}_0. In period $t+1$, $\tilde{d}_{i,t+1} = d^*_{i,t+1} + \varepsilon$ and $\tilde{d}_{j,t+1} = d^*_{j,t+1} - \varepsilon$ at the "effective history" h^{t+1}_0 with $y_{i,t} \geq y^*_{i,t}\left(\theta_t,\tilde{d}_{i,t}\right)$ and $y_{j,t} < y^*_{j,t}\left(\theta_t,\tilde{d}_{j,t}\right)$. Agents choose efforts $\tilde{e}_{i,t+1} = \bar{e}_i\left(\theta_{t+1},\tilde{d}_{i,t+1},\psi^*_i\left(\cdot\middle|\,h^{t+1}_0\right)\right)$, $\tilde{e}_{j,t+1} = \bar{e}_j\left(\theta_{t+1},\tilde{d}_{j,t+1},\psi^*_j\left(\cdot\middle|\,h^{t+1}_0\right)\right)$, and $\tilde{e}_{k,t+1} = e^*_{k,t+1}$ for all $k \notin \{i,j\}$.

Given on-path history h^{t+1}_d, let $S_i\left(y_i\right) = E_{\sigma^*}\left[S_i\left(\sigma^*,h^{t+2}_0\right)\middle|\,h^{t+1}_d,y_i\right]$. Then there exist functions $\bar{G}_i$ and $\bar{G}_j$ such that the distribution $S_i\left(\bar{G}_i\left(y_i\right)\right)\middle|\,\theta_{t+1},\tilde{d}_{t+1},\bar{e}_i$ equals $\psi^*_i\left(\cdot\middle|\,h^{t+1}_0\right)$ and similarly for agent j. In period $t+2$ onwards, treat the realized outputs for agents i and j as $\bar{G}_i\left(y_i\right)$ and $\bar{G}_j\left(y_j\right)$, but otherwise treat continuation play as in σ^*.

We claim that this perturbed strategy is an equilibrium, and that if $\varepsilon > 0$ is sufficiently small, this perturbed strategy generates strictly higher total surplus. First consider continuation play in periods $t+2$ onward. BFE are recursive, and $\tilde{S}$ was an equilibrium dyad-surplus vector under the original equilibrium σ^*. So continuation play forms an equilibrium.

In period $t+1$, agents i and j are willing to choose $\bar{e}_i, \bar{e}_j$ respectively, by definition. Agents $k \notin \{i,j\}$ are willing to choose $e^*_{k,t+1}$, because the distribution over reward schemes and continuation play are identical to σ^*. Therefore, the change in total surplus in period $t+1$ from this perturbation equals

$$\begin{aligned}&\left(E\left[y_{i,t+1}|\,\theta_{t+1}, \tilde{d}_{i,t+1}, \bar{e}_i\right] - c\left(\bar{e}_i\right) + E\left[y_{j,t+1}|\,\theta_{t+1}, \tilde{d}_{j,t+1}, \bar{e}_j\right] - c\left(\bar{e}_j\right)\right)\\ &-\left(E\left[y_{i,t+1}|\,\theta_{t+1}, d^*_{i,t+1}, e^*_{i,t+1}\right] - c\left(e^*_{i,t+1}\right) + E\left[y_{j,t+1}|\,\theta_{t+1}, d^*_{j,t+1}, e^*_{j,t+1}\right] - c\left(e^*_{j,t+1}\right)\right).\end{aligned}$$

Suppose $\hat{e}_i\left(\theta_{t+1}, d^*_{i,t+1}, d^*_{i,t+1}, e^*_{i,t+1}\right) = e^*_{i,t+1}$ and similarly for agent j. Then

$$E\left[y_{i,t+1}|\,\theta_{t+1}, \tilde{d}_{i,t+1}, \bar{e}_i\right] - c\left(\bar{e}_i\right) \geq E\left[y_{i,t+1}|\,\theta_{t+1}, \tilde{d}_{i,t+1}, \hat{e}_i\right] - c\left(\hat{e}_i\right)$$

by definition of $\bar{e}_i$, and similarly for agent j. By claim 4, $e^*_{i,t+1} = \hat{e}_i\left(\theta_{t+1}, d^*_{i,t+1}, d^*_{i,t+1}, e^*_{i,t+1}\right)$ and $e^*_{j,t+1} = \hat{e}_j\left(\theta_{t+1}, d^*_{j,t+1}, d^*_{j,t+1}, e^*_{j,t+1}\right)$. Therefore, we can bound the change in total surplus in period $t+1$ from below by

$$\begin{aligned}K\left(\varepsilon\right) \;=\;& s_i\left(\theta_{t+1}, d^*_{i,t+1}, \tilde{d}_{i,t+1}, e^*_{i,t+1}\right) - s_i\left(\theta_{t+1}, d^*_{i,t+1}, d^*_{i,t+1}, e^*_{i,t+1}\right)\\ &+ s_j\left(\theta_{t+1}, d^*_{j,t+1}, \tilde{d}_{j,t+1}, e^*_{j,t+1}\right) - s_j\left(\theta_{t+1}, d^*_{j,t+1}, d^*_{j,t+1}, e^*_{j,t+1}\right).\end{aligned}$$

This is the "cost" of these distorted policies. Note that $\lim_{\varepsilon\to 0} K\left(\varepsilon\right)/\varepsilon = 0$ by Claim 5.

