Policies in Relational Contracts

Daniel Barron and Michael Powell*

February 1, 2018

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 agents cannot coordinate to punish the principal following a deviation, 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 illustrate how they influence promotions, hiring, and performance.

*Barron: Northwestern University Kellogg School of Management, Evanston IL 60208; email: d-barron@kellogg.northwestern.edu. Powell: Northwestern University Kellogg School of Management, Evanston IL 60208; email: mike-powell@kellogg.northwestern.edu. We thank Johannes Hörner and two anonymous referees for their comments. The authors would like to thank Nageeb Ali, Joyee Deb, Willie Fuchs, John Geanakoplos, Bob Gibbons, Yingni Guo, Marina Halac, Jin Li, Elliot Lipnowski, Jim Malcomson, David Miller, Niko Matouschek, Luis Rayo, Larry Samuelson, Takuo Sugaya, Michael Waldman, 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 their partners' efforts. 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 contributes to the firm's objectives. Supervisors assign tasks to team members. Executives choose which subordinates to promote. Firms allocate capital to divisions. Human-resource managers hire and fire employees. These decisions make some individuals more integral and others less integral to the firm. And importantly, these decisions are often made on the basis of past performance. Supervisors promote those who have performed well, even if they would not make the best managers (Benson et al. (2017)). CFOs allocate scarce capital to divisions that have seen past success, even when other divisions have higher net-present value projects (Graham et al. (2015)). Firms delay efficient expansions that would entail hiring new workers who may be substitutes for their existing employees (Ariely et al. (2013)).

As these examples illustrate, biasing decisions towards an individual involves real costs—costs that could be avoided if the principal instead rewarded past success with money alone (Baker et al. (1988)). Why, then, do biased decisions arise? In this paper, we argue that if a principal cannot commit to a formal contract, then she may bias decisions towards an agent to make informal monetary rewards to that agent credible. Our basic argument is that, while biasing decisions decreases total surplus, it also increases the surplus produced by favored agents. An agent who produces a lot of future surplus can threaten to withhold that surplus to punish the principal following a deviation. The more surplus is destroyed by a punishment, the more willing is the principal to pay an agent rather than renege and face that punishment. Consequently, a principal who promises to bias future decisions towards an agent can also credibly promise larger monetary rewards to that agent.

Section 2 introduces this argument in the context of promotion decisions. Suppose a principal repeatedly employs two agents, each of whom privately chooses a binary effort that determines the distribution over his output. After one period of production, the principal makes a once-and-for-all decision to promote one of the agents, which increases that agent's expected productivity in all subsequent periods. Players have deep pockets, so the principal can motivate the agents by paying bonuses or demanding fines. However, formal contracts are not available, so the principal must be willing to follow through on both payments and promotion decisions.

We show that promotions complement monetary incentives if agents cannot coordinate to punish the principal following a deviation. Suppose that, if the principal reneges on a promised payment, then only the betrayed agent punishes her. In that case, an agent who has been promoted can punish the principal more severely. Hence, the principal is willing to pay a promoted agent a larger reward in equilibrium, which potentially motivates that agent to work harder. This argument suggests that a principal might optimally promote an agent who performs well, even if it is common knowledge that the other agent would make a better manager.

Consistent with this intuition, we show that the surplus-maximizing promotion policy takes the form of a tournament in which the agent who would be a worse manager might nevertheless be promoted. The principal can be made indifferent to her promotion decision so that she is willing to promote the less able agent even though she cannot formally commit to do so.[1] Biased promotions arise even though all parties have deep pockets, so this example provides an answer to Baker et al. (1988)'s puzzle of why firms use promotions in addition to pay to motivate their employees. It also rationalizes the finding in Benson et al. (2017) that some firms systematically promote high-performing employees over peers who are more likely to be good managers based on *ex ante* observable characteristics.

[1] As we discuss in Section 5 and Appendix F, indifference is not required for our basic intuition. Also see the example in Appendix E.

In Section 3, we extend this logic to a flexible model of policies in relational contracts with multiple agents. The key feature of our framework is that the principal makes a public decision in each period that influences how agents' efforts affect the firm's output. This decision simultaneously affects every agent's production, so a decision that makes one agent essential might make another expendable. This section has two objectives: first, it builds tools that can be used in many settings to explore how relational contracts impact optimal policies, and second, it clarifies the underlying tradeoffs that lead to biased decisions.

