Policies in Relational Contracts

Daniel Barron and Michael Powell*

March 22, 2016

Abstract

We consider how a firm's policies constrain its relational contracts. A policy is a sequence of decisions made by a principal; each decision determines how agents' efforts affect their outputs. We consider surplus-maximizing policies in a flexible dynamic moral hazard problem between a principal and several agents with unrestricted vertical transfers and no commitment. If each agent observes only his own output and pay, then the principal might optimally implement dynamically inefficient, history-dependent policies to credibly reward high-performing agents. We develop conditions under which such backward-looking policies are surplus-maximizing and then illustrate how they influence hiring, investment, and performance.

*Barron: Northwestern University Kellogg School of Management, Leverone Hall, 2001 Sheridan Road, Evanston IL 60208; email: d-barron@kellogg.northwestern.edu. Powell: Northwestern University Kellogg School of Management, Leverone Hall, 2001 Sheridan Road, Evanston IL 60208; email: mike-powell@kellogg.northwestern.edu. The authors would like to thank Joyee Deb, Willie Fuchs, John Geanakoplos, Bob Gibbons, Yingni Guo, Marina Halac, Jin Li, Jim Malcomson, Niko Matouschek, Luis Rayo, Larry Samuelson, Takuo Sugaya, Joel Watson, and Alexander Wolitzky for very helpful comments and discussion. Thanks also to participants in many seminars and conferences. Barron gratefully acknowledges support from the Yale University Cowles Foundation while working on this paper.

1 Introduction

Business relationships often rest upon parties' goodwill rather than the contracts they sign—the threat of future punishments can motivate individuals to exert effort and to reward the efforts of their partners. In the canonical relational-incentive contracting models that capture this intuition (Bull (1987); MacLeod and Malcomson (1989); Levin (2003)), the principal's only role is to promise and pay monetary compensation to her agents. She is otherwise passive.

Yet in any real-world enterprise, managers make decisions that affect how a group of individuals contribute to the firm's objectives. Supervisors assign tasks to team members. Supply-chain managers allocate business among suppliers. Executives allocate capital to divisions and promotions to subordinates. 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 sourcing decisions toward those who saw past success (Peter and Hull (1969); Graham et al. (2015); 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 biased decisions can arise in optimally managed relationships among a principal and her agents. While such decisions lead to lower continuation surplus, a principal who promises to bias future decisions towards an agent can credibly promise stronger monetary incentives to that agent today. To make this point, we develop a general framework that builds upon Levin (2003)'s repeated principal-agent model with moral hazard, transferable utility, and risk-neutral parties. We extend Levin's framework to accommodate persistent public states and multiple agents. The key feature of our model is that the principal can make a public **decision** in each period that influences how agents' 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 total continuation surplus. We say that such decisions are **biased**.

We show that backward-looking policies arise naturally if monitoring is bilateral, that is, if each agent cannot observe the principal's interactions with other agents. In this setting, players cannot coordinate punishments or rewards. As a result, the principal can promise larger rewards to an agent who is expected to produce a lot of future surplus. A decision that makes an agent more indispensable to the firm ensures that the principal and that agent have more to lose if they do not uphold their promises to one another. Decisions biased towards 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 an owner of an up-and-coming business who must choose how many people to hire. Achieving early success requires hard work from early employees, and motivating this hard work 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 indispensable in the future. Therefore, to motivate agents to work hard, the owner might adopt a policy of being slow to hire following an increase in demand for the firm's products, which makes existing workers more indispensable to 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.

In our general analysis, we develop a set of dynamic enforcement constraints that are necessary and sufficient for a policy to be implemented in a relational contract. These constraints link a principal's future decisions to the incentives she can credibly promise her agents today: after an agent produces a given output, his continuation payoff can be no less than his outside option and no more than the total value of his future production. His future production depends in turn on the principal's future decisions. Using these dynamic enforcement constraints, we develop sufficient conditions for backward-looking policies to be part of any surplus-maximizing relational contract. We show that decisions are biased toward those agents who have performed well in the past at the expense of those who have not.

Next, we apply our framework to two stylized examples to illustrate how backward-looking policies might manifest. Revisiting the hiring example, we confirm that additional hiring may optimally lag an increase in demand, and we 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 its managers (or employees).

Our game has imperfect private monitoring—agents observe neither 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 a solution concept that restricts attention to recursive relational contracts. This approach provides a tractable way to highlight the forces that lead to backward-looking policies in optimal relational contracts.

We explore the role of private monitoring with three extensions that analyze alternative solution concepts and monitoring structures. First, we demonstrate in the context of a simple class of games that the optimality of backward-looking policies is not driven by our restriction to recursive equilibria; they arise even if we consider the full (non-recursive) set of Perfect Bayesian Equilibria. Second, we show that if the game has public monitoring, then biased decisions never arise in surplus-maximizing equilibria. Unlike the setting with bilateral monitoring, agents can coordinate to jointly punish the principal if she does not uphold her promises if monitoring is public. Biased decisions decrease total continuation surplus, weakening the principal's incentive to uphold her promises, and so have no place in a surplus-maximizing equilibrium. Finally, we argue that the sequential efficiency result for public monitoring is fragile by considering a stylized monitoring assumption that implies that agents probabilistically observe deviations in one another's relationships. Restricting attention to a simple example, we show that backward-looking policies may be optimal so long as the principal's deviation in one relationship is not publicly observed with probability 1.

Literature Review: Andrews and Barron (forthcoming) is the most closely related paper to ours. Their model, which studies how relational contracts influence allocation dynamics in a supply chain, is a special case of our setting. One goal of our paper is to provide a general framework for analyzing policies in relational contracts. Our main objective is to study biased decisions and sequential inefficiencies, which only occur if first-best surplus is unattainable in equilibrium. In contrast, Andrews and Barron (forthcoming) restrict attention to settings in which first-best surplus can be attained.

Our paper builds on the literature on dynamics in optimal formal contracts to consider settings in which the principal cannot commit to payments. The seminal contribution by Fudenberg, Holmstrom, and Milgrom (1990)—henceforth FHM—considers whether a long-term formal contract can be replicated by a sequence of short-term contracts. An important part of FHM's analysis is to develop conditions under which the optimal long-term contract may be backward-looking in the sense that it (potentially inefficiently) conditions actions on payoff-irrelevant past information. For instance, either restrictions on transfers or asymmetric information about payoff-relevant variables may lead to such backward-looking formal contracts.

Many of the seminal papers in the relational contracting literature (Bull (1987); MacLeod and Malcomson (1989); Baker et al. (1994); Levin (2002, 2003)) study models in which the optimal formal contract would not be backward-looking. In these papers, stationary relational contracts—in which continuation play in each period is independent of the history on the equilibrium path—are optimal. Goldlucke and Kranz (2012) generalize the logic behind these papers to prove that if monitoring is public, utility is transferable, and players have symmetric information about the continuation game, then stationary equilibria are optimal.

A recent and growing literature, partially surveyed in Malcomson (2013), analyzes relational contracts if some of the conditions identified by FHM fail to hold. In these settings, optimal relational contracts might condition on payoff-irrelevant information and entail sequential inefficiencies. For instance, Halac (2012), Malcomson (forthcoming), and others study how relational contracts evolve if the players have asymmetric information about future payoff-relevant variables. Fong and Li (2016) show that the principal might inefficiently suspend production to punish poor performance if the agent has limited liability. Li et al. (forthcoming) show that if transfers are limited, then the principal might reward or punish the agent by changing how much control he can exercise in future periods. Board (2011) shows how limited transfers can lead to distorted allocation decisions in a supply chain. More broadly, an optimal equilibrium might entail history-dependent, inefficient on-path continuation play in a repeated game with imperfect public monitoring but no transfers (see, e.g., Green and Porter (1984); Abreu et al. (1990); Fudenberg et al. (1994)).

This paper takes a different approach by considering an environment that satisfies the conditions identified by FHM. Consequently, the optimal formal contract would not be backward-looking. Nevertheless, we give conditions under which an optimal relational contract must entail systematic history-dependent inefficiencies. In principle, each agent in our setting could be efficiently motivated using monetary payoffs alone. However, players must be willing to follow through on these payments in equilibrium. Biased future decisions ensure that players cannot profitably renege on these payments and so backward-looking policies make strong incentives credible in a relational contract.

2 An Example

This section introduces the key ideas of our model in an example.

Consider a principal who repeatedly interacts with two agents in periods $t = 0, 1, \ldots$. Players share a common discount factor δ. In $t = 0$, the principal and each agent make simultaneous non-negative payments to one another. Players have no liquidity constraints; let $w_{i,0} \in \mathbb{R}$ be the net payment 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 output is $y_{i,0} \in \{0, H_i\}$, with $\Pr\{y_{i,0} = H_i\} = pe_{i,0}$ and

$H_1 > H_2 > 0$. After output is realized, the principal again exchanges payments with each agent; the net payment to agent i is $\tau_{i,0} \in \mathbb{R}$.

At the start of the second period ($t = 1$), the principal makes a once-and-for-all decision by picking one of the two agents. She repeatedly plays the stage game with the chosen agent, while the other agent is constrained to choose $e_{i,t} = 0$ in every subsequent period. Let q_i be the probability that agent i is chosen, with $q_1 + q_2 \leq 1$. The principal and agent i respectively earn $(1-\delta)\sum_{i=1}^{2}(y_{i,t} - w_{i,t} - \tau_{i,t})$ and $(1-\delta)(w_{i,t} + \tau_{i,t} - ce_{i,t})$ in period t.

Suppose that 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. We argue that the principal might decide to continue her relationship with agent 2 even though doing so leads to lower continuation surplus in periods $t \geq 1$. Moreover, her decision optimally depends on the realized outputs in period 0.

Assume that δ is such that either agent can be motivated to work hard in periods $t \geq 1$ if the principal chooses him. Then in any equilibrium that maximizes total surplus, the chosen agent will work hard in each period $t \geq 1$ on the equilibrium path. 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 $t \geq 1$ (denoted $U_{i,1}$). So agent i's **reward scheme** following output $y_{i,0}$ is

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

Output is not contractible, so agent i's reward must be credible in equilibrium. Agent i can always earn 0 by choosing $e_{i,t} = \tau_{i,t} = w_{i,t} = 0$ in each period, so $B_i \geq 0$. The principal can similarly "walk away" from her relationship with agent i by refusing to pay that agent. Because this deviation would be observed only by agent i, the principal would deviate rather than pay an agent more than he produces in the future. So $B_i \leq \delta q_i(pH_i - c)$, where the right-hand side of this inequality is the total expected continuation surplus produced by agent i.

Section 4 shows that this **dynamic enforcement constraint**,

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

is the key constraint imposed by the absence of commitment. In particular, we can construct an equilibrium in which the principal implements any policy (q_1, q_2) and the agents earn any $\{B_i\}_i$ that satisfy this constraint.

Total continuation surplus is maximized if the principal chooses agent 1 in $t = 1$, because $H_1 > H_2$. But then the dynamic enforcement constraint implies $B_2 = 0$: agent 2 cannot be rewarded in equilibrium regardless of his output. Agent 2 is therefore unwilling to exert effort, because he is effectively in a one-shot interaction and so will not be rewarded for high output. Consequently, the principal can either maximize total continuation surplus or motivate agent 2 in period 0, but she cannot do both.

Consider an equilibrium in which both agents work hard in $t = 0$. Then agent 2 must be chosen with positive probability after some outputs. Critically, increasing q_2 relaxes the upper bound of agent 2's dynamic enforcement constraint but does not affect the lower bound. Agent 2 should be optimally rewarded with $B_2 > 0$ if and only if he produces high output, so the optimal equilibrium should have $q_2 > 0$ only if $y_{2,0} = H_2$. Such a history-dependent policy ensures that the principal can credibly reward agent 2 at exactly those histories for which agent 2's reward is constrained from above. In short, the principal's optimal policy may entail history-dependent dynamic inefficiencies.

Before we leave this example, it is revealing to compare this setting to a game with **public monitoring.** Suppose that all variables are publicly observed except efforts, which remain private. As before, each agent can earn no less than his outside option: $B_i \geq 0$. However, both agents observe and can jointly punish any deviation, so the principal can credibly promise to reward the agents if the sum of those rewards is no larger than the sum of the total continuation surpluses produced by those agents: $B_1 + B_2 \leq \delta[p(q_1 H_1 + q_2 H_2) - c]$. The right-hand side of this constraint equals total continuation surplus, so the principal can credibly promise larger rewards to both agents if total continuation surplus is higher. Hence, $q_1 = 1$ in any optimal equilibrium, since this policy both maximizes continuation surplus in $t \geq 1$ and relaxes the dynamic enforcement constraint in $t = 0$. In other words, backward-looking policies are not optimal if monitoring is public.

The rest of this paper generalizes the intuition highlighted by this example to show why backward-looking policies might maximize surplus in settings with bilateral monitoring. We consider public monitoring in Section 6.2.

3 The Model

A single principal (player 0, "she") and N agents (players $i \in \{1, .., N\}$, each "he") interact repeatedly. Time is discrete and indexed by $t \in \{0, 1, ...\}$. Players are risk-neutral and share a common discount factor $\delta \in (0, 1)$. In each period, the principal makes a decision, which determines how each agent i's effort maps into a distribution over that agent's output, from a set of feasible decisions. The principal and each agent can pay each other twice in each period: once after the principal makes a decision and once after output is realized. We refer to these transfers as wage and bonus payments, respectively. The principal sends a private message to each agent i along with the wage.

Formally, the stage game has the following timing:

1. State of the world θ_t and feasible decision set D_t are publicly realized according to $F(\cdot|\{\theta_{t'}, D_{t'}, d_{t'}\}_{t'=0}^{t-1})$.

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

3. The principal and each agent i simultaneously make non-negative transfers to each other. Define $w_{i,t} \in \mathbb{R}$ as the net wage paid to agent i. Only the principal and agent i observe $w_{i,t}$.

4. The principal chooses a message $m_{i,t} \in M$ to privately send to each agent i, where M is a large message space.[1]

[1]Formally, M is at least as large as the set of all functions $f : \mathbb{R} \to \mathbb{R}$. In practice, we can typically make do with a much smaller message space.

5. Each agent i chooses to participate ($a_{i,t} = 1$) or not ($a_{i,t} = 0$). If agent i does not participate, he receives $\bar{u}_i(d_t, \theta_t) \geq 0$ and produces output $y_{i,t} = 0$. Only the principal and agent i observe $a_{i,t}$.

6. If $a_{i,t} = 1$, agent i privately chooses effort $e_{i,t}$ from compact set $\mathcal{E}_i \subseteq \mathbb{R}_+$ at cost $c(e_{i,t})$.

7. Each agent i produces output $y_{i,t} \in \mathbb{R}$, with $y_{i,t} \sim P_i(\cdot|\theta_t, d_t, e_{i,t})$ such that $E[y_{i,t}|\theta_t, d_t, e_{i,t}] \geq 0$ for all $(\theta_t, d_t, e_{i,t})$. Denote $y_t = (y_{1,t}, ..., y_{N,t})$. Only the principal and agent i observe $y_{i,t}$.

8. The principal and each agent i simultaneously make non-negative transfers to one another. Define $\tau_{i,t} \in \mathbb{R}$ as the net bonus paid to agent i. Only the principal and agent i observe $\tau_{i,t}$.

We assume that parties have access to a public randomization device after each stage of the game. Agent i's and the principal's stage-game payoffs in each period t are

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

respectively. Each agent i's continuation payoff in period t is

$$U_{i,t} = \sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)u_{i,t'},$$

with an analogous definition for the principal's continuation payoff Π_t. For each agent i and period $t \geq 0$, denote $\xi_{i,t} = (m_{i,t}, w_{i,t}) \in \Xi_i = M_i \times \mathbb{R}$.