If $\hat{e}_i\left(\theta_{t+1}, d^*_{i,t+1}, d^*_{i,t+1}, e^*_{i,t+1}\right) \neq e^*_{i,t+1}$, then

$$E\left[y_{i,t+1}|\,\theta_{t+1}, \tilde{d}_{i,t+1}, \bar{e}_i\right] - c\left(\bar{e}_i\right) \geq E\left[y_{i,t+1}|\,\theta_{t+1}, \tilde{d}_{i,t+1}, e^*_{i,t+1}\right] - c\left(e^*_{i,t+1}\right)$$

as shown in the proof of Claim 5. So $\lim_{\varepsilon\to 0} K\left(\varepsilon\right)/\varepsilon = 0$ by Claim 5.

Now consider period t. Because $y^*_{j,t'} \leq y^*_j\left(\theta_{t'}, d_{j,t'}, e_{j,t'}\right)$ for all $t' \leq t$, it is without loss to assume that $B_j\left(y^*_{j,t'}\right) < S_j\left(h_0^{t'+1}\right)$ in every previous period. Suppose not for some $t' < t$; then we could replace the reward scheme with $\tilde{B}_j$ such that $\tilde{B}_j\left(y^*_{j,t'}\right) < S_j\left(h_0^{t'+1}\right)$ for all $y_{j,t'} < y^*_j\left(\theta_{t'}, d_{j,t'}, e_{j,t'}\right)$ such that IC-FOC holds with equality. Therefore, the perturbation of $S_j\left(h_0^{t+1}\right)$ described above does not affect agent j's effort in period t. Similarly, the distribution over $S_k\left(h_0^{t+1}\right)$ is identical for all $k \notin \{i,j\}$, so agent k is willing to choose the same effort in period t and all preceding periods. It remains to consider whether agent i is willing to choose effort $\tilde{e}_{i,t}$.

Because $e^*_{i,t} < e^{FB}_i\left(\theta_t, d^*_{i,t}\right)$, we know that $e^*_{i,t} = \hat{e}_i\left(\theta_t, d^*_{i,t}, d^*_{i,t}, e^*_{i,t}\right)$ by Claim 4. Therefore,

$$c'\left(e^*_{i,t}\right) = \int_0^{y^*_i\left(\theta_t, d^*_{i,t}, e^*_{i,t}\right)} \bar{U}_i\left(\theta_t\right) \frac{\partial f_i}{\partial e_i}\left(y_i \middle| \theta_t, d^*_{i,t}, e^*_{i,t}\right) dy_i + \int_{y^*_i\left(\theta_t, d^*_{i,t}, e^*_{i,t}\right)}^{\infty} S_i\left(y_i\right) \frac{\partial f_i}{\partial e_i}\left(y_i \middle| \theta_t, d^*_{i,t}, e^*_{i,t}\right) dy_i.$$

By Claim 3, $\bar{e}_i \geq e^*_i$ in period $t+1$. Furthermore, by assumption

$$\frac{\partial E\left[y_{i,t+1} \middle| \theta_{t+1}, d_{i,t+1}, e_{i,t+1}\right]}{\partial d_{i,t+1}} > 0.$$

Therefore, $S_i\left(y_i\right)$ strictly increases for every $y_i \geq y^*_i\left(\theta_t, d^*_{i,t}, e^*_{i,t}\right)$. Let $\tilde{S}_i\left(y_i\right)$ be the new dyad surpluses. Then there exist $\zeta > 0$ and a measurable function $\rho : [y^*_i\left(\theta_t, d^*_{i,t}, e^*_{i,t}\right), \infty) \to \{0, 1\}$ such that

$$\int_{y^*_i\left(\theta_t, d^*_{i,t}, e^*_{i,t}\right)}^{\infty} \rho\left(y_i\right) \frac{\partial f_i}{\partial e_i} dy_i > 0$$

and such that for all y_i such that $\rho\left(y_i\right) = 1$,

$$\tilde{S}_i\left(y_i\right) - S_i\left(y_i\right) \geq \zeta \rho\left(y_i\right).$$