As the promotion example highlights, a key assumption in our framework is that if the principal reneges on a payment to an agent, then only that agent punishes her. To provide a foundation for these uncoordinated punishments, we assume that agents do not observe one another's output or pay and cannot communicate with one another. The resulting game has imperfect private monitoring, which implies that standard equilibrium concepts are not recursive. We develop a refinement of Perfect Bayesian Equilibrium, called recursive equilibrium, which provides a rigorous but tractable way to model uncoordinated punishments in relational contracts.[2]

The first step of the general analysis derives necessary and sufficient conditions for a recursive equilibrium. These conditions take the form of incentive constraints, which ensure that each agent is willing to exert the desired effort, and dynamic enforcement constraints, which bound each agent's equilibrium reward following each output. The upper bound of an agent's dynamic enforcement constraint depends on the future surplus produced by that agent, which depends in turn on the principal's future decisions. As in the promotions example, the principal can be made indifferent among her equilibrium decisions, so that she is willing to implement decisions even if they do not maximize total continuation surplus.

These conditions highlight the costs and benefits of biased decisions. By definition, biased decisions have a *direct cost* because they do not maximize

[2] Appendix C extends our logic to the full (non-recursive) set of Perfect Bayesian Equilibria.

total continuation surplus. Biased decisions also make some agents expendable, which has an *incentive cost* because the principal cannot credibly promise large rewards to motivate expendable agents. However, biased decisions make other agents essential, which has an *incentive benefit* because those agents can be promised large rewards that potentially motivate them to work hard. Both incentive costs and incentive benefits depend on past performance, so optimal decisions do too. We identify a class of games for which these costs and benefits vary smoothly in the principal's decisions. For these games, we show that biased decisions arise in any surplus-maximizing relational contract if players are neither too patient nor too impatient.

In Section 4, we apply this framework to a simple model of hiring decisions. We consider a firm that faces stochastically and persistently increasing binary demand and decides whether to hire one or two employees in each period. Total continuation surplus is maximized by hiring more employees as demand increases. However, employees are substitutes in the sense that each additional employee decreases the expected output of the others. Consequently, we argue that the firm might optimally delay expanding in a surplus-maximizing relational contract, since keeping a small workforce makes strong rewards to early employees credible in equilibrium.

Section 5 discusses several extensions. We show that if agents could perfectly coordinate punishments, either by jointly observing a deviation or by communicating with one another, then biased decisions would never be optimal. However, coordination must be perfect; in a simple example, we show that biased decisions can be optimal if coordination is imperfect. We also extend our basic mechanism to the full set of Perfect Bayesian Equilibrium, and we explore optimally biased decisions if the principal must be given strict incentives to follow her equilibrium decision.

We view the assumption of uncoordinated punishments as plausible in many settings. For example, Bewley (1999) offers suggestive evidence that layoffs do not reduce productivity among remaining workers. More generally, our intuition requires only that the principal who betrays a single worker is not punished by her entire workforce. For instance, consider a firm with several plants, and suppose that if that firm betrays a worker, then it is punished by others at the same plant but not by workers at other plants. In our framework, we could treat the plants as different "agents," in which case surplus-maximizing relational contracts might entail decisions that inefficiently favor one plant over the others. Consistent with this interpretation, both Krueger and Mas (2004) and Mas (2008) provide evidence that workers at different factories do not coordinate punishments during periods of labor unrest.

Related Literature: 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 optimal relational contracts are stationary. In contrast, we focus on history-dependent inefficiencies. Our paper is therefore related to Fudenberg et al. (1990), which develops conditions under which an optimal *formal* contract may exhibit history-dependent inefficiencies. The contracting frictions highlighted by that paper—including limited liability and asymmetric information—have spurred a substantial literature in both formal and relational contracts.

We contribute to this literature by highlighting a mechanism by which history-dependent inefficiencies arise in settings without commitment, even if players have transferable utility and symmetric information about the continuation game. Andrews and Barron (2016) studies the same mechanism; our Lemma 1 generalizes Lemma 1 from that paper. Relative to Andrews and Barron (2016), we contribute in two ways. First, we consider sequential inefficiencies, while Andrews and Barron (2016) restricts attention to parameters under which the relational contract attains first-best. Therefore, we can address the apparently inefficient biases that arise in organizations and highlight how surplus-maximizing relational contracts optimally balance the costs and benefits of biased decisions. Second, Andrews and Barron (2016) considers allocation decisions in a supply chain, while we build more flexible tools for analyzing biased decisions. In particular, we argue that many superficially dissimilar distortions can be understood as different manifestations of the same fundamental need to make promised incentives credible.