Histories and Strategies Let $h_0^t = \{\theta_{t'}, D_{t'}, d_{t'}, w_{t'}, m_{t'}, a_{t'}, e_{t'}, y_{t'}, \tau_{t'}\}_{t'=0}^{t-1}$ be a history at the start of period t, with the set of such histories denoted $\mathcal{H}_0^t$. For any variable x realized during a period, let h_x^t be a within-period history immediately after that variable is realized, so for example $h_y^t = h_0^t \cup \{\theta_t, D_t, d_t, w_t, m_t, a_t, e_t, y_t\}$. Then $\mathcal{H}_x^t$ is the set of such histories, with $\mathcal{H}$ the set of all possible histories. For every agent i, let $\phi_i : \mathcal{H} \to 2^{\mathcal{H}}$ denote agent i's information set, so $\phi_i(h_x^t)$ is the set of histories that i cannot distinguish from h_x^t. Similarly, $\phi_0(h_x^t)$ is the principal's information set. Recall that $\phi_0(h_x^t)$ includes all variables except effort, while $\phi_i(h_x^t)$ includes only $\theta_{t'}, D_{t'}, d_{t'}$, and variables with subscript i. Let $\phi_i(\mathcal{H})$ be the set of player i's information sets.

A **relational contract** is a strategy profile $\sigma = \sigma_0 \times ... \times \sigma_N$, where σ_i maps $\phi_i(\mathcal{H})$ to feasible actions. Continuation play at $\phi_i(h^t)$ is denoted $\sigma_i|\phi_i(h^t)$. A **policy** is a mapping from the principal's information set after observing θ_t and D_t, $\phi_0(\mathcal{H}_D^t)$, to the distribution over decisions taken at that history, $\Delta(D_t)$.

Equilibrium In a game with private monitoring, players form beliefs about the true history given their private information. If players condition continuation play on these beliefs, then information—and hence play—grows increasingly complicated as the game progresses. To avoid these difficulties, many of our results restrict attention to **recursive equilibria (RE)**, which are a recursive and hence relatively tractable refinement of Perfect Bayesian Equilibrium (PBE).[2]

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 PBE of the continuation game.*

In any Perfect Bayesian Equilibrium, player i's actions at history h_x^t must be a best response to the other players' actions, given $\phi_i(h_x^t)$. A recursive equilibrium requires that at the start of each period on the equilibrium path, player i's actions are also a best response given the *true* history h_0^t. This additional restriction applies only at the start of each period: within a period, players best-respond to their Bayesian beliefs as usual. It also 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 is weaker than belief-free

[2] See Watson (2016)'s discussion of *plain PBE* for a rigorous definition in a setting with continuous action spaces.

equilibria (BFE), which is a standard recursive solution concept for repeated games with private monitoring.[3] A relational contract is **self-enforcing** if it is a recursive equilibrium.

Recursive equilibria are recursive on the equilibrium path, which is a non-trivial restriction on equilibrium play. Our main analysis restricts attention to RE, because they lead to clean and intuitive equilibrium constraints. Section 6.1 considers a simple class of games and extends our analysis to PBE, which illustrates that our core intuition is not driven by the restriction to RE.

A self-enforcing relational contract σ^* is **surplus-maximizing** if it yields the largest *ex ante* total expected surplus among recursive equilibria:[4]

$$\sigma^* \in \arg \max_{\sigma | \sigma \text{ is an RE}} E_\sigma \left[\Pi_0 + \sum_{i=1}^{N} U_{i,0} \right].$$

It is **sequentially surplus-maximizing** if, at each on-path $h_0^t \in \mathcal{H}_0^t$, continuation play $\sigma^*|h_0^t$ is surplus-maximizing in the continuation game starting at h_0^t. If $\sigma^*|h_0^t$ is not surplus-maximizing, then we say that the decisions following h_0^t are **biased** and the policy is **backward-looking**.[5] Our analysis gives conditions under which backward-looking policies arise in surplus-maximizing relational contracts.

We assume that $E[y_{i,t}|\theta_t, d_t, e_{i,t}] \geq 0$ for all $(\theta_t, d_t, e_{i,t})$, which implies that the harshest punishment agent i can impose on the principal is to choose $a_{i,t} = 0$ in each period t. Since $\bar{u}_i(\theta_t, d_{i,t}) \geq 0$ for all $(\theta_t, d_{i,t})$, agent i's min-max payoff can be attained if he chooses not to participate in each period, and

[3] See Ely et al. (2005) for 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 PBE, any BFE is also an RE. BFE and RE select the same set of equilibria in a simultaneous-move repeated game with full support over private signals.

[4] If we allow a round of transfers between the principal and each agent before the game begins, then maximizing total surplus is equivalent to maximizing the principal's payoff. It is also equivalent to maximizing the sum of agent payoffs. However, it is *not* equivalent to maximizing an individual agent's payoff.

[5] In our analysis, biased decisions always lead to some agents producing more surplus than they would in a surplus-maximizing continuation equilibrium, while other agents produce too little.

the principal chooses the decisions that minimize his expected outside option. Given a history h_x^t and $i \in \{1, ..., N\}$, we define this **punishment payoff** as

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

Discussion Several features of this model warrant further comment. First, agents do not observe one another's payments, participation decisions, or outputs, and they cannot communicate with one another about these variables. Consequently, the principal can renege on an equilibrium payment to one agent without the other agents ever observing or inferring that deviation. 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, if the principal reneges on a promise to an agent, that agent can punish the principal by taking his outside option. However, the other agents do not follow suit, because they do not observe the deviation. We explore this assumption in Section 6.[6]

Second, the distribution over θ_t and D_t depends only on the public history $\{\theta_{t'}, D_{t'}, d_{t'}\}_{t'=0}^{t-1}$. Consequently, agents have common information about the continuation game at the start of each period, which rules out adverse selection problems. Third, $w_{i,t}$ is paid before each agent i's participation decision ($a_{i,t}$), which simplifies equilibrium punishments by ensuring that agent i can immediately punish a deviation in $w_{i,t}$. We could add transfers after the participation decision but before efforts without changing any of our results. Finally, we allow the principal to send a message to each agent only once per period; allowing the principal to send additional messages at other times would not change any of our results.

[6]If agents could costlessly communicate with one another, then they can use those messages to implement joint punishments. The resulting equilibrium would resemble those in the game with public monitoring studied in Section 6.2.

4 Backward-Looking Policies

This section demonstrates how policies constrain incentives in equilibrium. Section 4.1 develops necessary and sufficient conditions for a relational contract to be self-enforcing. Section 4.2 uses those conditions to show how backward-looking policies are an integral feature of surplus-maximizing relational contracts.

4.1 Necessary and Sufficient Conditions for Equilibrium

Consider a history immediately before agents choose effort. Define agent i's reward scheme B_i at this history as the mapping from each possible output realization to i's expected payoff following that realization. As in Section 2, B_i must satisfy dynamic enforcement constraints in equilibrium: for each output, the upper bound on B_i depends on the principal's equilibrium policy following that output. We introduce the notion of credible reward schemes to formalize these constraints.

Denote the **net cost** of $(a_{i,t}, e_{i,t})$ by $C_{i,t} = a_{i,t}c(e_{i,t}) - (1 - a_{i,t})\bar{u}_i(d_t, \theta_t)$. For each agent i, define i**-dyad surplus** in period t as the total continuation surplus produced by i:

$$S_{i,t} = \sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)(y_{i,t'} - C_{i,t'}). \tag{1}$$

Then total continuation surplus equals $\sum_{i=1}^{N} S_{i,t}$.

Definition 2 *Given relational contract σ, a* ***reward scheme*** $B_i : \mathcal{H}_d^t \times \Xi_i \times \mathbb{R} \to \mathbb{R}$ *is* ***credible in*** σ *if it satisfies:*

1. *IC constraint: for each on-path h_d^t, $\xi_{i,t}$, $a_{i,t}$, and $e_{i,t}$,*

$$(a_{i,t}, e_{i,t}) \in \arg\max_{\tilde{a}_{i,t}, \tilde{e}_{i,t}} E_y\left[B_i(h_d^t, \xi_{i,t}, y_{i,t}) | h_d^t, \xi_{i,t}, \tilde{a}_{i,t}, \tilde{e}_{i,t}\right] - (1-\delta)C_i. \tag{2}$$

2. *Dynamic enforcement constraint: for each on-path h_y^t,*

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

A credible reward scheme satisfies two sets of constraints. First, agent i must be willing to exert effort $e_{i,t}$ if he expects to earn the corresponding reward after each output $y_{i,t}$. This **IC constraint** is given by (2) and implies that B_i 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 than $\delta E_\sigma \left[S_{i,t+1}|h_y^t, \xi_{i,t}, y_{i,t}\right]$ at information set $(h_y^t, \xi_{i,t}, y_{i,t})$, since the principal would prefer to renege than to pay agent i more than his total future production. Agent i's payoff is bounded from below by his min-max payoff, which equals $\bar{U}_i(h_0^{t+1})$. This **dynamic enforcement constraint** (3) must hold with respect to agent i's beliefs about continuation play once he observes his own output. The lower bound of this constraint depends only on the history h_d^t. In contrast, expected i-dyad surplus depends on agent i's future production, which in turn depends on the principal's future decisions. So the continuation policy affects the upper bound of (3).

Recall that the recursive equilibrium solution concept requires players to best-respond given the true history at the start of each period but to form Bayesian expectations given their private histories within each period. Consequently, the expectations in (2) and (3) condition on the full history before period t, plus the variables that agent i observes in period t.

We show that a policy and sequence of effort choices are part of a self-enforcing relational contract if and only if they are supported by a credible reward scheme for each agent i.

Lemma 1 1. *If σ^* is a self-enforcing relational contract, then for all $i \in \{1, ..., N\}$, there exists a reward scheme B_i^* that is credible in σ^*.*

2. *If σ is a strategy with a credible reward scheme B_i for each $i \in \{1, ..., N\}$, then there exists a self-enforcing relational contract σ^* that induces the same joint distribution over states of the world, decisions, efforts, and*

outputs as σ, so $E_\sigma[S_{i,t}] = E_{\sigma^*}[S_{i,t}]$ for each $i \in \{1,..,N\}$ and $t \geq 0$.

Proof: See Appendix A.

To prove part 1 of Lemma 1, consider agent i's moral hazard problem 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 $U_{i,t+1}$ with output $y_{i,t}$. For every $(h_d^t, \xi_{i,t}, y_{i,t})$, define agent i's reward scheme under σ^* as

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

Agent i chooses $e_{i,t}$ in equilibrium if and only if B_i^* satisfies (2). Our goal is to derive the dynamic enforcement constraint (3) that bounds B_i^*.

Following output $y_{i,t}$, agent i would rather renege and be punished than pay more than his entire continuation utility from the relational contract. Therefore, $B_i^* \geq \bar{U}_i$ in any equilibrium. Similarly, the principal can walk away from her relationship with agent i by not paying wages or bonuses to i. Importantly, she can do so without alerting the other agents, who 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. Agent i can therefore earn no more than the total surplus he expects to produce in the future: $B_i^* \leq \delta E_{\sigma^*}[S_{i,t+1} | h_d^t, \xi_{i,t}, y_{i,t}]$. These arguments prove part 1.

The proof of part 2 is more involved. Intuitively, we construct a self-enforcing relational contract using the strategy σ and credible reward scheme B_i. In each period of our construction, the principal chooses the same decision as in σ. She then sends a message to each agent specifying his equilibrium effort and a schedule of output-dependent fines that that agent must pay. Equilibrium wages are such that the principal earns 0 from each agent in every period at the time she chooses d_t. Each agent exerts the effort specified in the message and then pays the fine specified in the message that corresponds to his realized output. A deviation is punished by the breakdown of the corresponding relationship.

Each agent can perfectly infer the principal's stage-game payoff from his wage, his message, and the schedule of fines he pays. Hence, an agent can punish the principal if she would earn a strictly positive payoff in a period. Consequently, the principal earns 0 in each period both on and off the equilibrium path, so she is willing to follow the equilibrium policy. The agent earns his entire i-dyad surplus in each period, but he pays fines following low output. He is willing to exert effort and make the specified payments, because these fines are derived from a credible reward scheme.

4.2 Backward-Looking Policies in Smooth Games

Backward-looking policies can affect equilibrium surplus in three ways. First, backward-looking policies have a **direct cost**, because they lead to lower total continuation surplus. However, such decisions might be biased toward an agent i, in the sense that they lead to a larger $E_{\sigma^*}\left[S_{i,t+1}|h_d^t, \xi_{i,t}, y_{i,t}\right]$ than in any sequentially surplus-maximizing equilibrium for some $y_{i,t}$. Increasing i-dyad surplus relaxes (3) for agent i, leading to an **incentive benefit**: agent i can earn larger rewards in equilibrium following that $y_{i,t}$, which might motivate him to exert more effort. Of course, decisions biased towards agent i are biased away from some agent $j \neq i$. So biased decisions also have an **incentive cost**: biasing decisions away from an agent makes motivating that agent more difficult.

The direct cost of a backward-looking policy depends only on continuation play and so is independent of the payoff-irrelevant history. In contrast, the incentive cost and incentive benefit vary history-by-history, because agent i's dynamic enforcement constraint (3) might bind at some outputs but not others. The upper bound of agent i's dynamic enforcement constraint is likely to bind at a history in which agent i "performs well," that is, $y_{i,t}$ statistically suggests that i exerted effort. At such histories, biasing future decisions towards i has a large incentive benefit, because it relaxes a binding constraint and so facilitates more effort from agent i. Similarly, the upper bound of agent j's constraint is unlikely to bind if he "performs poorly." Tightening j's constraint at such histories has a small incentive cost. A surplus-maximizing relational contract entails biased decisions exactly when the incentive benefits outweigh both the incentive costs and direct costs. Consequently, decisions will tend to be biased towards agents who have performed well in the past—in the sense of producing output that indicates high effort—at the expense of those who have performed poorly.

This intuition is particularly clear in games where equilibrium surplus varies smoothly in decisions and effort. Our main result focuses on these "smooth" games.

Definition 3 *A game is **smooth** if:*

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

2. *Outside options depend only on θ_t. For every $i \in \{1, ..., N\}$, $\mathcal{E}_i$ is an interval and $c_i(\cdot)$ is smooth, strictly increasing, and strictly convex.*

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

4. *Higher d_i lead to weakly more informative P_i: for any θ, $x \in \mathbb{R}$, and $d_i \geq \tilde{d}_i$, there exists a conditional distribution $R_i(\cdot|x) \geq 0$ such that for any e_i, y_i,*

$$p_i(y_i|\theta, \tilde{d}_i, e_i) = \int_{-\infty}^{\infty} R_i(y_i|x) p_i(x|\theta, d_i, e_i) dx. \tag{4}$$

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 determines the distribution of $y_{i,t}$, where a higher weight $d_{i,t}$ leads to both a larger expected $y_{i,t}$ and a weakly more informative distribution in the Blackwell sense. Expected outputs are smooth in all arguments, weakly concave in $e_{i,t}$, and strictly concave in $d_{i,t}$. The distribution over outputs has interval support and satisfies the

Mirrlees-Rogerson conditions, which ensure that we can replace the incentive-compatibility constraint (2) with its first-order condition.[7]

Given these assumptions, first-best effort for agent i is

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

For each (d_i, θ, e_i), there exists a unique $y_i^*(d_i, \theta, e_i) \in \mathbb{R}$ that satisfies

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

Loosely, output $y_i > y_i^*(d_i, \theta, e_i)$ statistically suggests that, conditional on (d_i, θ), agent i chose no less than effort e_i.

Our main result gives conditions under which every surplus-maximizing relational contract in a smooth game entails a backward-looking policy. These conditions are phrased in terms of endogenous objects—decisions, effort, and outputs—to make the intuition clear. After discussing the result, we prove a corollary that restates it in terms of primitives for a simple class of games.