Consider the function

$$\begin{aligned} \tilde{H}\left(\tilde{e}_i, \zeta\right) &= -c\left(\tilde{e}_i\right) + \int_0^{y^*_i\left(\theta_t, d^*_{i,t}, e^*_{i,t}\right)} \bar{U}_i\left(\theta_t\right) \frac{\partial f_i}{\partial e_i}\left(y_i \middle| \theta_t, d^*_{i,t}, \tilde{e}_i\right) dy_i \\ &+ \delta \int_{y^*_i\left(\theta_t, d^*_{i,t}, e^*_{i,t}\right)}^{\infty} \Big(S_i\left(G_i\left(y_i \middle| \theta_t, d^*_{i,t}, d^*_{i,t}, e^*_{i,t}, \tilde{e}_i\right)\right) + \zeta \rho\left(G_i\left(y_i \middle| \theta_t, d^*_{i,t}, d^*_{i,t}, e^*_{i,t}, \tilde{e}_i\right)\right) \frac{\partial f_i}{\partial e_i}\left(y_i \middle| \theta_t, d^*_{i,\right. \end{aligned}$$

Effort $e^*_{i,t}$ satisfies the original IC-FOC constraint with equality, so

$$\tilde{H}\left(e^*_{i,t}, 0\right) = 0.$$

We claim that $\left.\frac{\partial \tilde{e}_{i,t}}{\partial \zeta}\right|_{\left(e^*_{i,t}, 0\right)} > 0$. An argument similar to the one used in Claim 1 shows that $\frac{\partial \tilde{H}}{\partial \tilde{e}_i} < 0$. So this partial derivative exists. It is strictly positive, because $\frac{\partial \tilde{H}}{\partial \zeta} > 0$.

Define

$$\tilde{s}_i\left(\tilde{e}_{i,t}\left(\zeta\right)\right) = E\left[y_i \middle| \theta_t, d^*_{i,t}, \tilde{e}_{i,t}\left(\zeta\right)\right] - c\left(\tilde{e}_{i,t}\left(\zeta\right)\right).$$

Note that ζ depends on the difference between $\tilde{S}_i\left(y_i\right)$ and $S_i\left(y_i\right)$ and is therefore a function of ε. Note also that $\tilde{s}_i$ is differentiable in $\tilde{e}_i$ and $\frac{\partial \tilde{s}_i}{\partial \tilde{e}_i} > 0$ at $\tilde{e}_i = e^*_{i,t}$, because $e^*_{i,t} < e^{FB}_i\left(\theta_t, d^*_{i,t}\right)$. Furthermore,

$$\zeta\left(\varepsilon\right) \rho\left(y_{i,t}\right) \geq E\left[y_{i,t+1} \middle| \theta_{t+1}, d^*_{i,t+1} + \varepsilon, e^*_{i,t+1}\right] - E\left[y_{i,t+1} \middle| \theta_{t+1}, d^*_{i,t+1}, e^*_{i,t+1}\right]$$

because effort $e_{i,t}^*$ is weakly increasing in ε. Define $\underline{\zeta}(\varepsilon)$ to solve

$$\underline{\zeta}(\varepsilon)\rho(y_{i,t}) = E\left[y_{i,t+1}|\,\theta_{t+1}, d_{i,t+1}^* + \varepsilon, e_{i,t+1}^*\right] - E\left[y_{i,t+1}|\,\theta_{t+1}, d_{i,t+1}^*, e_{i,t+1}^*\right].$$

Then $\underline{\zeta}(\varepsilon)$ is differentiable and satisfies $d\underline{\zeta}/d\varepsilon > 0$. So the difference in period-t surplus may be bounded from may be written

$$\mathcal{B}(\varepsilon) = \tilde{s}_i\left(\tilde{e}_{i,t}\left(\underline{\zeta}(\varepsilon)\right)\right) - \tilde{s}_i\left(e_{i,t}^*\right)$$

where $\tilde{s}_i = E\left[y_i|\,e_i\right] - c(e_i)$. Since only agent i's effort changes in period t.