A growing literature within relational contracting, surveyed by Malcomson (2013), focuses on other sources of history-dependent inefficiencies, including asymmetric information (Halac (2012); Malcomson (2016)), learning (Watson (1999, 2002)), subjective evaluations and other forms of private monitoring between a principal and a single agent (Levin (2003); Fuchs (2007); Fong and Li (2017a)), or limited transfers (Board (2011); Fong and Li (2017b); Li et al. (2017); Lipnowski and Ramos (2017)).[3] Among these papers, Board (2011), which studies how limited transfers can lead to allocation dynamics in a supply chain, is related in terms of application. History-dependent inefficiencies arise in that paper because agents cannot pay the principal, and they exist even if the principal can commit to a formal contract.

We emphasize a distinct friction that arises even if parties have deep pockets. Our model suggests that biases are likely to be worse when formal contracts are difficult to write and agents have trouble coordinating punishments. In addition, our results make predictions about the *types* of sequential inefficiencies that are likely to arise. In our setting, biased decisions are valuable only if at least one agent produces high output, and so are unlikely to arise when all agents perform poorly. Our mechanism also predicts that biased decisions complement higher pay, so compensation should be positively correlated with biased decisions conditional on performance. Section 4 suggests that biases should arise in hiring decisions, which implies that firm size and productivity dynamics might be history dependent. Finally, biased decisions are always biased *towards* some agents at the expense of others; they never arise in bilateral relationships, nor do they entail simultaneously damaging every relationship. Consequently, our setting requires the principal to interact with multiple agents, though it differs from moral-hazard-in-teams settings (Holmstrom (1982); Rayo (2007)) because each agent produces a separate output.

More broadly, a substantial literature has studied cooperation in settings with imperfectly coordinated punishments, for example in settings with communal enforcement (Kandori (1992); Ellison (1993); Ali and Miller (2016)).

[3] Many repeated games with public monitoring rule out transfers entirely, leading to equilibrium dynamics. See Goldlucke and Kranz (2012).

While we also assume that agents cannot coordinate punishments, we allow transfers and focus on how the principal's decisions constrain those transfers in equilibrium. Ali et al. (2016) rules out certain kinds of coordinated punishments by imposing bilateral renegotiation-proofness, but shows that this refinement is not enough to generate sequential inefficiencies if utility is transferable.

Finally, our framework is related to papers that study how organizations should optimally make decisions, including how to allocate decision rights (Aghion and Tirole (1997); Dessein (2002)), assign tasks and promote employees (see Waldman (2013) for a survey), allocate capital (see Gertner and Scharfstein (2013) for a survey), and design hiring, firing, and skill-development policies (which Lazear and Oyer (2013) argues is an understudied set of issues). Our framework suggests that relational considerations might lead to dynamic inefficiencies in these (and other) decisions.

2 Biased Promotions: An Example

We introduce the key ideas of our analysis using a simple model of promotions. The intuition in this section is self-contained, though some formal proofs draw on later results.

Consider a principal who repeatedly interacts with two agents in periods $t = 0, 1, ...$ Players share a common discount factor δ. In each period, each agent privately exerts binary effort that determines the distribution over that agent's output. Parties are risk-neutral and have deep pockets, so the principal can pay (or be paid by) each agent both before effort is chosen and after output is realized. We make two key assumptions. First, at the start of the second period ($t = 1$), the principal can decide to promote exactly one of the two agents, which increases that agent's expected output in every subsequent period. Second, agents do not observe one another's output or pay.

The timing of the stage game is:

1. In $t = 1$, the principal publicly chooses one agent to promote: $d_1 \in \{1, 2\}$.

2. The principal and each agent $i \in \{1,2\}$ 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]

3. Each agent $i \in \{1,2\}$ chooses to participate or not, $a_{i,t} \in \{0,1\}$. Only the principal and agent i observe $a_{i,t}$.