Proposition 1 *In any surplus-maximizing relational contract σ^* of a smooth game:*

1. ***The policy is backward-looking:*** *For any agents i and j, let E_{t+1} be the 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) $d_{i,t+1}^*, d_{j,t+1}^* \in (0,1)$ with positive probability. For almost every $h_0^{t+1} \in E_{t+1}$, $\sigma^*|h_0^{t+1}$ is not surplus-maximizing.*
2. *For all $t \geq 0$, $E_{\sigma^*}\left[\sum_{i=1}^N d_{i,t}\right] = 1$.*

Proof: See Appendix A.

The second statement of Proposition 1 implies that any surplus-maximizing relational contract will use the full "budget" for decisions. Holding all efforts

[7] See Rogerson (1985).

and all other weights fixed, condition 3 of Definition 3 implies that a larger $d_{i,t}$ increases expected output. Condition 4 ensures that a larger $d_{i,t}$ also makes motivating agent i easier in equilibrium. Therefore, increasing $d_{i,t}$ increases expected total surplus and relaxes (3) for agent i, and hence $\sum_{i=1}^{N} d_{i,t} = 1$ in each period of any surplus-maximizing relational contract. The only question is what $d_{i,t}$ is assigned to each agent.

The first statement shows that policies will be backward-looking at histories that satisfy four conditions. Agent i must exert positive but less than first-best effort (condition (i)) and produce "high" output given that effort (condition (ii)), while some other agent j must produce "low" output in every previous period (condition (iii)). Finally, it must be feasible to bias future decisions towards agent i and away from agent j (condition (iv)).

To prove this result, consider a history $h_0^{t+1} \in E_{t+1}$. We assume that $\sigma^*|h_0^{t+1}$ is surplus-maximizing and construct a perturbed equilibrium that strictly dominates σ^*. For now, suppose that increasing $d_{i,t+1}$ smoothly increases i-dyad surplus in the continuation game. If $\sigma^*|h_0^{t+1}$ is surplus-maximizing, then a small increase in $d_{i,t+1}$ or $d_{j,t+1}$ must have identical marginal effects on i-dyad and j-dyad surplus, respectively. Hence, biasing d_{t+1} towards agent i by slightly increasing $d_{i,t+1}$ and decreasing $d_{j,t+1}$ entails a second-order decrease in total continuation surplus, a first-order increase in $S_{i,t+1}$, and a first-order decrease in $S_{j,t+1}$. The upper bound of agent j's dynamic enforcement constraint has never been binding, because he has never produced high output. So he is willing to work equally hard if decisions at h_0^{t+1} are biased away from him. In contrast, the upper bound of agent i's dynamic enforcement constraint binds in period t, because $y_{i,t} > y_{i,t}^*$. Hence, biasing d_{t+1} towards agent i means that agent i can be promised a larger reward in equilibrium, which induces a first-order increase in the maximum $e_{i,t}$ that i is willing to exert. Hence, a small bias towards agent i (and away from agent j) entails a second-order direct cost, no incentive cost, and a first-order incentive benefit. This perturbed equilibrium therefore dominates σ^*.

Two subtleties complicate this intuition. First, increasing $e_{i,t}$ changes the distribution over $y_{i,t}$, which potentially affects other agents' expected dyad-surpluses and hence their incentives. In the proof, we construct a mapping from the perturbed distribution over $y_{i,t}$ to the original distribution over continuation play that induces agent i to work harder while ensuring that all other agents' incentives are unchanged. The second challenge is that our argument requires a change in $d_{i,t+1}$ to have a smooth effect on i-dyad surplus. Condition 3 in Definition 3 implies that, holding $e_{i,t+1}$ fixed, increasing $d_{i,t+1}$ smoothly increases i-dyad surplus in period $t+1$. Conditions 1, 2, and 4 ensure that, holding the distribution over continuation play from period $t+2$ onwards constant, the maximum equilibrium $e_{i,t+1}$ is smoothly increasing in $d_{i,t+1}$. Increasing $d_{i,t+1}$ and decreasing $d_{j,t+1}$ therefore smoothly increases i-dyad surplus and smoothly decreases j-dyad surplus, respectively.

While Proposition 1 is stated in terms of equilibrium efforts and decisions, the result can be restated in terms of primitives for a simple subset of smooth games. To that end, we consider a set of games in which d_t affects agents' expected output, but not the informativeness of that output as a signal of effort.

Definition 4 *A game is* ***mean-shifting*** *if for every* $i \in \{1, ..., N\}$, $\mathcal{E}_i \subseteq [0,1]$ *is a non-singleton and there exist distributions* $\tilde{P}_i^L, \tilde{P}_i^H : \mathbb{R} \to \mathbb{R}$ *and function* $\gamma_i : \Theta \times D \to \mathbb{R}$ *such that*

$$P_i(y_i|\theta, d, e_i) = (1 - e_i)\tilde{P}_i^L(y_i - \gamma_i(\theta, d)) + e_i\tilde{P}_i^H(y_i - \gamma_i(\theta, d)).$$

In a mean-shifting game, agent i's output is drawn from a mixture distribution between $\tilde{P}_i^L$ and $\tilde{P}_i^H$, where effort increases the weight on $\tilde{P}_i^H$. Output $y_{i,t}$ is shifted up by a constant that depends on the state of the world and the principal's decision, $\gamma_i(\theta_t, d_t)$. Suppose that F, c, and $\{\bar{u}_i\}_i$ satisfy conditions 1 and 2 of Definition 3. Condition 3 is satisfied if $\tilde{P}_i^L$ and $\tilde{P}_i^H$ are smooth, with densities $\tilde{p}_i^L$ and $\tilde{p}_i^H$ such that $\frac{\tilde{p}_i^H}{\tilde{p}_i^L}$ is increasing, and if γ_i is smooth in d_i with $\frac{\partial \gamma_i}{\partial d_i} > 0$ and $\frac{\partial^2 \gamma_i}{\partial d_i^2} < 0$.[8] Condition 4 is satisfied for $R_i(\cdot|x)$ equal to the Dirac function at $x + \gamma_i(d, \theta) - \gamma_i(\tilde{d}, \theta)$. We call a mean-shifting game that satisfies

[8]Note that mean-shifting games immediately satisfy CDFC: $\frac{\partial^2 P_i}{\partial e_i^2} = 0$.

Definition 3—which includes examples with a variety of natural distributions $\tilde{P}_i^L$ and $\tilde{P}_i^H$—a **smooth mean-shifting game**.

The (incentive and direct) costs and (incentive) benefits of a backward-looking policy are particularly easy to compare in a smooth mean-shifting game, because decisions only affect expected output. Consequently, we can restate Proposition 1 in terms of primitives for these games.

Corollary 1 *Consider a smooth mean-shifting game such that θ_t is i.i.d.. Suppose $\lim_{d_i \to 0} \frac{\partial \gamma_i}{\partial d_i} = \infty$ and $\min_{e_i} c'(e_i) = 0$ for every $i \in \{1, ..., N\}$. Then there exist $\underline{\delta} < \bar{\delta}$ such that if $\delta \in [\underline{\delta}, \bar{\delta}]$, no surplus-maximizing relational contract σ^* is sequentially surplus-maximizing.*

The limit conditions on the mean-shifting terms $\{\gamma_i\}_{i=1}^N$ imply that $d_{i,t} \in (0, 1)$ in any sequentially surplus-maximizing relational contract. If players are neither too patient nor too impatient, then there exists some agent who exerts positive effort that is less than first-best. Therefore, the conditions for part 1 of Proposition 1 hold with positive probability, proving Corollary 1.

5 Examples

This section uses the tools from Section 4.1 to characterize surplus-maximizing relational contracts in two applied examples that are not smooth. 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. Both examples assume $N = 2$, with $\bar{u}_i = 0$, $e_{i,t} \in \{0, 1\}$, and $c(e_{i,t}) = ce_{i,t}$ for each $i \in \{1, 2\}$.

5.1 Biased Hiring Decisions

Consider a firm that faces persistent demand shocks and decides how many agents to employ in each period. This example illustrates how persistent shocks in demand and diminishing returns in hiring can lead a firm to delay hiring after demand expands.

Definition 5 *The **hiring game** has the following features:*

- *The set of possible states 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 $d_t = 1$, agent 1 is hired.*[9]
- *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 $\frac{1}{2} < \alpha < 1$ if $d_t = 2$.*

The principal faces persistent and growing demand in each period: weak demand ($\theta_t = W$) eventually becomes robust ($\theta_t = R$) and thereafter remains robust. After observing demand in each period, the firm hires either one or two workers. Each hired agent produces output if and only if he works hard. This output is increasing in demand but exhibits diminishing returns—represented by $\alpha < 1$—in the number of workers hired in a period.

Surplus-maximizing relational contracts can exhibit hiring delays in this setting: if $\theta_t = R$, then hiring two workers would be sequentially surplus-maximizing, but the firm might refrain from doing so.

Proposition 2 *In the hiring game, suppose $R > \frac{c}{2\alpha - 1} > W > c$ and $\alpha R > W$. Then there exist $\underline{\delta} < \bar{\delta}$ such that if $\delta \in (\underline{\delta}, \bar{\delta})$, any surplus-maximizing relational contract σ^* satisfies:*

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

There exists a surplus-maximizing relational contract with the following features: agent i always chooses $e_{i,t} = 1$ if hired. If $\theta_t = R$ for the first time in period $t > 0$, then $d_t = 1$ with probability $\chi \in (0, 1)$ and otherwise $d_t = 2$. Then $d_{t'} = d_t$ for every $t' > t$.

[9]This restriction is without loss of generality for our result.

Proof: See Appendix A.

The two conditions in the statement of Proposition 2 ensure that (i) if agents exert effort, then myopic profit is maximized by hiring two workers if $\theta_t = R$ and one worker if $\theta_t = W$, and (ii) 1-dyad surplus is larger if $d_t = 2$ and $\theta_t = R$ than if $d_t = 1$ and $\theta_t = W$. If a firm initially faces robust demand, the optimal relational contract prescribes the same actions in each period. However, if demand is initially weak, then the firm hires only one worker and might delay expanding by continuing to hire only one worker after demand becomes robust. Under the conditions of Proposition 2, agent 1 might be willing to exert effort while $\theta_t = W$ only if decisions are biased towards him once demand becomes robust. The principal biases future decisions towards agent 1 by refraining from hiring agent 2, which decreases total surplus but increases the surplus produced by agent 1, because $\alpha < 1$.

In this example, one surplus-maximizing policy is to make a once-and-for-all expansion decision: once demand becomes robust, the principal expands either immediately or never. This stark policy is optimal because of the linear relationship between decisions and output in this example, but it nevertheless illustrates that the surplus-maximizing relational contract may entail substantial and long-lasting distortions.

This example provides a potential rationale for the empirical puzzle noted by Ariely et al. (2013), which argues that firms that rely on relational contracts tend to expand more slowly than those that rely on formal contracts. Here, hiring remains slow, because fast expansion undermines the firm's ability to credibly motivate its existing workers. New firms face no such trade-off, so they can more easily expand to take advantage of improved demand. Hence, this example would suggest that new entry may drive increased employment immediately after a recession or other period of low demand. Consistent with this implication, Haltiwanger et al. (2013) find that young firms tended to drive net job growth in the US from 1976-2004.

5.2 Irreversible Investments

Suppose the principal chooses one agent to receive a permanent investment that increases that agent's productivity. This investment can be interpreted as training, a promotion, or any other organizational decision that increases the returns from one agent's efforts. The returns from this investment might differ across agents. In which agent should the principal invest?

We show that surplus-maximizing policies might entail a distorted tournament for the investment. The agent who performs "best" according to this tournament is chosen, even if investing in the other agent would lead to a larger increase in total continuation surplus. In this example, the principal chooses one of the two agents to receive an investment in period $t = 1$. Agents have identical productivities without the investment, but agent 1's expected output increases more from the investment than agent 2's. Once made, the investment is permanent.

Definition 6 *The* ***investment game*** *is mean-shifting, with* $|\Theta| = 1$ *and the following features:*[10]

- $d_t = 0$ *denotes that neither agent receives an investment, while* $d_t \in \{1, 2\}$ *indicates agent* d_t *receives investment. Investment occurs in* $t = 1$ *and is permanent:* $D_0 = \{0\}$*,* $D_1 = \{1, 2\}$*, and* $D_t = \{d_{t-1}\}$ *for any* $t > 1$*.*

- *For each* $i \in \{1, 2\}$*,* $\tilde{P}_i^L \equiv \tilde{P}^L$ *and* $\tilde{P}_i^H \equiv \tilde{P}^H$ *from Definition 4 are smooth with density* $\tilde{p}^L$*,* $\tilde{p}^H$*, respectively, and* $\tilde{p}^H$ *strictly MLRP-dominates* $\tilde{p}^L$*.*

- *For each* i*,* $\gamma_i(d_t) = 0$ *for all* $d_t \neq i$*, while* $\gamma_i(i) > 0$*. Assume* $\gamma_1(1) - \gamma_2(2) \equiv \Delta > 0$ *and* $E[y_i|d, e_i = 1] - c > E[y_i|d, e_i = 0] = 0$ *for every* $d \in D$*.*

Define

$$L(y_i) = \frac{\tilde{p}^H(y_i)}{\tilde{p}^L(y_i)}.$$

[10] Consequently, we suppress dependence on θ_t in all expressions.

Then L is strictly increasing in y_i, and there exists a unique y^* with $L(y^*) = 1$.

An agent who expects to receive the investment with high probability can be credibly promised a large reward in period $t = 0$. As in Section 2, the principal can potentially motivate both agents in period $t = 0$ by conditioning investment on realized output in that period. The result is a tournament for investment: the less-efficient agent might receive the investment if he performs well in the first period.

Proposition 3 *In the investment game, there exists $\bar{\Delta} > 0$ such that for $\Delta < \bar{\Delta}$, there exist $0 \leq \underline{\delta} < \bar{\delta} < 1$ such that if $\delta \in (\underline{\delta}, \bar{\delta})$, any surplus-maximizing relational contract σ^* satisfies:*

1. $e_{1,0} = e_{2,0} = 1$;
2. $d_1 = 2$ *if and only if (i)* $L(y_{2,0}) > 1$ *and (ii)*

$$\frac{1}{L(y_{2,0})} < \alpha + \beta \left(\frac{1}{L(y_{1,0})} \right) \tag{5}$$

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

Proof: See Appendix A.

If agents' productivities following investment are not too different ($\Delta < \bar{\Delta}$), then the principal finds it optimal to motivate both agents to exert effort in period $t = 0$. She makes these incentives credible by running a tournament between the two agents, with the cut-off for agent 2 to win the investment given by (5). The chosen agent in period $t = 1$ continues working hard in subsequent periods, because he produces a lot of dyad surplus in equilibrium. The other agent can no longer be motivated to work hard in equilibrium and so shirks in periods $t \geq 1$.

6 The Role of Private Monitoring

This section explores the assumption of private monitoring in our setting. Section 6.1 proves that for smooth games with mean-shifting decisions, backward-looking policies are surplus-maximizing in the full (non-recursive) set of Perfect Bayesian Equilibria. Section 6.2 proves that surplus-maximizing relational contracts are always sequentially surplus-maximizing if monitoring is public. Section 6.3 analyzes an example with a hybrid between private and public monitoring in which deviations are publicly observed with some probability. Surplus-maximizing relational contracts might entail backward-looking policies so long as this probability is strictly less than 1.

6.1 Biased Decisions in Perfect Bayesian Equilibria

This section considers the full set of PBE in the context of smooth mean-shifting repeated games. We show that the intuition from Proposition 1 does not depend on the restriction to recursive equilibria: an analogue of Corollary 1 holds for the full set of PBE.