This perturbation leads to strictly higher total surplus if

$$\mathcal{B}(\varepsilon) + K(\varepsilon) > 0.$$

Dividing by ε and taking the limit as $\varepsilon \to 0$ implies that this perturbation generates strictly higher surplus if

$$\lim_{\varepsilon\to 0}\frac{\mathcal{B}(\varepsilon)}{\varepsilon} > -\lim_{\varepsilon\to 0}\frac{K(\varepsilon)}{\varepsilon}.$$

Recall that $\lim_{\varepsilon\to 0}\tilde{s}_i\left(\tilde{e}_{i,t}\left(\underline{\zeta}(\varepsilon)\right)\right) = \tilde{s}_i\left(e_{i,t}^*\right)$. So

$$\lim_{\varepsilon\to 0}\frac{\mathcal{B}(\varepsilon)}{\varepsilon} = \frac{d\tilde{s}_i}{d\tilde{e}_i}\frac{d\tilde{e}_i}{d\underline{\zeta}}\frac{d\underline{\zeta}}{d\varepsilon} > 0,$$

because each of these derivatives is strictly positive. Similarly,

$$\lim_{\varepsilon\to 0}\frac{K(\varepsilon)}{\varepsilon} = \frac{\partial s_i}{\partial\tilde{d}_{i,t+1}}\left(\theta_{t+1}, d_{i,t+1}^*, d_{i,t+1}^*, e_{i,t+1}^*\right) - \frac{\partial s_j}{\partial\tilde{d}_{j,t+1}}\left(\theta_{t+1}, d_{j,t+1}^*, d_{j,t+1}^*, e_{j,t+1}^*\right) = 0.$$

We conclude that this perturbed equilibrium generates strictly higher surplus than σ^*, so σ^* is not surplus-maximizing. ■

7.0.2 Proof of Statement 1

If $\sum_{i=1}^N d_{i,t} < 1$ in some period t, consider the alternative decision $\tilde{d}_t$ with $\sum_{i=1}^N \tilde{d}_{i,t} = 1$ and $\tilde{d}_{i,t} \geq d_{i,t}$ for all i, and each agent i chooses effort $\bar{e}_i(\theta, d_i, \psi_i^*(\cdot|h_d^t)$. It is straightforward to show that this alternative is a continuation equilibrium that generates strictly higher surplus total surplus **and** i-dyad surplus for all agents $i \in \{1, .., N\}$. Therefore, it also weakly relaxes all dynamic-enforcement constraints in previous periods. So we have constructed an equilibrium that generates more total ex ante surplus than σ^*. Contradiction. ■

Proof of Proposition 2

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

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

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

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

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

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

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

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

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

while 1-dyad surplus is bounded above by

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

This bound on total surplus may be rewritten

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

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

Note that

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

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

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

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

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

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

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

or

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

Now, define

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

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

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

Proof of Proposition 3

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

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

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

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

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

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

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

This inequality holds if ξ is not too large.

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

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

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

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

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

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

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

Proof of Proposition 4

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 $b_i : \phi_0 \left(\mathcal{H}_y^t \right) \rightarrow \mathbb{R}$ satisfying

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

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

   $$\delta \sum_{i \in I} E_{\sigma^*} \left[\bar{U}_i \left(h_0^{t+1} \right) \middle| h_y^t \right] \leq \sum_{i \in I} b_i \left(\phi_0 \left(h_y^t \right) \right) \leq \delta \sum_{i \in I} E_{\sigma^*} \left[\sum_{i \in I} U_i \left(\sigma^*, h_0^{t+1} \right) + \Pi \left(\sigma^*, h_0^{t+1} \right) \middle| h_y^t \right]. \tag{10}$$

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

   $$\begin{aligned} U_i(\sigma^*, h_d^t) &\geq \bar{U}_i(h_d^t) & (11) \\ \Pi \left(\sigma^*, h_d^t \right) &\geq \sum_{i \in I} \left(E_{\sigma^*} \left[b_i \left(\phi_0 \left(h_y^t \right) \right) - (1-\delta) c_{i,t} \middle| h_d^t \right] - U_i \left(\sigma^*, h_d^t \right) \right) \end{aligned}$$

2. For strategy σ, suppose there exists $b_i : \phi_0\left(\mathcal{H}_y^t\right) \to \mathbb{R}$ satisfying (2), (10), and (11). Then there exists a BFE σ^* that induces the same distribution over $\left\{D_t, d_t, e_t, y_t, \{u_i\}_{i=0}^N\right\}_{t=0}^\infty$ as σ.

Proof of Lemma A.1

1: Suppose σ^* is a PDE. 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 by

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

Then agent i chooses $e_{i,t}$ to solve (2). As in Lemma 1, $b_i(\phi_0(h_y^t)) \geq E[\bar{U}_i(h_0^{t+1})|h_y^t]$. Suppose there exists a set $I \subset \{1,\ldots,N\}$ such that

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

Then the principal may profitably deviate by choosing $\tau_{i,t} = 0$ for all $i \in I$, earning no less than 0 in the continuation game. Together, these arguments imply (10).

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

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

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

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

implying the first line of (11).

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

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

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

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

If this expression holds for the crucial set of agents I, then a fortiori it holds for any other set of agents, implying the second line of (10).

2 : Define $\zeta\left(h^{t}\right)=\left\{D_{t'},d_{t'},e_{t'},y_{t'}\right\}_{t'=0}^{t}$. 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 $\sigma|\left\{h^t,D_t\right\}$.

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

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

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

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

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

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

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

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

We claim 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

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

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

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

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 (11). Similarly, the principal has no profitable deviation: setting $I = \emptyset$ in (2) implies $\Pi\left(\sigma, h_d^t\right) \geq 0$.

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

Agent i chooses effort to maximize

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

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

Finally, consider deviations in $\{\tau_{i,t}\}_{i=1}^N$. If $\tau_{i,t} < 0$, agent i has no profitable deviation by the first inequality in (10). Let $J = \{i \in \{1, \ldots, N\}|\, \tau_{i,t} \leq 0\}$. The principal has no profitable deviations as long as

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

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 (10) implies that the principal has no profitable deviation.

Completing Proof of Proposition 4

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 $\{D_t, d_t, e_t, y_t\}_{t=0}^{\infty}$ as σ^*, but with $U_i(\tilde{\sigma}, h^t) = \bar{U}_i(h^t)$ 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

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

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), (10), and (11) 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 $\{D_t, d_t, e_t, y_t,\}_{t=0}^{\infty}$ 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 (10) and (11). Since the principal's continuation payoff is higher under $\bar{\sigma}$ than under $\tilde{\sigma}^*$, we conclude that $\bar{\sigma}$ satisfies Lemma A.1. So σ^* cannot be surplus-maximizing, which is a contradiction.

Proof of Proposition 5

We begin by finding **necessary** conditions for equilibrium. As before, define $b_i(h_y^t) = E_\sigma[(1-\delta)\tau_{i,t} + \delta U_i(h_0^{t+1}) | h_y^t]$ as agent i's reward. By an argument similar to Lemma 1, agent i can earn no less than 0 in the continuation game, so $b_i(h_y^t) \geq 0$. Consider a deviation in the principal's relationship with agent i. The principal is min-maxed if agent i chooses his outside option. With probability ϵ, this choice is not publicly observed, in which case the principal loses no more than $\Pi^i \equiv \sum_{t'=1}^{\infty} \delta^{t'}(1-\delta)(y_{i,t+t'} - w_{i,t+t'} - \tau_{i,t+t'})$. With probability $1-\epsilon$, this choice is publicly observed, in which case the principal loses no more than $\delta \sum_{j \neq 1}^{N} S_i + \delta \Pi^i$ because $\Pi^j \leq S_j$. Therefore, the principal is willing to pay $\tau_{i,t}$ only if

$$\tau_{i,t} \leq \frac{\delta}{1-\delta}\left((1-\epsilon)\sum_{j\neq i} S_i + \Pi^i\right).$$

Plugging this expression into the definition of b_i yields that $b_i(h_y^t) \leq \delta\left((1-\epsilon)\sum_{j\neq i} S_j + S_i\right)$ Finally, the principal must not be able to profitably deviate by refusing to pay all agents, a condition given by $\sum_{j=1}^{N} b_j(h_y^t) \leq \delta \sum_{j=1}^{N} S_j$.

Suppose these three sets of conditions are the only constraints on the equilibrium. Define $\tilde{C}^{R1} = R - c$, $\tilde{C}^{R2} = (2-\epsilon)(\alpha R - c)$, $\tilde{C}^{W1} = (1-\delta)(W-c) + \delta(\rho \tilde{C}^{R1} + (1-\rho)\tilde{C}^{W1})$, and $\tilde{C}^{W2} = (1-\delta)(W-c) + \delta(\rho C^{R2} + (1-\rho)C^{W2})$. We make the following equilibrium assumptions:

$$\begin{aligned}
c &\leq \frac{\delta}{1-\delta}(\alpha R - c) \\
c &> \frac{\delta}{1-\delta}\tilde{C}^{W2} \\
c &\leq \frac{\delta}{1-\delta}\tilde{C}^{W1} \\
2(\alpha R - c) &> R - c > (2-\epsilon)(\alpha R - c) > W - c > 2(\alpha W - c) \\
(1-\delta)(\alpha R - W) &> \delta\rho(1-\epsilon)(\alpha R - c)
\end{aligned}$$

As long as $\epsilon > 0$, it is easy to show that these constraints are simultaneously satisfied by an open, non-empty set of parameters. Define $\gamma \in (0,1)$ as the solution to $c = \frac{\delta}{1-\delta}(\gamma \tilde{C}^{W1} + (1-\gamma)\tilde{C}^{W2})$.

Note that for appropriately-chosen parameters, we can satisfy the constraints and make γ arbitrarily close to 0. If these constraints are satisfied, any sequentially efficient equilibrium must hire one agent whenever $\theta_t = W$ and two agents whenever $\theta_t = R$. However, under this hiring scheme, the agent hired when $\theta_t = W$ cannot be motivated to work hard.