The central difficulty in extending Corollary 1 is that different players observe different variables and so potentially form different beliefs about the true history in each period. A player's strategy need only be a best response given that player's beliefs, so play at a given history is not necessarily an equilibrium of the continuation game. Our definition of sequentially surplus-maximizing relational contracts relies on the fact that equilibrium play—and hence the equilibrium payoff set—is recursive, so it does not immediately extend to PBE.

We therefore define sequentially surplus-maximizing PBE in terms of *ex ante* expected payoffs rather than continuation payoffs. That is, let $\bar{V} = \max_{\sigma^* \in PBE} E_{\sigma^*}\left[\sum_{i=1}^{N} S_{i,0}\right]$ be the maximum total surplus attainable in a PBE. Then a PBE is **PBE-sequentially surplus-maximizing** if in each $t \geq 0$, $E_{\sigma^*}\left[\sum_{i=1}^{N} S_{i,t}\right] = \bar{V}$. If (θ_t, D_t) is i.i.d., then we show that $E_{\sigma^*}\left[\sum_{i=1}^{N} S_{i,t}\right] \leq \bar{V}$ for any $t \geq 0$. Hence, a PBE-sequentially surplus-maximizing equilibrium maximizes *ex ante* expected continuation surplus in each period.

Lemma 2 *Assume that* (θ_t, D_t) *are i.i.d.. Then for any* $t \geq 0$*, there exists a PBE* σ^* *such that* $E_{\sigma^*}\left[\sum_{i=1}^{N} S_{i,t}\right] = V$ *if and only if there exists a PBE* $\tilde{\sigma}$ *such that* $E_{\tilde{\sigma}}\left[\sum_{i=1}^{N} S_{i,0}\right] = V$.

Proof: See Appendix A.

Lemma 2 shows that total continuation surplus V is attainable beginning in period t in an equilibrium if and only if it is attainable by some equilibrium beginning in period 0. The proof of this result has two steps. First, it establishes an extension of the necessary and sufficient conditions from Lemma 1 for the full set of PBE. This proof is similar to that of Lemma 1, though care must be taken to track each agent's beliefs in each history. As in Lemma 1, this construction guarantees that the principal earns no more than 0 at each history and is indifferent among histories on the equilibrium path.

Second, we use the PBE σ^* satisfying $E_{\sigma^*}\left[\sum_{i=1}^N S_{i,t}\right] = V$ to construct a PBE $\tilde{\sigma}$ with $E_{\tilde{\sigma}}\left[\sum_{i=1}^N S_{i,0}\right] = V$. At the start of the game in $\tilde{\sigma}$, the principal privately chooses a history $h_0^t \in \mathcal{H}_0^t$ according to the distribution over such histories induced by σ^*. She uses her private messages in $t = 0$ to report $\phi_i(h_0^t)$ to each agent i. Play then proceeds as in $\sigma^*|h_0^t$. In this construction, each agent has exactly the same information about continuation play that he would have in $\sigma^*|h_0^t$, so he is willing to play according to $\sigma^*|h_0^t$. The principal is willing to randomize over her initial choice of h_0^t, because she earns 0 at every history on the equilibrium path. Therefore, $\tilde{\sigma}$ is a PBE that replicates in period 0 the distribution over period-t continuation play induced by σ^*.

Lemma 2 shows that our definition of PBE-sequentially surplus-maximizing is well-defined, in the sense that an equilibrium that satisfies it attains the maximum *ex ante* expected continuation surplus in every period. Our main result in this section is an analogue of Corollary 1: in smooth mean-shifting games, there exists a range of discount factors for which no surplus-maximizing PBE is PBE-sequentially surplus-maximizing.

Proposition 4 *Consider a smooth mean-shifting game such that θ_t is i.i.d. and $\lim_{d_i \to 0} \frac{\partial \gamma_i}{\partial d_i} = \infty$ for every $i \in \{1, ..., N\}$. Let $\delta \in (\underline{\delta}, \bar{\delta})$, where $\underline{\delta}$ and $\bar{\delta}$ are the bounds from Corollary 1. Then no surplus-maximizing PBE is PBE-sequentially surplus-maximizing.*

Proof: See Appendix A.

As in Proposition 1 and Corollary 1, backward-looking policies are surplus-maximizing in Proposition 4, because they make strong effort incentives credible. In any sequentially surplus-maximizing PBE, the decision d_t is chosen to maximize total surplus in period t, so

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

must hold for any agents i, j. This condition uniquely pins down d_t^* in each period. Consequently, the continuation surplus produced by each agent does not depend on the actions of any other agents, so agents' beliefs about the true history are irrelevant. As a result, any sequentially surplus-maximizing PBE is payoff-equivalent to a sequentially surplus-maximizing RE. Under the conditions of Corollary 1, surplus-maximizing RE are not sequentially surplus-maximizing, so neither are surplus-maximizing PBE. Hence, backward-looking policies can be surplus-maximizing even if we consider the full set of Perfect Bayesian Equilibria.

6.2 No Biased Decisions Under Public Monitoring

The **game with public monitoring** is identical to the general game in Section 3 with one exception: all variables except e_t are publicly observed, while e_t remains private.[11] Under this monitoring structure, all agents can observe whether the principal deviates. Consequently, the principal faces an aggregate reneging constraint: she is only willing to pay rewards if the *sum* of those rewards is smaller than the *sum* of dyad surpluses. Backward-looking policies tighten this bound and so undermine the principal's ability to credibly promise rewards. This logic, familiar from Levin (2003), implies that backward-looking policies are never surplus-maximizing in the game with public monitoring.

Proposition 5 *In the game with public monitoring, every surplus-maximizing relational contract is sequentially surplus-maximizing.*

[11] Recursive equilibria are equivalent to Perfect Public Equilibria if monitoring is public.

Proof: See Appendix A.

In sharp contrast to Proposition 1, Proposition 5 implies that surplus-maximizing relational contracts never inefficiently condition on past play in the game with public monitoring. The intuition for this result is a natural extension of the intuition given in Section 2, and the proof is a straightforward adaptation of techniques used by Levin (2003) and Goldlucke and Kranz (2012). The principal's most tempting deviation in the game with public monitoring is to simultaneously renege on all agents, since she can be held to her min-max payoff following any deviation. This punishment is substantially harsher than in the game with private monitoring. Importantly, its severity also depends on total continuation surplus rather than i-dyad surplus, which drives the differences between Propositions 1 and 5.

6.3 Biased Decisions Under Imperfect Coordination

In Section 6.2, agents immediately and perfectly coordinate to punish the principal in the game with public monitoring. We believe that these perfectly coordinated punishments are unrealistic in many settings: for instance, they would imply that an employer loses her entire workforce if she withheld a bonus from even a single deserving worker. This section modifies the hiring example from Section 5.1 to allow only imperfect coordination among agents. We show that biased decisions might be surplus-maximizing so long as agents cannot perfectly coordinate punishments.

In the hiring game, suppose that deviations are $\epsilon-$**private:** the first time a given agent chooses $a_{i,t} = 0$, all agents observe this choice with probability $1 - \epsilon$ and otherwise only the principal observes it. Subsequent $a_{i,t} = 0$ are observed only by the principal. In any surplus-maximizing equilibrium of this game, $a_{i,t} = 0$ only following a deviation. Therefore, this monitoring structure gives agents a "once and for all" chance to communicate and coordinate their punishments after the principal deviates.

So long as $\epsilon > 0$, Proposition 6 shows that there exist parameter values for which any surplus-maximizing relational contract has a backward-looking policy.

Proposition 6 *Consider the hiring game with ϵ-private monitoring. If $\epsilon > 0$, 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 intuition for Proposition 6 is fairly straightforward. If the principal reneges on a payment to agent i, then all agents observe i's subsequent rejection with probability $1 - \epsilon$. If so, they all punish the principal, destroying total surplus $\delta E_{\sigma^*}\left[\sum_{j=1}^{N} S_{j,t+1}\right]$. Otherwise, only agent i punishes the principal, destroying total surplus $\delta E_{\sigma^*}\left[S_{i,t+1}\right]$. If $\epsilon > 0$, then agent i's future production is always lost if the principal reneges on i but not if she reneges on agent $j \neq i$. So as in Section 5.1, the principal can make larger rewards to i credible by biasing future hiring decisions towards i.

This basic intuition masks considerable complexity that arises from the monitoring structure. Unlike the proof of Lemma 1, the principal may not be willing to implement some policies in equilibrium. While these additional constraints make a general analysis difficult, the surplus-maximizing policy in the hiring game depends only on current and past demands and hiring, all of which are publicly observed. So we can ensure the principal strictly prefers to follow this policy rather than deviating. Proposition 6 illustrates that, in our hiring example, backward-looking policies might be surplus-maximizing so long as agents can only imperfectly observe one another's relationships.

7 Discussion and Conclusion

We have argued that biased decisions can arise in surplus-maximizing relational contracts, even if the principal may freely pay or be paid by her agents. Biased decisions increase the future surplus produced by one agent at the cost of reducing the surplus produced by others and so complement and make credible large monetary rewards. Consequently, 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 incentives and the promise of future business.

The results of section 6.2 imply that the principal would prefer her payments to be publicly observed, since that would allow agents to jointly punish a deviation. In practice, the principal might try to make some aspects of her payments public information to help her agents coordinate. For instance, the principal might pay agents out of a fixed, public bonus pool. For such a scheme to work, the principal must be able both to commit to make the bonus pool public and to refrain from secretly diverting funds from it. Of course, such pools might also entail dynamic inefficiencies, particularly if the size of the pool depends on the outputs realized in each period.

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 agents' efforts are either substitutes or complements. 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 desired policy in such environments. We conjecture that conditions similar to those in Lemma 1 are necessary, but not sufficient, if efforts are substitutes. If efforts are complements, then the relational contract must deter the principal from reneging on multiple agents at once, which further complicates the analysis.

We have presented a few simple examples of how these biases might manifest, but further research is needed to expand the scope of these examples. In particular, Section 5.1 suggests that a richer analysis of hiring and firing decisions in firms might lead to further empirical predictions.

References

Abreu, D., D. Pearce, and E. Stacchetti (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica 58*(5), 1041–1063.

Andrews, I. and D. Barron (forthcoming). The allocation of future business: Dynamic relational contracts with multiple agents. *American Economic Review*.

Ariely, D., S. Belenzon, and U. Tsolmon (2013). Health insurance and relational contracts in small american firms. Unpublished Manuscript.

Asanuma, B. (1989). Manufacturer-supplier relationships in japan and the concept of relation-specific skill. *Journal of the Japanese and International Economics 3*(1), 1–30.

Baker, G., R. Gibbons, and K. J. Murphy (1994). Subjective performance measures in optimal incentive contracts. *Quarterly Journal of Economics 109*(4), 1125–1156.

Board, S. (2011). Relational contracts and the value of loyalty. *American Economic Review 101*(7), 3349–3367.

Bull, C. (1987). The existence of self-enforcing implicit contracts. *Quarterly Journal of Economics 102*(1), 147–159.

Ely, J., J. Horner, and W. Olszewski (2005). Belief-free equilibria in repeated games. *Econometrica 73*(2), 377–415.

Fong, Y.-F. and J. Li (2016). Relational contracts, limited liability, and employment dynamics. Unpublished Manuscript.

Fudenberg, D., B. Holmstrom, and P. Milgrom (1990). Short-term contracts and long-term agency relationships. *Journal of Economic Theory 51*(1), 1–31.

Fudenberg, D. and D. Levine (1994). Efficiency and observability in games with long-run and short-run players. *Journal of Economic Theory 62*, 103–135.

Fudenberg, D., D. Levine, and E. Maskin (1994). The folk theorem with imperfect public information. *Econometrica 62*(5), 997–1039.

Goldlucke, S. and S. Kranz (2012). Infinitely repeated games with public monitoring and monetary transfers. *Journal of Economic Theory 147*(3), 1191–1221.

Graham, J., C. Harvey, and M. Puri (2015). Capital allocation and delegation of decision-making authority within firms. *Journal of Financial Economics 115*(3), 449–470.

Green, E. and R. Porter (1984). Noncooperative collusion under imperfect price information. *Econometrica 52*(1), 87–100.

Halac, M. (2012). Relational contracts and the value of relationships. *American Economic Review 102*(2), 750–779.

Haltiwanger, J., R. Jarmin, and J. Miranda (2013). Who creates jobs? small versus large versus young. *Review of Economics and Statistics 95*(1), 347–361.

Levin, J. (2002). Multilateral contracting and the employment relationship. *Quarterly Journal of Economics 117*(3), 1075–1103.

Levin, J. (2003). Relational incentive contracts. *American Economic Review 93*(3), 835–857.

Li, J., N. Matouschek, and M. Powell (forthcoming). Power dynamics in organizations. *American Economic Journal: Microeconomics*.

MacLeod, B. and J. Malcomson (1989). Implicit contracts, incentive compatibility, and involuntary unemployement. *Econometrica 57*(2), 447–480.

Malcomson, J. (2013). Relational incentive contracts. In R. Gibbons and J. Roberts (Eds.), *Handbook of Organizational Economics*, pp. 1014–1065.

Malcomson, J. (forthcoming). Relational incentive contracts with persistent private information. *Econometrica*.

Peter, L. and R. Hull (1969). *The Peter Principle*. New York, NY: William Morrow.

Rogerson, W. (1985). The first-order approach to principal-agent problems. *Econometrica 53*(6), 1357–1367.

Watson, J. (2016). Perfect bayesian equilibrium: General definitions and illustrations. Unpublished Manuscript.

A For Online Publication: Proofs

A.1 Proof of Lemma 1

Part 1: Given RE σ^*, define $B_i : \mathcal{H}_d^t \times \Xi \times \mathbb{R} \to \mathbb{R}$ by

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

Following on-path history h_0^t, $\sigma^*|h_0^t$ is a Perfect Bayesian Equilibrium. So for any successor h_d^t, ξ_t, agent i is willing to choose $a_{i,t}$, $e_{i,t}$ only if (2) holds.

Suppose $B_i(h_d^t, \xi_{i,t}, y_{i,t}) < \delta E_{\sigma^*}\left[\bar{U}_i(h_0^{t+1})|h_d^t\right]$. Then $\tau_{i,t} < 0$ because $E\left[U_{i,t+1}|h_0^{t+1}\right] \geq \bar{U}_i(h_0^{t+1})$, so agent i may profitably deviate by choosing $\tau_{i,t} = 0$, which implies (3). Suppose $B_i(h_d^t, \xi_{i,t}, y_{i,t}) > \delta E_{\sigma^*}\left[S_{i,t+1}|h_d^t, \xi_{i,t}, y_{i,t}\right]$. Then there exists some history h_y^t consistent with $(h_d^t, \xi_{i,t}, y_{i,t})$ such that this inequality holds. Suppose the principal deviates by paying $\tau_{i,t'} = w_{i,t'} = 0$ for all $t' \geq t$ but otherwise playing according to the distribution $\sigma^*| \cup_{j \neq i} \phi_j(h_0^{t+1})$. Agent i detects this deviation but can punish the principal no more harshly than $y_{i,t'} = w_{i,t'} = \tau_{i,t'} = 0$ in all future periods. The other agents do not detect this deviation and so do not condition their play on it. Outputs and transfers do not affect the continuation game, so this deviation is feasible. The principal's payoff following it is bounded below by

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

Therefore, the principal is willing to pay $\tau_{i,t}$ only if

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

Adding $\delta U_{i,t+1}$ to both sides of this expression and taking expectations conditional on $h_d^t, \xi_{i,t}, y_{i,t}$ yields the right-hand inequality in (3). $\square$

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

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

2. At history $h_0^{t,*}$, after θ_t^* and D_t^* are realized, the principal draw $h_e^t \in \mathcal{H}_e^t$ from the distribution $\sigma|\{h_0^t, \theta_t^*, D_t^*\}$. The principal chooses d_t^* as in h_e^t.