Now, consider the following strategy. If $\theta_t = W$, then $d_t = 1$, $w_{i,t} = 0$, $e_{1,t} = 1$, $e_{2,t} = 0$, $\tau_{1,t} = c$, and $\tau_{2,t} = 0$. If $\theta_t = R$ for the first time in period t, then the same actions as above are played with probability γ. With probability $1-\gamma$, $d_t = 2$, $w_{i,t} = 0$, $e_{i,t} = 1$, and $\tau_{i,t} = c$. In subsequent periods, players choose the same actions as in the first period such that $\theta_t = R$. Following any deviation, the agents who observe that deviation reject production and choose low effort in every period, and the principal maximizes surplus given those strategies.

This strategy is an equilibrium. The principal's decision d_t depends only on publicly observed variables (the public randomization device and θ_t), so any deviation from this decision results in payoff 0. So the principal is willing to follow d_t. If $\theta_t = W$, then the principal is willing to pay $\tau_{1,t} = c$ because if she does not, she earns payoff 0 with probability $1-\epsilon$ and otherwise earns $\frac{\delta\rho}{1-\delta+\delta\rho}(1-\gamma)(\alpha R - c)$ continuation surplus. By choice of γ,

$$c = \frac{\delta}{1-\delta}(\gamma \tilde{C}^{W1} + (1-\gamma)\tilde{C}^{W2})$$

so the principal would rather pay $\tau_{1,t}$ than face punishment. Agent 1 is indifferent between rejecting production and accepting and working hard, so is willing to follow the equilibrium. Similar arguments show that there are no profitable deviations once $\theta_t = R$ since $c \leq \frac{\delta}{1-\delta}(\alpha R - c)$ by assumption.

If γ is sufficiently close to 0, then this equilibrium strictly dominates any sequentially surplus-maximizing equilibrium because agent 1 works hard while $\theta_t = R$. So no surplus-maximizing equilibrium is sequentially surplus-maximizing, as desired.

7.1 Proof of Lemma 2

We begin by proving the appropriate extension of Lemma 1 to PBE.

Definition: PBE-Credible Recall $\xi_{i,t} = (m_{i,t}, w_{i,t})$. A reward scheme $B_i : \Xi \times \mathbb{R} \times \phi_i(\mathcal{H}_d^t) \rightarrow \mathbb{R}$ is **PBE-credible in** σ if:

1. For each $h_d^t, \xi_{i,t}$, and $C_{i,t}$ on the equilibrium path,

$$C_{i,t} \in \arg\max_{C_i | d_i, \theta} E_\sigma \left[B_i(\xi_{i,t}, y_{i,t} | \phi_i(h_d^t)) | \phi_i(h_d^t), \xi_{i,t}, C_i \right] - (1-\delta) C_i. \quad (12)$$

2. For each on-path h_y^t,

$$\delta E_\sigma \left[\bar{U}_i(h_0^{t+1}) | \phi_i(h_d^t) \right] \leq B_i(\xi_{i,t}, y_{i,t} | \phi_i(h_d^t)) \leq \delta E_\sigma \left[S_i(\sigma, h_0^{t+1}) | \phi_i(h_d^t), \xi_{i,t}, y_{i,t} \right]. \quad (13)$$

Lemma A.2: Extension of Lemma 1 to PBE

1. If σ^* is a PBE in which no player conditions on past effort choices. Then for each agent i there exists a PBE-credible reward scheme for σ^*.

2. Suppose σ is a strategy with a PBE-credible reward scheme B_i for each agent i. Then there exists a PBE σ^* that induces the same joint distribution over states of the world, decisions, efforts, and outcomes as σ.

Proof of Lemma A.2 This argument is adapted from Andrews and Barron (2014).

Statement 1: Suppose σ^* is a PBE and define B_i by

$$B_i(\xi_{i,t}, y_{i,t}|\phi_i(h_d^t)) = E_{\sigma^*}\left[(1-\delta)\tau_{i,t} + \delta U_i|\phi_i(h_d^t), \xi_{i,t}, y_{i,t}\right].$$

Then B_i must satisfy (12) and the left-hand inequality of (13) by an argument identical to Lemma 1. The principal knows the true history apart from effort, which is irrelevant for her payoff by assumption. So (9) must hold at each history. But then it holds *a fortiori* in expectation, implying the right-hand side of (13)