3. For each $i \in \{1, ..., N\}$, the principal pays

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

Note that $w_{i,t}^* \geq 0$, because $E_\sigma [y_{i,t} | h_d^t, \xi_{i,t}] \geq 0$ by assumption and (3) holds. The principal sends messages

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

4. Agent i chooses $a_{i,t}^* = a_{i,t}$, $e_{i,t}^* = e_{i,t}$, where $(a_{i,t}, e_{i,t})$ are inferred from $m_{i,t}^*$.

5. Following output y_t^*, for each agent $i \in \{1, ..., N\}$,

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

where agent i infers the right-hand side from $m_{i,t}^*$. Note $\tau_{i,t}^* \leq 0$ by (3).

6. Let $h_0^{t+1,*}$ be the realized history at the start of $t+1$. The principal draws $h_0^{t+1} \in \mathcal{H}_0^{t+1}$ from $\sigma|\{h_e^t, y_t\}$. Then $h_0^{t+1,*}$ and h_0^{t+1} induce identical continuation games. Repeat this construction with $h_0^{t+1}, h_0^{t+1,*}$.

7. Following a deviation: if agent i observes a deviation (except in $e_{i,t}$), he takes his outside option and pays no transfers in this and every subsequent period. If the principal observes the deviation, then $m_{j,t'} = w_{j,t'} = \tau_{j,t'} = 0$ for each $j \in \{1, ..., N\}$ 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.

By construction, h_0^t and $h_0^{t,*}$ induce the same continuation game in each period on the equilibrium path. Therefore, total continuation surplus and i-dyad surplus for each $i \in \{1, ..., N\}$ are identical in $\sigma^*|h_0^{t,*}$ and $\sigma|h_0^t$ by construction.

Deviations by the Principal: For any on-path $h_d^{t,*}$ and agent $i \in \{1, ..., N\}$, the distribution over $y_{i,t}^*$ is identical to $\sigma|h_d^t$. So

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

and hence $E_{\sigma^*}\left[\Pi_{i,t}|h_d^{t,*}\right] = 0$. If the principal deviates in d_t^*, $w_{i,t}^*$, or $m_{i,t}^*$, then each agent i either observes this deviation or not. If agent i observes the deviation, then the principal earns 0 from that and every other agent. If agent i does not observe the deviation, then $m_{i,t}^*$ must include a history $\tilde{h}_d^{t,*}$ such that $E_{\sigma^*}[y_{i,t} - \tilde{w}_{i,t} - \tau_{i,t}|\tilde{h}_d^{t,*}] = 0$ given the wage $\tilde{w}_{i,t}$ included in $m_{i,t}^*$. But agent i determines the distribution over $y_{i,t}$ and $\tau_{i,t}$, so the principal must earn 0 following such a deviation. A nearly identical argument applies off the equilibrium path. The principal takes no other costly actions, so we conclude she has no profitable deviation.

Deviations by Agent i: If agent i deviates in period t, then the principal min-maxes him, so he earns continuation surplus $E_{\sigma^*}\left[U_{i,t+1}|h_0^{t+1,*}\right] = \bar{U}_i(h_0^{t+1,*}) = \bar{U}_i(h_0^{t+1})$. Off-path, i has no profitable deviation, because $\bar{u}_i(d_t, \theta_t) \geq 0$.

At each on-path $h_0^{t,*}$, we must show that agent i has no profitable deviation in $e_{i,t}^*$ or $\tau_{i,t}^*$ (agent i can never profitably deviate from $w_{i,t}^* \geq 0$). In σ^*, $E_{\sigma^*}\left[U_{i,t}|h_0^{t,*}\right] = E_{\sigma^*}\left[S_{i,t}|h_0^{t,*}\right]$. So agent i chooses $a_{i,t}^*, e_{i,t}^*$ to maximize

$$E_{\sigma^*}\left[(1-\delta)\tau_{i,t}^* + \delta S_{i,t+1}|h_d^{t,*}, \xi_{i,t}^*, a_{i,t}, e_{i,t}\right] - c(e_{i,t}),$$

because he infers $h_d^{t,*}$ from D_t^*, θ_t^*, d_t^*, and $m_{i,t}^*$. Plugging in $\tau_{i,t}^*$ yields

$$E_{\sigma^*}\left[B_i(h_d^t, \xi_{i,t}, y_{i,t}^*) - \delta E_\sigma\left[S_{i,t+1}|h_d^t, \xi_{i,t}, e_{i,t}\right] + \delta S_{i,t+1}|h_d^{t,*}, \xi_{i,t}^*, a_{i,t}, e_{i,t}\right] - c(e_{i,t}).$$

Now, $E_{\sigma^*}\left[B_i(h_d^t,\xi_{i,t},y_{i,t})|h_d^{t,*},\xi_{i,t}^*,a_{i,t},e_{i,t}\right] = E_\sigma\left[B_i(h_d^t,\xi_{i,t},y_{i,t})|h_d^t,\xi_{i,t},a_{i,t},e_{i,t}\right]$ because the distribution over $y_{i,t}$ is identical in $\sigma|h_d^t$ and $\sigma^*|h_d^{t,*}$. By construction, $\sigma^*|h_e^{t,*}$ and $\sigma|h_e^t$ generate the same distributions over i-dyad surplus in period $t+1$ onwards, so $E_{\sigma^*}\left[S_{i,t+1}|h_d^{t,*},\xi_{i,t}^*,a_{i,t},e_{i,t}\right] = E_\sigma\left[S_{i,t+1}|h_d^t,\xi_{i,t},a_{i,t},e_{i,t}\right]$. Therefore, (2) implies that agent i has no profitable deviation from $e_{i,t}^*$.

Agent i is willing to pay $\tau_{i,t}^* < 0$ if

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

As above, $E_{\sigma^*}\left[S_{i,t+1}|h_d^{t,*},\xi_{i,t}^*,y_{i,t}^*\right] = E_\sigma\left[S_{i,t+1}|h_d^t,\xi_{i,t},y_{i,t}^*\right]$ by construction. Further, $E_{\sigma^*}\left[\bar{U}_i(h_0^{t+1})|h_d^{t,*},\xi_{i,t}^*,y_{i,t}^*\right] = E_{\sigma^*}\left[\bar{U}_i(h_0^{t+1})|h_d^t\right]$, because $h_0^{t,*}$ and h_0^t induce the same continuation game, and (θ_t,d_t) are the same in h_d^t and $h_d^{t,*}$. Agent i is willing to pay $\tau_{i,t}^*$ if

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

which is implied by the left-hand inequality in (3).

We conclude that σ^* is an RE with the desired properties. ■

A.2 Proof of Proposition 1

A.2.1 A Guide for the Reader

The first statement is the complicated part of the proof. Broadly, this proof proceeds by contradiction and includes three elements.

Suppose that continuation play at $h_0^{t+1} \in E_{t+1}$ is surplus-maximizing. First, we show that we can perturb the equilibrium to smoothly increase $E[S_{i,t+1}|h_0^{t+1}]$ as $E[S_{j,t+1}|h_0^{t+1}]$ decreases. This step involves increasing $d_{i,t+1}$, decreasing $d_{j,t+1}$, and showing that these changes affect period $t+1$ effort in a smooth way holding continuation play fixed. Second, we show that if i-dyad surplus $E[S_{i,t+1}|h_0^{t+1}]$ for $h_0^{t+1} \in E_{t+1}$ increases, then we can smoothly increase agent i's equilibrium effort in period t holding all other agents' efforts fixed. This step involves constructing a perturbation such that each agent $j \neq i$ faces the same mapping from j's output to j-dyad surplus, even as i's effort changes. Finally, we argue that increasing i-dyad surplus and decreasing j-dyad surplus leads to a second-order loss in total surplus for periods $t+1$ onwards, but allows for a first-order gain in agent i's effort (holding all other efforts fixed). Hence, such a perturbation increases total *ex ante* expected surplus, and so no surplus-maximizing equilibrium can be sequentially surplus-maximizing if $\Pr\{E_{t+1}\} > 0$ for any $t+1 > 0$.

We outline the six steps involved in this proof below. The parenthetical comments at the start of each step roughly link that step to the corresponding elements described above.

1. (Sets up elements 1 and 2) We define a function $G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \tilde{e}_i)$ that takes as input the state of the world θ, an "original" weight and effort pair for agent i (d_i, e_i), a "new" weight and effort pair $(\tilde{d}_i, \tilde{e}_i)$, and a realized output y_i. If y_i is drawn from the "new" distribution $P_i(\cdot|\theta, \tilde{d}_i, \tilde{e}_i)$, then $G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \tilde{e}_i)$ is distributed according to the "original" distribution $P_i(\cdot|\theta, d_i, e_i)$.

2. (Sets up elements 1 and 2) We define $\hat{e}_i$, one of the key functions for the argument. Given a reference (θ, d_i, e) and a new decision $\tilde{d}_i$, $\hat{e}_i$ gives one feasible effort that can be induced in equilibrium, holding the distribution over continuation play fixed at the distribution under (θ, d_i, e). To implement $\hat{e}_i$, transform the realized output y_i by $G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \hat{e}_i)$ and then reward agent i according to a "one step" reward scheme that punishes the agent if $y_i < y_i^*(\theta, d_i, e_i)$ and otherwise rewards the agent. Claim 2 gives conditions under which $\hat{e}_i$ is differentiable in $\tilde{d}_i$.

3. (Used in elements 1 and 2) Claim 3 rearranges (2) and (3) to give a single necessary and sufficient condition for effort $e_{i,t}^*$ to be induced in equilibrium, holding the mapping from output to i-dyad surplus fixed. Since P_i satisfies MLRP and CDFC, we can replace (2) with its first-order condition. To maximize i's effort, the lower bound of (3) should bind for $y_i < y_i^*(\theta, d_i, e_i)$, and the upper bound should bind otherwise.

4. (Used in elements 1 and 2) Claim 4 serves two purposes. First, it confirms a condition required by Claim 2. Second, if the inequality identified in Claim 3 holds with equality, then $e^*_{i,t} = \hat{e}_i(\theta_t, d^*_{i,t}, d^*_{i,t}, e^*_{i,t})$.

5. (Completes element 1, sets up element 3) Claim 5 gives a necessary condition for a continuation equilibrium $\sigma^*|h_0^t$ to be surplus-maximizing. For any $i, j \in \{1, ..., N\}$, if increasing $d_{i,t}$ and decreasing $d_{j,t}$ is feasible, doing so cannot increase total continuation surplus. To prove this result, we use Claim 4 to show that either (i) the necessary and sufficient condition from Claim 3 is slack, or (ii) $e_{i,t} = \hat{e}_i(\theta_t, d_{i,t}, d_{i,t}, e_{i,t})$. If (i), we perturb $d_{i,t}$ to $\tilde{d}_{i,t}$, transform $y_{i,t}$ by $G_i(y_{i,t}|\theta_t, d_{i,t}, \tilde{d}_{i,t}, e_{i,t}, e_{i,t})$, and map this perturbed output to continuation play as in the original equilibrium. For a small enough perturbation, $e_{i,t}$ continues to satisfy the condition from Claim 3, so it can be induced in equilibrium. If (ii), then $e_{i,t}$ might violate the condition from Claim 3 under $\tilde{d}_{i,t}$. However, in that case $e_{i,t} = \hat{e}_i(\theta_t, d_{i,t}, d_{i,t}, e_{i,t})$, and Claim 2 implies that $\hat{e}_i$ is differentiable in its third argument. So we can implement effort $\hat{e}_i(\theta_t, d_{i,t}, \tilde{d}_{i,t}, e_{i,t})$, transform output by $G_i(y_{i,t}|\theta_t, d_{i,t}, \tilde{d}_{i,t}, e_{i,t}, \hat{e}_i)$, and preserve the same distribution over continuation play from period $t+2$ onwards.

6. (Completes elements 2 and 3) We consider $h_0^{t+1} \in E_t$. If $\sigma^*|h_0^{t+1}$ is surplus-maximizing, Claim 5 implies that increasing $d_{i,t+1}$ and decreasing $d_{j,t+1}$ has a second-order effect on total continuation surplus. Condition 4 of Definition 3 implies that the *most efficient* e_i satisfying (2) and (3), holding the distribution over continuation play fixed, is more efficient if d_i is larger. Because $E[y_i|\theta_i, d_i, e_i]$ is strictly increasing in d_i, a small increase in $d_{i,t+1}$ increases $E[S_{i,t+1}|h_0^{t+1}]$. Because $e^*_{i,t} < e_i^{FB}(\theta_t, d^*_{i,t})$, increasing $E[S_{i,t+1}|h_0^{t+1}]$ following a realization $y_{i,t} > y_i^*(\theta_t, d^*_{i,t}, e^*_{i,t})$ allows for strictly higher effort for agent i in period t, even if we otherwise hold the distribution over continuation play fixed. Agent j's effort in period t is unchanged because the upper bound of (3) is not binding for j. Consequently, perturbing $\sigma^*|h_0^{t+1}$ in this way leads to a first-order increase in period-t surplus, which is strictly larger than the second-

order loss in period $t+1$ surplus from the perturbation of d_{t+1}. So in a surplus-maximizing relational contract, continuation play at h_0^{t+1} cannot be surplus-maximizing.

A.2.2 Proof of Statement 1

The inverse distribution P_i^{-1} is continuously differentiable because P_i is strictly increasing and continuously differentiable. Because $\bar{U}_i(h_d^t)$ depends only on θ_t, we abuse notation to write these punishment payoffs $\bar{U}_i(\theta_t)$.

Definition A.1: Define G_i by

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

When unambiguous, we will suppress the conditioning variables in G_i.

Claim 1: If y_i has distribution $P_i(y_i|\theta, \tilde{d}_i, \tilde{e}_i)$, then $x_i \equiv G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \tilde{e}_i)$ has distribution $P_i(x_i|\theta, d_i, e_i)$.

Proof of Claim 1: It suffices to show that

$$P_i(y_i|\theta, \tilde{d}_i, \tilde{e}_i) = P_i\left(G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \tilde{e}_i)|\theta, d_{i,} e_i\right)$$

which is true by definition of G_i. $\square$

Definition A.2: For monotonically increasing $S_i : \mathbb{R} \to \mathbb{R}$, define $\hat{e}_i(\theta, d_i, \tilde{d}_i, e_i|S_i)$ implicitly by

$$0 = \begin{array}{c} \int_{-\infty}^{y_i^*(\theta,d_i,e_i)} \bar{U}_i(\theta) \frac{\partial p_i}{\partial e_i}(y_i|\theta, \tilde{d}_i, \hat{e}_i) dy_i + \\ \int_{y_i^*(\theta,d_i,e_i)}^{\infty} S_i\left(G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \hat{e}_i)\right) \frac{\partial p_i}{\partial e_i}(y_i|\theta, \tilde{d}_i, \hat{e}_i) dy_i - c'(\hat{e}_i) \end{array}. \tag{6}$$

Claim 2: Suppose $(\theta, d_i, \tilde{d}_i, e_i)$ satisfies $d_i = \tilde{d}_i$ and $\hat{e}_i(\theta, d_i, \tilde{d}_i, e_i|S_i) = e_i$. Then $\hat{e}_i$ is differentiable in $\tilde{d}_i$ on a neighborhood about that point.