Statement 2: We construct a PBE σ^* from the strategy σ. Consider a construction identical to the one used in the proof of Lemma 1, statement 2, except that

$$w_{i,t}^* = E_\sigma\left[y_{i,t} - \frac{1}{1-\delta}\left(B_i(\xi_{i,t}, y_{i,t}|\phi_i(h_d^t)) - \delta S_i(\sigma, h_0^{t+1})\right)|\phi_i(h_d^t), \xi_{i,t}\right],$$

$$m_{i,t}^* = \left\{\phi_i(h_0^{t,*}), a_{i,t}, e_{i,t}, \left\{B_i(\xi_{i,t}, y_{i,t}|\phi_i(h_d^t)) - \delta E_\sigma\left[S_i(\sigma, h_0^{t+1})|\phi_i(h_d^t), \xi_{i,t}, y_{i,t}\right]\right\}_{y_{i,t}\in\mathbb{R}}\right\},$$

and the transfer following output y_t^* equals

$$(1-\delta)\tau_{i,t}^* = B_i(\xi_{i,t}, y_{i,t}^*|\phi_i(h_d^t)) - \delta E_\sigma\left[S_i(\sigma, h_0^{t+1})|\phi_i(h_d^t), \xi_{i,t}, y_{i,t}^*\right].$$

By construction, σ^* implements the same joint distribution over states of the world, decisions, efforts, and outcomes as σ. It remains to show that σ^* is a PBE. As in the proof of Lemma 1, the principal earns 0 from each agent i at each history h_0^t on and off the equilibrium path. So as in Lemma 1, the principal is willing to follow the equilibrium.

Consider deviations by agent i. It is straightforward to show that σ^* induces a coarser information partition for agent i than σ: if h_0^t,$h_0^{t,*}$ and $\tilde{h}_0^t, \tilde{h}_0^{t,*}$ are two pairs of histories from the construction of σ^*, then $\phi_i(h_0^{t,*}) = \phi_i(\tilde{h}_0^{t,*})$ whenever $\phi_i(h_0^t) = \phi_i(\tilde{h}_0^t)$. Following any deviation in period t, agent i earns payoff $\bar{U}_i(h_0^{t+1})$. Agent i cannot profitably deviate from $w_{i,t}^* \geq 0$. Agent i is willing to choose $e_{i,t}$ if

$$e_{i,t} \in \arg\max_{e_i} E_{\sigma^*}\left[(1-\delta)\tau_{i,t}^* + \delta U_i|\phi_i(h_d^{t,*}), e_i\right] - (1-\delta)c(e_i).$$

Now, $S_i(\sigma, h_0^{t+1}) = S_i(\sigma^*, h_0^{t+1,*}) = U_i(\sigma^*, h_0^{t+1,*})$, where the first equality is from the definition of h_0^{t+1} and $h_0^{t+1,*}$ and the second equality follows from the fact that the principal earns 0 at each history. Further, because $\phi_i(h_0^{t,*}) = \phi_i(\tilde{h}_0^{t,*})$ whenever $\phi_i(h_0^t) = \phi_i(\tilde{h}_0^t)$ and

$E_\sigma\left[S_i(\sigma, h_0^{t+1})|\phi_i(h_d^t)\right]$ is the same at all histories in $\phi_i(h_d^{t,*})$ by construction of σ^*, it can be shown that

$$E_{\sigma^*}\left[E_\sigma\left[S_i(\sigma, h_0^{t+1})|\phi_i(h_d^t)\right]|\phi_i(h_d^{t,*})\right] = E_\sigma\left[S_i(\sigma, h_0^{t+1})|\phi_i(h_d^{t,*})\right].$$

Plugging these expressions into agent i's incentive constraint yields (12), which is satisfied because B_i is credible.

Agent i is willing to pay $\tau_{i,t}^*$ so long as

$$-(1-\delta)\tau_{i,t}^* \leq \delta E_{\sigma^*}\left[U_i - \bar{U}_i|\phi_i(h_0^{t+1,*})\right].$$

Note that $E_{\sigma^*}\left[\bar{U}_i|\phi_i(h_0^{t+1,*})\right] = \bar{U}_i(h_0^{t+1})$ because $\bar{U}_i$ depends only on the public history, which is identical at $h_0^{t+1,*}$ and h_0^{t+1}. Therefore, the agent is willing to pay $\tau_{i,t}^*$ so long as the left-hand side of (13) holds, which is true because B_i is credible. We conclude that no player has a profitable deviation, so σ^* is a PBE. ■

Completing Proof of Lemma 2 ($\rightarrow$) Suppose σ^* attains continuation surplus V from period t. Consider the following alternative $\tilde{\sigma}$:

1. At $t = 0$, the principal chooses h_0^t from the distribution over $\mathcal{H}_0^t$ induced by σ^*.

2. Play continues as in $\sigma^*|h_0^t$, with the following exception: the message $m_{i,t}$ in period $t = 0$ includes $\phi_i(h_0^t)$.