Proof of Claim 2: Let S_i be a monotonically increasing function. Denote the right-hand side of (6) by H. Then H is continuously differentiable in $\tilde{d}_i$ and $\hat{e}_i$, so $\frac{\partial \hat{e}_i}{\partial \tilde{d}_i}$ exists about $(\theta, d_i, \tilde{d}_i, e_i)$ by the Implicit Function Theorem if $\frac{\partial H}{\partial \hat{e}_i} \neq 0$.[12]

To show that $\frac{\partial H}{\partial \hat{e}_i} \neq 0$, we bound H from above by a function $\bar{H}$ satisfying $H = \bar{H}$ at (θ, d_i, d_i, e_i), with $\frac{\partial \bar{H}}{\partial \hat{e}_i} < 0$ on a neighborhood about that point. For $\epsilon > 0$, let

$$\bar{H} = \begin{array}{c} \int_{-\infty}^{y_i^*(\theta,d_i,e_i)} \bar{U}_i(\theta) \frac{\partial p_i}{\partial e_i}(y_i|\theta, d_i, \hat{e}_i) dy_i + \\ \int_{y_i^*(\theta,d_i,e_i)}^{y_i^*(\theta,d_i,e_i)+\epsilon} S_i\left(G_i(y_i|\theta, d_i, d_i, e_i, \hat{e}_i)\right) \frac{\partial p_i}{\partial e_i}(y_i|\theta, d_i, \hat{e}_i) dy_i + \\ \int_{y_i^*(\theta,d_i,e_i)+\epsilon}^{\infty} S_i(y_i) \frac{\partial p_i}{\partial e_i}(y_i|\theta, d_i, \hat{e}_i) - c'(\hat{e}_i) \end{array}.$$

At $\hat{e}_i = e_i$, $G_i(y_i) = y_i$ and so $\bar{H} = H$. For $\hat{e}_i > e_i$ sufficiently close, we claim that $\bar{H} \geq H$. Note that $G_i(y_i) \leq y_i$ if $\hat{e}_i \geq e_i$ because P_i is FOSD increasing in e_i. Since S_i is monotonically increasing, we must have $S_i(G_i(y_i)) \leq S_i(y_i)$. Further, for $\hat{e}_i$ sufficiently close to e_i, $\frac{\partial p_i}{\partial e_i}(y_i|\theta, d_i, \hat{e}_i) \geq 0$ for $y_i \geq y_i^*(\theta, d_i, e_i) + \epsilon$ because $\frac{\partial p_i}{\partial e_i}(\cdot|\theta, d_i, e_i)$ is strictly increasing in y_i and equals 0 at $y_i^*(\theta, d_i, e_i)$. This proves that $\bar{H} \geq H$.

If $\epsilon = 0$, then $\frac{\partial \bar{H}}{\partial \hat{e}_i} < 0$ by CDFC. It can be shown that $\frac{\partial \bar{H}}{\partial \hat{e}_i}$ is continuous in ϵ, so $\frac{\partial \bar{H}}{\partial \hat{e}_i} < 0$ for $\epsilon > 0$ sufficiently small. So $\bar{H}$ satisfies the desired properties, and hence $\frac{\partial H}{\partial \hat{e}_i} < 0$. $\square$

Claim 3: Consider an equilibrium σ^*. Fix $(h_d^t, \xi_{i,t}^*)$ on the equilibrium path. For each agent i and on-path effort $e_{i,t}^*$, there exists a reward scheme B_i that satisfies (2) and (3) if and only if either (i) $e_{i,t}^* = \min \mathcal{E}_i$, or (ii)

$$c'(e_{i,t}^*) \leq \begin{array}{c} \int_{-\infty}^{y_i^*(\theta_t,d_{i,t},e_{i,t}^*)} \bar{U}_i(\theta_t) \frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) dy_i + \\ \int_{y_i^*(\theta_t,d_{i,t},e_{i,t}^*)}^{\infty} E_{\sigma^*}[S_i|h_d^t, \xi_{i,t}^*, y_i] \frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) dy_i \end{array}. \tag{7}$$

[12] The first term in H is continuously differentiable in $\tilde{d}_i$ and $\hat{e}_i$ because p_i and y_i^* are both continuously differentiable. To show that the second term is differentiable, apply the change of variable $x = G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \hat{e}_i)$.

Proof of Claim 3: Suppose $e_{i,t}^* > \min \mathcal{E}_i$ does not satisfy (7). Because p_i satisfies MLRP and CDFC, we can replace (2) with its first-order condition as in Rogerson (1985):

$$c'(e_{i,t}^*) = \int_{-\infty}^{\infty} B_i(h_d^t, \xi_{i,t}^*, y_i) \frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) dy_i. \tag{8}$$

Consider choosing B_i to maximize the right-hand side of this equality, subject to the constraint (3). We can solve this problem for each y_i: if $\frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) < 0$, then $B_i(h_d^t, \xi_{i,t}^*, y_i) = \bar{U}_i(\theta_t)$, and otherwise $B_i(h_d^t, \xi_{i,t}^*, y_i) = E_{\sigma^*}[S_i|h_d^t, \xi_{i,t}^*, y_i]$. But this is exactly the B_i implemented in (7). Contradiction.

If $e_{i,t}^* = \min \mathcal{E}_i$, then the reward scheme $B_i(h_d^t, \xi_{i,t}^*, y_i) = \bar{U}_i(\theta_t)$ induces $e_{i,t}^*$ because $c(e_{i,t})$ is monotonically increasing. Suppose $e_{i,t}^* > \min \mathcal{E}_i$ satisfies (7). Clearly, the right-hand side of (8) is strictly smaller than the left-hand side if $B_i(h_d^t, \xi_{i,t}^*, y_i) = \bar{U}_i(\theta_t)$. The right-hand side of (8) is continuous in B_i, so we can apply the Intermediate Value Theorem to conclude that there exists some reward scheme B_i such that (8) is satisfied. $\square$

Claim 4: Let σ^* be a surplus-maximizing equilibrium, and fix some $(h_d^t, \xi_{i,t}^*)$ on the equilibrium path. Define $S_i(y_{i,t}) = E_{\sigma^*}\left[S_{i,t+1}|h_d^t, \xi_{i,t}^*, y_{i,t}\right]$. Without loss, $S_i(y_{i,t})$ is increasing in $y_{i,t}$. Moreover, if (7) holds with equality at $e_{i,t}^*$, then $e_{i,t}^* = \hat{e}_i(\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*|S_i)$.

Proof of Claim 4: Suppose there exists $y_i < \tilde{y}_i$ such that $S_i(y_i) > S_i(\tilde{y}_i)$. Consider the following alternative: with probability $\epsilon > 0$, outcome $\tilde{y}_i$ is treated as y_i. With probability $\frac{p_i(\tilde{y}_i|\theta_t, d_{i,t}, e_{i,t}^*)}{p_i(y_i|\theta_t, d_{i,t}, e_{i,t}^*)}\epsilon$, outcome y_i is treated as outcome $\tilde{y}_i$. Agents $j \neq i$ face identical distributions over continuation play and so exert the same effort in each period. For agent i, this perturbation relaxes (7) if and only if

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

Both terms on the left-hand side are strictly positive: the first by assumption, the second by strict MLRP. So this perturbation strictly relaxes (7) for agent i without affecting it for $j \neq i$. So we can assume S_i is increasing without loss.

Suppose (7) holds with equality. Note that $G_i(y_i|\theta_t, d^*_{i,t}, d^*_{i,t}, e^*_{i,t}, e^*_{i,t}) = y_i$ for all y_i. Therefore, $\hat{e}_i(\theta_t, d^*_{i,t}, d^*_{i,t}, e^*_{i,t}|S_i)$ and $e^*_{i,t}$ are both defined implicitly by (7) holding with equality. $\square$

Claim 5: Define

$$s_i(\theta_t, d_{i,t}, e_{i,t}) = E\left[y_{i,t}|\theta_t, d_{i,t}, e_{i,t}\right] - c(e_{i,t}).$$

For any $h_0^t \in \mathcal{H}_0^t$, suppose $\sigma^*|h_0^t$ is surplus-maximizing with $d_{i,t}, d_{j,t} \in (0,1)$. Define $\mathbb{I}_{i,t} = 1$ if (7) holds with equality at a successor history h_d^t, and $\mathbb{I}_{i,t} = 0$ otherwise. Define $\hat{e}_i = \hat{e}_i(\theta_t, d_{i,t}, d_{i,t}, e_{i,t})$. Then for any $i, j \in \{1, ..., N\}$,

$$\frac{\partial s_i}{\partial d_i} + \mathbb{I}_{i,t}\frac{\partial s_i}{\partial e_i}\frac{\partial \hat{e}_i}{\partial \tilde{d}_i} = \frac{\partial s_j}{\partial d_j} + \mathbb{I}_{j,t}\frac{\partial s_j}{\partial e_j}\frac{\partial \hat{e}_j}{\partial \tilde{d}_j} \tag{9}$$

with probability 1 following h_0^t.

Proof of Claim 5: Suppose towards contradiction that the left-hand side of (9) is strictly larger than the right-hand side. Consider the following perturbation (denoted by tildes): $\tilde{d}_{i,t} = d_{i,t} + \epsilon$, $\tilde{d}_{j,t} = d_{j,t} - \epsilon$, $\tilde{e}_{i,t} = \hat{e}_i(\theta_t, d_{i,t}, \tilde{d}_{i,t}, e_{i,t})$ if $\mathbb{I}_{i,t} = 1$ and $\tilde{e}_{i,t} = e_{i,t}$ otherwise, and $\tilde{e}_{j,t} = \hat{e}_j(\theta_t, d_{j,t}, \tilde{d}_{j,t}, e_{j,t})$ if $\mathbb{I}_{j,t} = 1$ and $\tilde{e}_{j,t} = e_{j,t}$ otherwise. For all agents $k \notin \{i, j\}$, $\tilde{d}_{k,t} = d_{k,t}$ and $\tilde{e}_{k,t} = e_{k,t}$. Continuation play is as in σ^*, except $y_{i,t}$ is transformed by $G_i(\cdot|\theta, d_{i,t}, \tilde{d}_{i,t}, e_{i,t}, \tilde{e}_{i,t})$, and similarly with $y_{j,t}$ and G_j.

We claim that there exists a credible reward scheme for each agent in this perturbation, and hence this perturbation is also a continuation equilibrium. By Claim 3, it suffices to show that this alternative satisfies (7). For each agent $k \in \{1, ..., N\}$, this perturbation induces an identical marginal distribution over continuation play from $t+1$ onwards. So for agents $k \notin \{i, j\}$, the credible reward scheme in the original equilibrium remains credible in this perturbation.

Consider agent $k \in \{i,j\}$. If $\mathbb{I}_{k,t} = 0$, then (7) was slack in the original equilibrium. But (7) and G_i are continuous in $d_{k,t}$, so $e_{k,t}$ continues to satisfy it in the perturbed equilibrium if ϵ is sufficiently small. If $\mathbb{I}_{k,t} = 1$, the reward scheme

$$\tilde{B}_k(y_{k,t}) = \begin{cases} \bar{U}_k(\theta_t) & y_{k,t} \leq y_k^*(\theta_t, d_{k,t}, e_{k,t}) \\ S_k(G_k(y_{k,t})) & y_{k,t} > y_k^*(\theta_t, d_{k,t}, e_{k,t}) \end{cases}$$

is credible. These reward schemes satisfy (8) at $\hat{e}_k$ by definition.

Finally, we argue that this perturbation yields strictly higher total surplus than $\sigma^*|h_0^t$, which contradicts the claim that $\sigma^*|h_0^t$ is surplus-maximizing. Because total surplus in period $t+1$ onwards is identical in the original and perturbed equilibrium. It suffices to consider total surplus in period t. Agents $k \notin \{i,j\}$ produce identical period-t surplus in both equilibria. Consider the difference in surplus for agents i and j. The perturbed equilibrium generates no more total surplus than the original equilibrium only if

$$s_i(\theta_t, d_{i,t}+\epsilon, \tilde{e}_{i,t}) + s_j(\theta_t, d_{j,t}-\epsilon, \tilde{e}_{j,t}) - (s_i(\theta_t, d_{i,t}, e_{i,t}) + s_j(\theta_t, d_{j,t}, e_{j,t})) \leq 0 \quad (10)$$

Dividing by $\epsilon > 0$, and taking the limit as $\epsilon \to 0$ results in (9) with a weak inequality $\leq$. Contradiction; we assumed $>$. $\square$

Completing the proof of Statement 1 Let $h_0^{t+1} \in E_{t+1}$. If $\sigma^*|h_0^{t+1}$ is surplus-maximizing, then (9) holds by Claim 5. Let $h_d^t \in \mathcal{H}_d^t$ be a predecessor to h_0^{t+1}, and consider the following perturbation at $\sigma^*|h_d^t$: $\tilde{e}_{i,t} = e_{i,t}^* + \eta$ for some $\eta > 0$ determined below, while $\tilde{e}_{k,t} = e_{k,t}^*$ for all $k \neq i$. At the end of period t, agent i's output is transformed by $G_i(y_{i,t}|\theta_t, d_{i,t}^*, d_{i,t}^*, e_{i,t}^*, \tilde{e}_{i,t})$, and this transformed output is henceforth treated as the realized output.

If $y_{i,t} \geq y_{i,t}^*(\theta_t, \tilde{d}_{i,t})$ and $y_{j,t} < y_{j,t}^*(\theta_t, \tilde{d}_{j,t})$, then $\tilde{d}_{i,t+1} = d_{i,t+1}^* + \epsilon$, $\tilde{d}_{j,t+1} = d_{j,t+1}^* - \epsilon$, and $\tilde{d}_{k,t+1} = d_{k,t+1}^*$ for $k \notin \{i,j\}$. Agent i's effort equals the more efficient of $e_{i,t+1}^*$ and $\hat{e}_i(\theta_{t+1}, d_{i,t+1}^*, \tilde{d}_{i,t+1}, e_{i,t+1}^*)$, while agent j's effort is $\tilde{e}_{j,t+1} = e_{j,t+1}^*$ if $\mathbb{I}_{j,t+1} = 0$ and $\tilde{e}_{j,t+1} = \hat{e}_j(\theta_{t+1}, d_{j,t+1}^*, \tilde{d}_{j,t+1}, e_{j,t+1}^*)$ if $\mathbb{I}_{j,t+1} = 1$. For $k \notin \{i,j\}$, $\tilde{e}_{k,t+1} = e_{k,t+1}^*$. Otherwise, play is as in $\sigma^*|h_0^{t+1}$. At the end of period $t+$

1, agent j's output is transformed by $G_j(y_j|\theta_{t+1}, d^*_{j,t+1}, \tilde{d}_{j,t+1}, e^*_{j,t+1}, \tilde{e}_{j,t+1})$, and similarly for agent i if $\tilde{e}_{i,t+1} = \hat{e}_i$. If $\tilde{e}_{i,t+1} = e^*_{i,t+1}$, then output is transformed by the distribution R_i given in Condition 4 of Definition 3. Continuation play then proceeds as in σ^*.

We claim this perturbed strategy is an equilibrium, and that if $\epsilon > 0$ is sufficiently small, it generates strictly higher total surplus than σ^*. Because RE are recursive, play from $t+2$ onward is an equilibrium. The distribution over continuation play in $t+2$ is constructed to be identical to σ^*. In period $t+1$, a credible reward scheme for $\tilde{e}_{j,t+1}$ exists by the argument made in Claim 5. Similarly, a credible reward scheme exists for $\tilde{e}_{i,t+1} = \hat{e}_i$. If $\tilde{e}_{i,t+1} = e^*_{i,t+1}$, agent i's transformed distribution over output is identical to the output distribution in the original equilibrium for any $e_{i,t+1}$. Therefore, $e^*_{i,t+1}$ satisfies (7) under $\tilde{d}_{i,t+1}$ because it satisfied this inequality under $d^*_{i,t+1}$. We conclude that continuation play from period $t+1$ onwards is an equilibrium.

The change in total surplus in period $t+1$ from this perturbation equals

$$0 \geq K(\epsilon) = \begin{array}{l} s_i(\theta_{t+1}, \tilde{d}_{i,t+1}, \tilde{e}_{i,t+1}) + s_j(\theta_{t+1}, \tilde{d}_{j,t+1}, \tilde{e}_{j,t+1}) - \\ \left(s_i(\theta_{t+1}, d^*_{i,t+1}, e^*_{i,t+1}) + s_j(\theta_{t+1}, d^*_{j,t+1}, e^*_{j,t+1})\right) \end{array}.$$

This is the "direct cost" of backward-looking policies, which comes from the biased decision in period $t+1$. Importantly, $\tilde{e}_{j,t+1}$ equals the perturbed effort from the proof of Claim 5, while $\tilde{e}_{i,t+1}$ is weakly more efficient than the perturbed effort from Claim 5. Therefore, $K(\epsilon)$ is bounded from below by the left-hand side of (10). But then (9) implies that $\lim_{\epsilon \to 0} \frac{K(\epsilon)}{\epsilon} = 0$.

Now consider period t. Because $y^*_{j,t'} \leq y^*_j(\theta_{t'}, d_{j,t'}, e_{j,t'})$ for all $t' \leq t$, (7) implies that it is without loss to assume that the upper bound of (3) does not bind for agent j. The perturbation does not affect j's punishment payoff $\bar{U}_j(h_0^{t'})$ for $t' \leq t$, so agent j is willing to exert the same effort as in σ^*. Agents $k \notin \{i, j\}$ face the same distribution over $S_{k,t+1}$ and so are willing to choose the same efforts as well.

We claim that $E_{\tilde{\sigma}}\left[S_{i,t+1}|h_d^t, \xi_{i,t}, y_{i,t}\right]$ is strictly larger in the perturbed equilibrium relative to the original equilibrium. Holding $e_{i,t+1}$ fixed, $E_{\tilde{\sigma}}\left[S_{i,t+1}|h_d^t, \xi_{i,t}, y_{i,t}\right]$ is increasing in $d_{i,t+1}$ by Condition 3 of Definition 3. Furthermore, $\tilde{e}_{i,t+1}$ is weakly more efficient than $e^*_{i,t+1}$ by construction. Hence, $E_{\tilde{\sigma}}\left[S_{i,t+1}|h^t_d, \xi_{i,t}, y_{i,t}\right] > E_{\sigma^*}\left[S_{i,t+1}|h^t_d, \xi_{i,t}, y_{i,t}\right]$ as desired.

By assumption, $e^*_{i,t} < e^{FB}_i(\theta_t, d^*_{i,t})$. Consequently, (7) must hold with equality for agent i in period t; otherwise, we could increase $e^*_{i,t}$, transform output by the appropriate G_i, and increase i-dyad surplus in period t while continuing to satisfy (7). As a result, agent i is willing to exert strictly more effort in the perturbed equilibrium: $\tilde{e}_{i,t} > e^*_{i,t}$. Moreover, a straightforward but tedious application of the Implicit Function Theorem—similar to the proof of Claim 2—shows that the effort $\tilde{e}_{i,t}$ in the perturbed equilibrium is a function of ϵ, with $\frac{\partial \tilde{e}_{i,t}}{\partial \epsilon}|_{\epsilon=0} > 0$.

Consider the change in total surplus from period t onwards. As $\epsilon \to 0$, this change equals

$$\lim_{\epsilon \to 0}\left(\frac{s_i(\theta_t, d^*_{i,t}, \tilde{e}_{i,t}) - s_i(\theta_t, d^*_{i,t}, e^*_{i,t})}{\epsilon} + \frac{\delta K(\epsilon)}{\epsilon}\right) = \frac{\partial s_i}{\partial e_i}\frac{\partial \tilde{e}_i}{\partial \epsilon}|_{\epsilon=0} > 0.$$

The first term in this product is positive because $\lim_{\epsilon\to 0}\tilde{e}_{i,t-1} = e^*_{i,t-1} < e^{FB}_i(\theta_{t-1}, d_{i,t-1})$. The second term is positive by the argument above. Hence, this perturbation increases total continuation surplus in period $t-1$ onwards. It also increases i-dyad surplus, so there exists a credible reward scheme to support agent i's actions in periods $t' < t-1$ as well. We conclude that this perturbation is a self-enforcing relational contract that generates strictly higher total surplus than σ^*. ■

A.2.3 Proof of Statement 2

If $\sum_{i=1}^N d_{i,t} < 1$ at h^t_d, consider an 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 $i \in \{1, ..., N\}$. As in the proof of Statement 1, all agents can be induced to choose the same efforts given these decisions. Therefore, this alternative generates higher total surplus and relaxes (3) in all previous periods. But σ^* is surplus-maximizing; contradiction. ■

A.3 Proof of Corollary 1

In a smooth mean-shifting game, $\mathcal{E}_i = [\underline{e}_i, \bar{e}_i]$ for some $0 \leq \underline{e}_i < \bar{e}_i \leq 1$. Suppose continuation equilibrium $\sigma^*|h_0^t$ is surplus-maximizing at h_0^t. Claim 6 of Proposition 1 implies that decisions in period t must satisfy

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

for all $i, j \in \{1, ..., N\}$ and every θ_t. There exists a unique d_t^* that satisfies this condition because $\{\gamma_i\}_i$ are strictly concave.

Suppose σ^* is sequentially surplus-maximizing. Then by the above argument, d_t^* depends only on θ_t in each $t \geq 0$. Because on-path decisions are independent of observed play, it is straightforward to argue that equilibrium play in any sequentially surplus-maximizing equilibrium entails $e_{i,t} = e_i^*$ for each $t \geq 0$ and some $e_i^* \in [\underline{e}_i, e_i^{FB}]$. For $i \in \{1, ..., N\}$, define x_i^* as the unique value satisfying $\frac{\tilde{p}_i^H(x_i)}{\tilde{p}_i^L(x_i)} = 1$. From (7), e_i^* is defined implicitly by

$$c'(e_i^*) = \int_{-\infty}^{x_i^* + \gamma_i(d_t^*, \theta_t)} \bar{U}_i(\theta_t) \left[\tilde{p}_i^H(y_i) - \tilde{p}_i^L(y_i)\right] dy_i + \int_{x_i^* + \gamma_i(d_t^*, \theta_t)}^{\infty} S_i^* \left[\tilde{p}_i^H(y_i) - \tilde{p}_I^L(y_i)\right] dy_i,$$

where $S_i^* = E[y_i - c(e_i^*)|e_i^*]$ is a strictly concave function of e_i^*. Because $c'(\underline{e}_i) = 0$, $e_i^{FB} > \underline{e}_i$ and so there exist $\underline{\delta} < \bar{\delta}$ such that $e_i^* \in (0, e_i^{FB})$ for $\delta \in (\underline{\delta}, \bar{\delta})$. It immediately follows that e_i^* is a differentiable function of δ on this interval.

For $e_{i,t} = e_i^*$, $y_{i,t} - \gamma_i(\theta_t, d_t^*) > x_i^*$ with positive probability in each period t. Similarly, $y_{j,t'} < y_j^*(d_{j,t}^*, \theta_t, e_j^*)$ for all $t' \leq t$ with positive probability in each t. Therefore, the conditions of Proposition 1, part 1, hold for a set of histories E_t that occur with positive probability in each $t > 0$ in any sequentially surplus-maximizing equilibrium. Proposition 1 then implies that continuation play at these histories cannot be surplus-maximizing. So σ^* cannot be surplus-maximizing. ■

A.4 Proof of Proposition 2

Define $S^{R2} = \alpha R - c$, $S^{R1} = R - c$, 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.

Suppose $\theta_0 = R$. Define $\underline{\delta} \in (0,1)$ by $c = \frac{\underline{\delta}}{1-\underline{\delta}} S^{R2}$. Then for $\delta \geq \underline{\delta}$, Lemma 1 implies that there exists an equilibrium with $d_t = 2$ and $e_{i,t} = 1 \; \forall i \in \{1,2\}$ in each period. Any surplus-maximizing equilibrium therefore attains first-best.

If $\theta_0 = W$, then $d_0 = 1$ in any surplus-maximizing equilibrium. Suppose $d_0 = 2$: then either $e_{i,0} = 0$ for $i \in \{1,2\}$, in which case $d_0 = 1$ generates the same surplus, or $e_{i,0} = 1$ for at least one i, in which case $d_0 = 1$ generates strictly higher surplus. Similarly, in any period $t \geq 0$ with $\theta_t = W$, $d_t = 1$ both maximizes total continuation surplus and relaxes all prior binding dynamic enforcement constraints.

Define $\bar{\delta}$ as the solution to

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

Suppose $\delta \in [\underline{\delta}, \bar{\delta})$. Then in any equilibrium with $d_t = 2$ whenever $\theta_t = R$, $e_{1,t} = 0$ whenever $\theta_t = W$. Consider a relational contract of the form specified in Proposition 2, where $\chi > 0$ is chosen so that agent 1's constraint (3) holds with equality for $\theta_t = W$. For δ close to $\bar{\delta}$, it is straightforward to show that $\chi \approx 0$ and so this alternative dominates any equilibrium in which $d_t = 2$ whenever $\theta_t = R$.

It remains to show that an equilibrium of this form is surplus-maximizing. In any surplus-maximizing relational contract, agents work hard whenever they are hired. Therefore, once $\theta_t = R$, 1-dyad and total continuation surplus are linear functions of $\Pr\{d_{t'} = 1\}$ and $\Pr\{d_{t'} = 2\}$:

$$E[S_{1,t}|\theta_t = R] = \sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)\left(\Pr\{d_{t'} = 1\}(R-c) + \Pr\{d_{t'} = 2\}(\alpha R - c)\right)$$

and

$$E[S_{1,t}+S_{2,t}|\theta_t = R] = \sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)\left(\Pr\{d_{t'} = 1\}(R-c) + 2\Pr\{d_{t'} = 2\}(\alpha R - c)\right)$$

For any surplus-maximizing relational contract, construct a relational contract of the form described above by letting $\chi = \sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)\Pr\{d_{t'} = 1\}$. It is clear that total surplus is maximized if χ is chosen so that (3) binds, proving the claim. ■

A.5 Proof of Proposition 3

Define

$$S^B = \int_0^{\infty} y_i \tilde{p}^H(y_i) dy_i - c.$$

By assumption, for each $i \in \{1,2\}$ S^B equals i-dyad surplus if $e_{i,t} = 1$ and $d_t \neq i$ in each t. Let $\gamma^i = \gamma_i(i)$. From period $t = 1$ onwards, an equilibrium exists in which agent i exerts effort if and only if

$$c \leq \frac{\delta}{1-\delta} \int_{y^*}^{\infty} \left(S^B + 1\{d_t = i\}\gamma^i\right) \left[\tilde{p}^H(y_i^*) - \tilde{p}^L(y_i^*)\right] dy_i. \tag{11}$$

Let $\underline{\delta}$ satisfy (11) with equality if $d_t = i = 2$. Let $\bar{\delta}$ satisfy (11) with equality if $d_t \neq i$. Then $\bar{\delta} > \underline{\delta}$.

If $\delta \in [\underline{\delta}, \bar{\delta})$, then (11) holds for agent i only if $d_t = i$, so $e_{i,t} = 0$ for $t \geq 1$ if $d_t \neq i$. Consider effort choices in $t = 0$. It is straightforward to show that in any surplus-maximizing equilibrium, either both agents exert effort in $t = 0$, or only agent 1 exerts effort in $t = 0$.

If only agent 1 exerts effort in $t = 0$, then $d_1 = 1$ with probability 1. If both agents exert effort in $t = 0$, then $d_1 = 2$ with positive probability because $\delta < \bar{\delta}$. Following output $y_0 \in \mathbb{R}^2$ in $t = 0$, let $\rho(y_0)$ denote the probability that agent 1 is chosen in $t = 1$. Then Lemma 1 implies that the surplus-maximizing equilibrium must maximize the probability of $\rho(y) = 1$, conditional on motivating both agents to work hard:

$$\max_{\rho:\mathbb{R}^2\to[0,1]} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(y)\tilde{p}^H(y_1)\tilde{p}^H(y_2)dy_1dy_2$$

subject to (2) and (3). Given $\delta \in [\underline{\delta}, \bar{\delta})$, these two sets of constraints may be combined as:

$$c \leq \tfrac{\delta}{1-\delta}\int_{-\infty}^{\infty}\int_{y^*}^{\infty} \rho(y)\left[S^B+\gamma^1\right]\left[\tilde{p}^H(y_1)-\tilde{p}^L(y_1)\right]\tilde{p}^H(y_2)dy_1dy_2$$
$$c \leq \tfrac{\delta}{1-\delta}\int_{-\infty}^{\infty}\int_{y^*}^{\infty}(1-\rho(y))\left[S^B+\gamma^2\right]\left[\tilde{p}^H(y_2)-\tilde{p}^L(y_2)\right]\tilde{p}^H(y_1)dy_2dy_1$$

for agents 1 and 2, respectively.

The Lagrangian for this constrained optimization problem may be solved separately for each y. If $L(y_2) < 0$, then $\rho(y) = 1$. Otherwise, the derivative of this Lagrangian with respect to $\rho(y)$ equals

$$1+\lambda_1(S^B+\gamma^1)\left(1-\frac{1}{L(y_1)}\right)-\lambda_2(S^B+\gamma^2)\left(1-\frac{1}{L(y_2)}\right)$$

with λ_i the multiplier associated with the constraint for agent i. This expression is independent of ρ, so $\rho(y) = 0$ whenever this expression is strictly negative. Rearranging yields (5).

If $\Delta < \frac{1-\delta}{\delta}S^B$, the equilibrium in which both agents work hard in $t = 0$ dominates the equilibrium in which only agent 1 works hard. This proves the claim. ■

A.6 Proof of Lemma 2

We first prove an extension of Lemma 1 to PBE.

Definition: A reward scheme $B_i : \phi_i(\mathcal{H}_d^t) \times \Xi_i \times \mathbb{R} \to \mathbb{R}$ is **PBE-credible in** σ if:

1. For each h_d^t, $\xi_{i,t}$, and $(a_{i,t}, e_{i,t})$ on the equilibrium path,

$$(a_{i,t}, e_{i,t}) \in \arg\max_{a_i, e_i} E_\sigma \left[B_i(\phi_i(h_d^t), \xi_{i,t}, y_{i,t}) | \phi_i(h_d^t), \xi_{i,t}, a_i, e_i \right] - (1-\delta) C_i. \tag{12}$$

2. For each on-path h_y^t,

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

A.6.1 Claim 1

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^*.[13]

2. Suppose σ is a strategy with a PBE-credible reward scheme B_i for $i \in \{1, ..., N\}$. Then $\exists$ PBE σ^* with the same joint distribution over θ_t, d_t, e_t, and y_t as σ.

A.6.2 Proof of Claim 1

This proof is extended from Andrews and Barron (ming), who provide more detail.

Part 1. This argument is nearly identical to Lemma 1, part 1. Suppose σ^* is a PBE and define B_i by

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

Then B_i must satisfy (12) and the first inequality of (13) or else the agent would deviate from $(a_{i,t}, e_{i,t})$ or $\tau_{i,t}$, respectively. The second inequality of

[13] Every PBE in this game is payoff-equivalent to a PBE in which players do not condition on past effort choices. The proof of this result is similar to Fudenberg and Levine (1994), who prove a similar result for games with imperfect public monitoring and a product monitoring structure.

(13) must hold history-by-history or else the principal would deviate from $\tau_{i,t}$, so *a fortiori* must hold in expectation. $\square$

Part 2. Consider the construction identical to Lemma 1, part 2, except that

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

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

and the transfer after output y_t^* equals

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

By construction, σ^* implements the same joint distribution over θ_t, d_t, e_t, and y_t as σ. We claim 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 the principal has no deviation from σ^*.