By construction, players have the same beliefs in $\tilde{\sigma}$ as in $\sigma^*|h_0^t$. Therefore, $\tilde{\sigma}$ is an equilibrium. In expectation, $\tilde{\sigma}$ generates total surplus V because h_0^t is drawn according to the distribution induced by σ^*.

($\leftarrow$) Suppose σ^* leads to net ex ante total surplus V. Construct $\tilde{\sigma}$ in the following way: for all periods $t' < t$, players follow the static equilibrium ($a_{i,t'} = w_{i,t'} = \tau_{i,t'} = 0$, $d_{t'}$ chosen to maximize surplus in period t' given these actions). From period t onwards, players follow σ^*. Then $\tilde{\sigma}$ is clearly an equilibrium that attains continuation surplus V in period t. ■

7.2 Proof of Proposition 6

Let σ^* be a surplus-maximizing PBE. Then we claim that there exists a $d^* : \Theta \rightarrow D$ such that $d_0 = d^*(\theta_0)$ with probability 1. For each $i \in \{1, .., N\}$, define

$$\tilde{S}_i(x|h_e^0) = E_{\sigma^*}\left[S_i(\sigma^*, h_0^1)|\theta_0, d_0, e_0, y_0 = x + \gamma(\theta_0, d_0)\right].$$

Let $d^*(\theta)$ be the unique vector such that $\sum_{i=1}^N d_i^*(\theta) = 1$ and $\forall i, j \in \{1, ..., N\}$,

$$\frac{\partial \gamma_i}{\partial d_i}(\theta, d_i^*(\theta)) = \frac{\partial \gamma_j}{\partial d_j}(\theta, d_j^*(\theta)). \tag{14}$$

The vector $d^*(\theta)$ exists and is unique because $\lim_{d_i \to 0} \frac{\partial \gamma_i}{\partial d_i} = \infty$ and γ_i is strictly concave.

Suppose $d_0 \neq d^*(\theta_0)$ with positive probability. Then there exists agents i, j such that $\frac{\partial \gamma_i}{\partial d_i}(\theta, d_{0,i}) > \frac{\partial \gamma_j}{\partial d_j}(\theta, d_{0,j})$. Then it must be that $d_{0,j} > 0$ and $d_{0,i} < 1$. Consider an alternative strategy that is identical to σ^* except for $t = 0$. In $t = 0$: (i) $\tilde{d}_{0,i} = d_{0,i} + \epsilon$, $\tilde{d}_{0,j} = d_{0,j} - \epsilon$, and $\tilde{d}_{0,k} = d_{0,k}$, (ii) agents choose e_0 as in σ^*, and (iii) if output y_0 is realized, continuation play is chosen as if the output realized was $\tilde{y}_0 = y_0 - \gamma(\theta_0, \tilde{d}_0) + \gamma(\theta_0, d_0)$.

It is straightforward to show that this alternative strategy is a PBE. It generates strictly higher total surplus in period $t = 0$ because σ^* does not satisfy (14). Following period 0, it generates the same surplus as σ^*. So σ^* cannot be a surplus-maximizing equilibrium; contradiction.

Now, suppose that σ^* is a sequentially surplus-maximizing PBE. Then in every period t, $d_t = d^*(\theta_t)$ with probability 1 by the previous argument; otherwise, we could construct an equilibrium that is surplus-maximizing but with $d_0 \neq d^*(\theta_0)$. Further, we claim that there exists a function $e^* : \Theta \to [0, \infty)$ such that $e_t = e^*(\theta)$ with probability 1 in each period of σ^*. Define $\bar{x}_i(e_i)$ as the unique x_i such that $\frac{\partial \tilde{p}_i}{\partial e_i}(x_i, e_i) = 0$. Given that d_t depends only on θ_t in each period, each agent faces a stationary environment. The techniques of Levin (2003) can be extended to this setting, which proves that $e_t = e^*(\theta_t)$ in each period t.

In any sequentially surplus-maximizing PBE, $d_t = d^*(\theta_t)$ and $e_t = e^*(\theta_t)$. These actions are independnet of history, so agent beliefs about the true history are irrelevant for satisfying (12) and (13). Therefore, any sequentially surplus-maximizing PBE σ^* is also a Recursive Equilibrium. At almost every on-path history h_0^t, $\sigma^*|h_0^t$ is a PBE that maximizes continuiation surplus. So σ^* is a sequentially surplus-maximizing Recursive Equilibrium.

Suppose that surplus-maximizing Recursive Equilibria are not sequentially surplus-maximizing. Then no sequentially surplus-maximizing RE exists, so by the argument above no sequentially surplus-maximizing PBE exists. This proves the claim. ■

8 References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Li, J., Matouschek, N., and Powell, M., "Power Dynamics in Organizations," *Working Paper* (2015).

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)