Consider the possible deviations by agent i. Agent i earns $\bar{U}_i(h_0^{t+1})$ if he deviates in period t. Agent i is willing to choose $(a_{i,t}, e_{i,t})$ if

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

As in Lemma 1, $E_{\sigma^*}\left[U_{i,t+1}|\phi_i(h_d^{t,*}), a_{i,t}, e_{i,t}\right] = E_{\sigma^*}\left[S_{i,t+1}|\phi_i(h_d^{t,*}), a_{i,t}, e_{i,t}\right]$. Furthermore, it can be shown that for every agent i, σ^* induces a coarser information partition over histories 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 $\phi_i(h_0^t) = \phi_i(\tilde{h}_0^t)$. Therefore, $E_{\sigma^*}\left[S_{i,t+1}|\phi_i(h_d^{t,*}), a_{i,t}, e_{i,t}\right] = E_\sigma\left[S_{i,t+1}|\phi_i(h_d^t), a_{i,t}, e_{i,t}\right]$. Plugging these expressions into agent i's IC constraint yields (12).

Agent i is willing to pay $\tau_{i,t}^*$ if

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

This constraint is satisfied because (13) holds. So σ^* is the desired PBE. $\blacksquare$

A.6.3 Completing Proof of Lemma 2

($\rightarrow$) If $E_{\sigma^*}\left[\sum_{t'=t}^{\infty}\delta^{t'-t}(1-\delta)(\pi_{t'}+\sum_{i=1}^{N}u_{i,t'})\right]=\bar{V}$, consider the strategy $\tilde{\sigma}$ in which the principal chooses h_0^t from the distribution over $\mathcal{H}_0^t$ induced by σ^*, then play continues as in $\sigma^*|h_0^t$. By construction, players have the same beliefs in $\tilde{\sigma}$ and $\sigma^*|h_0^t$, so $\tilde{\sigma}$ is an equilibrium that generates total surplus V.

($\leftarrow$) Suppose σ^* satisfies $E_{\sigma^*}\left[\sum_{t'=0}^{\infty}\delta^{t'}(1-\delta)(\pi_{t'}+\sum_{i=1}^{N}u_{i,t'})\right]=\bar{V}$. Consider strategy $\tilde{\sigma}$ in which the static equilibrium is played in all periods $t'<t$, then play σ^* from period t onwards. This is clearly an equilibrium that attains continuation surplus $\bar{V}$ from period $t>0$ onwards. ■

A.7 Proof of Proposition 4

Let σ^* be a surplus-maximizing PBE. As in Corollary 1, in any PBE-sequentially surplus-maximizing equilibrium, decisions and efforts depend only on the payoff-relevant history. But this history is publicly observed, so any PBE-sequentially surplus-maximizing PBE must be payoff-equivalent to an RE. It is straightforward to show that in that case, the surplus-maximizing RE is sequentially surplus-maximizing. So if no surplus-maximizing RE is sequentially surplus-maximizing, then no surplus-maximizing PBE is PBE-sequentially surplus-maximizing. ■

A.8 Proof of Proposition 5

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

Statement of Claim 1

If σ^* is a RE, then $\forall i\in\{1,...,N\}$ there exists a function $B_i:\phi_0(\mathcal{H}_y^t)\rightarrow\mathbb{R}$ satisfying:

1. **Public Effort IC:** for any $i\in\{1,...,N\}$ and h_e^t,

$$(a_{i,t},e_{i,t})\in\arg\max_{a_i,e_i}E_{\sigma^*}\left[B_i(\phi_0(h_y^t))-(1-\delta)C_i|h_a^t,e_i\right].\qquad(14)$$

2. **Public Dynamic Enforcement:** for any $I \subseteq \{1, ..., N\}$ and h_y^t,

$$\delta \sum_{i \in I} E_{\sigma^*} \left[\bar{U}_i(h_0^{t+1}) | h_y^t \right] \leq \sum_{i \in I} B_i(\phi_0(h_y^t)) \leq \delta E_{\sigma^*} \left[\sum_{i \in I} U_{i,t+1} + \Pi_{t+1} | h_y^t \right]. \tag{15}$$

3. **Individual Rationality:** for any $h_d^t \in \mathcal{H}_d^t$ and every agent $j \in \{1, ..., N\}$,

$$E_{\sigma^*}[U_{j,t+1} | h_d^t] \geq \bar{U}_j(h_d^t). \tag{16}$$

For every subset of agents $I \subseteq \{1, ..., N\}$,

$$E_{\sigma^*} \left[\Pi_{t+1} | h_d^t \right] \geq \sum_{i \in I} \left(E_{\sigma^*} \left[B_i(\phi_0(h_y^t)) - (1-\delta) C_{i,t} | h_d^t \right] - E_{\sigma^*} \left[U_{i,t} | h_d^t \right] \right). \tag{17}$$

Proof of Claim 1

Suppose σ^* is a RE. Define B_i by

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

Analogous to Lemma 1, agent i chooses $e_{i,t}$ to solve (14). Agent i's continuation surplus is bounded below by $\bar{U}_i(h_0^{t+1})$ in h_0^{t+1}, so $B_i(\phi_0(h_y^t)) \geq E \left[\bar{U}_i(h_0^{t+1}) | h_y^t \right]$. If $\exists I \subseteq \{1, ..., N\}$ such that

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

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. These arguments imply (15).

If $w_{i,t} < 0$, then agent i is willing to pay only if $E[U_{i,t} | h_d^t] \geq \bar{U}_i(h_d^t)$. Let $I = \{i | E_{\sigma^*}[w_{i,t} | h_d^t] \leq 0\}$. Then the principal is willing to pay $\sum_{i \notin I} w_{i,t} > 0$

only if

$$E_{\sigma^*}\left[(1-\delta)\left(\sum_{i=1}^{N} y_{i,t}-\sum_{i\notin I} w_{i,t}\right)-\sum_{i=1}^{N}\left(B_i(\phi_0(h_y^t))-\delta U_{i,t+1}\right)+\delta\Pi_{t+1}|h_d^t\right]\geq 0.$$

Rewriting this expression in terms of $U_{i,t}$ and Π_t yields

$$E_{\sigma^*}\left[\Pi_t|h_d^t\right]\geq\sum_{i\in I}E_{\sigma^*}\left[B_i(\phi_0(h_y^t))-(1-\delta)C_{i,t}-\delta U_{i,t}|h_d^t\right].$$

This expression holds *a fortiori* for any other set of agents. These arguments together imply (16) and (17). ■

Completing Proof of Proposition 5

Suppose σ is a surplus-maximizing RE that is not sequentially surplus-maximizing. Consider a strategy profile $\tilde{\sigma}$ that is identical to σ except for wages which satisfy $E[U_{i,t}|h_d^t]=\bar{U}_i(h_d^t)$. Then it is easy to show $\tilde{\sigma}$ satisfies (14) for the same B_i as σ. $\tilde{\sigma}$ satisfies (17) because $E_{\tilde{\sigma}}\left[\Pi_{t+1}+\sum_{i\in I}\bar{U}_i(h_d^t)|\phi_0(h_y^t)\right]\geq E_{\sigma}\left[\Pi_{t+1}+\sum_{i\in I}U_{i,t+1}|\phi_0(h_y^t)\right]$.

The strategies σ and $\tilde{\sigma}$ generate the same ex ante total surplus, and moreover there exists some history h_0^t such that $\tilde{\sigma}|h_0^t$ is not surplus-maximizing. Consider an alternative strategy $\tilde{\sigma}^*$ that is identical to $\tilde{\sigma}$, except $\tilde{\sigma}^*|h_0^t$ is surplus-maximizing and holds all agents at their outside options. It is easy to see that $\tilde{\sigma}^*$ satisfies (14)-(17) because $\tilde{\sigma}$ does, and $\tilde{\sigma}^*$ generates strictly higher total continuation surplus than $\tilde{\sigma}$. Thus, it suffices to show that the policy and efforts in $\tilde{\sigma}^*$ are part of an equilibrium.

Consider the following strategies σ^*, defined recursively from $\tilde{\sigma}^*$. For histories $\tilde{h}_0^t, h_0^{t,*}\in\mathcal{H}_0^t$, use the public randomization device to choose $\tilde{h}_d^t\in\mathcal{H}_d^t$ according to $\tilde{\sigma}^*|\{\tilde{h}_0^t,\theta_t,D_t\}$. The principal chooses $d_t\in D_t$ as in $\tilde{h}_d^t$. For each agent i, the wage is $w_{i,t}=E_{\tilde{\sigma}^*}\left[-\tau_{i,t}^*+C_{i,t}+\frac{1}{1-\delta}\bar{U}_i(\tilde{h}_d^t)-\frac{\delta}{1-\delta}\bar{U}_i(\tilde{h}_d^{t+1})|\tilde{h}_d^t\right]$, with $\tau_{i,t}^*$ defined below. The public randomization device chooses $\tilde{h}_e^t\in\mathcal{H}_e^t$ as in $\tilde{\sigma}^*|\tilde{h}_d^t$. Agent i chooses $a_{i,t},e_{i,t}$ as in $\tilde{h}_e^t$. Following output y_t, agent i's bonus equals $\tau_{i,t}^*=\frac{1}{1-\delta}E_{\tilde{\sigma}^*}\left[B_i(\phi_0(\tilde{h}_y^t))-\bar{U}_i(h_0^{t+1})|\tilde{h}_e^t,y_t\right]$. History $\tilde{h}_0^{t+1}$ is drawn by the public randomization device according to $\tilde{\sigma}^*|(\tilde{h}_e^t, y_t)$. This process is repeated with $\tilde{h}_0^{t+1}$. Following a deviation by agent j, $a_{j,t'} = 0$ and $w_{j,t'} = \tau_{j,t'} = 0$ in all $t' \geq t$, and the principal chooses $d_{t'}$ to hold agent i at $\bar{U}_i(h_0^t)$. Following any other deviation, play as if agent 1 deviated.

We claim σ^* is a recursive equilibrium. Indeed, it is straightforward to show that agent i earns $\bar{U}_i(h_0^t)$ at each h_0^t. The principal is willing to pay $w_{i,t} \geq 0$, or the agent is willing to pay $w_{i,t} \leq 0$, because $\tilde{\sigma}^*$ satisfies (16) and (17). Each agent i is willing to choose $a_{i,t}$ and $e_{i,t}$ because $\tilde{\sigma}^*$ satisfies (14). And the principal is willing to pay $\tau_{i,t}^*$ because $\tilde{\sigma}^*$ satisfies (15). Furthermore, σ^* generates the same total ex ante expected surplus as $\tilde{\sigma}^*$, and so generates strictly higher ex ante expected surplus than σ. So σ^* cannot be surplus-maximizing. ■

A.9 Proof of Proposition 6

Given equilibrium σ^*, define $B_i(h_d^t, \xi_{i,t}, y_{i,t}) = E_{\sigma^*}\left[(1-\delta)\tau_{i,t} + \delta U_{i,t+1} | h_d^t, \xi_{i,t}, y_{i,t}\right]$ as in Lemma 1. Then $B_i(h_d^t, \xi_{i,t}, y_{i,t}) \geq 0$. Consider a deviation in the principal's relationship with agent i. If agent i chooses his outside option, the principal earns her minimum payoff 0 in that period. This choice is publicly observed with probability $1-\epsilon$, in which case the principal earns 0 continuation surplus. Otherwise, the principal loses $\Pi^i \equiv \sum_{t'=1}^{\infty} \delta^{t'}(1-\delta)(y_{i,t+t'} - w_{i,t+t'} - \tau_{i,t+t'})$ by an argument similar to Lemma 1. So in any equilibrium,

$$B_i(h_d^t, \xi_{i,t}, y_{i,t}) \leq \frac{\delta}{1-\delta} E\left[(1-\epsilon)\sum_{j \neq i} S_j + S_i | h_d^t, \xi_{i,t}, y_{i,t}\right].$$

Define $\tilde{S}^{R1} = R - c$, $\tilde{S}^{R2} = (2-\epsilon)(\alpha R - c)$, $\tilde{S}^{W1} = (1-\delta)(W-c) + \delta(\rho \tilde{S}^{R1} + (1-\rho)\tilde{S}^{W1})$, and $\tilde{S}^{W2} = (1-\delta)(W-c) + \delta(\rho \tilde{S}^{R2} + (1-\rho)\tilde{S}^{W2})$. Suppose the principal deviates in period t, when $\theta_t = \theta$. Then $\tilde{S}^{\theta d}$ equals the expected surplus destroyed following a deviation if $d_t = d$ whenever $\theta_t = R$ on the equilibrium path. We make assumptions such that (i) the principal cannot motivate agent 1 while $\theta_t = W$ if $d_t = 2$ whenever $\theta_t = R$, but can motivate agent 1 if $d_t = 1$ whenever $\theta_t = R$; and (ii) conditional on high effort, $d_t = 2$

is surplus-maximizing if $\theta_t = R$, $d_t = 1$ is surplus-maximizing if $\theta_t = W$, and more surplus is lost following a deviation if $d_t = 1$ in every subsequent period than if $d_t = 2$.

$$\tilde{S}^{W2} < \tfrac{1-\delta}{\delta} c \leq \min \left\{ \alpha R - c, \tilde{S}^{W1} \right\},$$
$$2(\alpha R - c) > \tilde{S}^{R1} > \tilde{S}^{R2} > W - c > 2(\alpha W - c).$$

For $\epsilon > 0$, there exists an open set of parameters that simultaneously satisfy these conditions.

Suppose that the only constraints in equilibrium are (2) and that agent i's reward scheme must satisfy

$$0 \leq B_i(h_d^t, \xi_{i,t}, y_{i,t}) \leq \frac{\delta}{1-\delta} E\left[(1-\epsilon) \sum_{j \neq i} S_j + S_i | h_d^t, \xi_{i,t}, y_{i,t} \right].$$

By the first assumption, there exists a reward scheme such that $e_{1,t} = e_{2,t} = 1$ if $\theta_t = R$ and $d_t = 2$. Therefore, any sequentially surplus-maximizing equilibrium must have $d_t = 2$ whenever $\theta_t = R$. But the first assumption also implies that $e_{1,t} = 0$ whenever $\theta_t = W$ if $d_t = 2$ whenever $\theta_t = R$. So agent 1 does not exert effort while $\theta_t = W$ in any sequentially surplus-maximizing equilibrium.

Consider the alternative strategy described in the proof of Proposition 2, with $\chi \in (0,1)$ chosen to solve $c = \frac{\delta}{1-\delta}(\chi \tilde{S}^{W1} + (1-\chi)\tilde{S}^{W2})$. By construction, all hired agents can be motivated to choose $e_{i,t} = 1$ in each t under this strategy. So surplus in this alternative is $W - c$ in each period with $\theta_t = W$. Once $\theta_t = R$, surplus equals $2(\alpha R - c)$ with probability χ and otherwise equals $R - c$. We can choose parameters such that χ is arbitrarily close to 0, in which case this alternative generates strictly higher total surplus than any sequentially surplus-maximizing equilibrium.

The final step is to prove that this alternative strategy is in fact an equilibrium. Both θ_t and the public randomization device are publicly observed, and the proposed d_t conditions only on these variable. Hence, both agents detect any deviation in d_t and so the principal earns 0 following such a deviation. Therefore, the principal has no profitable deviation in d_t. Each agent is paid

$w_{i,t} = 0$. The principal pays $\tau_{i,t} = c$ if she hires agent i and otherwise pays $\tau_{i,t} = 0$. Following a deviation in $\tau_{i,t}$, the principal earns 0 with probability $1-\epsilon$ or loses i-dyad surplus with probability ϵ. By choice of χ, the principal is indifferent between paying $\tau_{i,t}$ or not. Agents have no profitable deviation from $e_{i,t}$ or $a_{i,t}$, so this is an equilibrium. Moreover, this equilibrium dominates any sequentially surplus-maximizing equilibrium for an open set of parameters. ■