4. Each agent $i \in \{1,2\}$ privately chooses effort $e_{i,t} \in \{0,1\}$ at cost $ce_{i,t}$.

5. For $i \in \{1,2\}$, output $y_{i,t} \in \mathbb{R}$ is realized, with $y_t = (y_{1,t}, y_{2,t})$. If $a_{i,t} = 0$, then $y_{i,t} = 0$. If $a_{i,t} = 1$, then $y_{i,t}$ has distribution $P(\cdot|e_{i,t})$ with density $p(\cdot|e_{i,t})$, where $\frac{p(\cdot|1)}{p(\cdot|0)}$ is strictly increasing and $E[y_{i,t}|e_{i,t} = 1] - c > E[y_{i,t}|e_{i,t} = 0] \geq 0$. Only the principal and agent i observe $y_{i,t}$.

6. The principal and each agent $i \in \{1,2\}$ make non-negative transfers to each other. Define $\tau_{i,t} \in \mathbb{R}$ as the net bonus paid to agent i. Only the principal and agent i observe $\tau_{i,t}$.

Let $1_{i,t} \equiv 0$ if either $t = 0$ or $d_1 \neq i$, and $1_{i,t} \equiv 1$ otherwise. In period t, the principal's and agent i's payoffs are

$$\begin{aligned} \pi_t &= (1-\delta) \textstyle\sum_{i=1}^{2} \left(y_{i,t} + a_{i,t} 1_{i,t} \gamma_i - w_{i,t} - \tau_{i,t} \right) \\ u_{i,t} &= (1-\delta)(w_{i,t} + \tau_{i,t} - c a_{i,t} e_{i,t}), \end{aligned}$$

respectively, where γ_i is the extra output produced by agent i if he is promoted. We assume $\gamma_1 > \gamma_2 > 0$ so that, holding efforts fixed, total surplus would be maximized by promoting agent 1. Define $U_{i,t}$ as agent i's normalized discounted continuation payoff at the start of period t, and define **i-dyad surplus** as the total continuation surplus produced by agent i,

$$S_{i,t} = \sum_{t'=t}^{\infty} \delta^{t'-t} (1-\delta) \left(y_{i,t} + a_{i,t} 1_{i,t} \gamma_i - c a_{i,t} e_{i,t} \right).$$

[4] In Section 3, we allow the principal to send messages to each agent during this stage. These messages are unnecessary in this example, so we omit them.

This model assumes that agents do not observe one another's actions, outputs, or payments, and cannot communicate with one another. While this assumption is stylized, we believe that it provides a foundation for an important feature of many business relationships: widespread punishments are difficult to coordinate, especially when 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. However, the other agents do not follow suit, because they do not observe the deviation. Section 5 revisits these monitoring assumptions.

Our basic solution concept is Perfect Bayesian Equilibrium (PBE), specifically Watson (2016)'s definition of *plain PBE*. In a PBE of a game with private monitoring, players might have different beliefs about the true history, which implies that the set of continuation payoffs in a PBE might depend on the history. To avoid this complication, our results consider recursive equilibria, which are recursive and hence a relatively tractable refinement of PBE. Let h_0^t be a history at the start of period t and $\mathcal{H}_0^t$ be the set of such histories.

Definition 1 *A Perfect Bayesian Equilibrium is a* ***recursive equilibrium (RE)*** *if, for any period t and on-path history $h_0^t \in \mathcal{H}_0^t$, the continuation strategy profile and associated beliefs form a PBE of the continuation game.*

In a PBE, agent i's actions at a history must be a best response to the other players' actions, given agent i's beliefs. Recursive equilibrium further requires that at the start of each period on the equilibrium path, each player's actions are also a best response given the true history. Within a period, players best-respond given their beliefs, and continuation play need not be recursive after a deviation. We impose this equilibrium refinement in order to focus on the dynamics that arise from uncoordinated punishments without confounding dynamics from persistent private beliefs about the true history.[5] Appendix C extends our logic to all PBE.

A recursive equilibrium is **surplus-maximizing** if it maximizes *ex ante*

[5] See Kandori (2002) for a discussion of dynamics from private beliefs.

total expected surplus among all recursive equilibria.[6] The goal of this section is to give conditions under which every surplus-maximizing equilibria entails **biased promotions**, in the sense that $d_1 = 2$ with positive probability, even though there exists a continuation equilibrium with $d_1 = 1$ that yields strictly higher total continuation surplus. In particular, defining

$$l(\cdot) = \frac{p(\cdot|e_i = 1)}{p(\cdot|e_i = 0)}$$

as the likelihood ratio, we prove the following result.

Proposition 1 *There exist $\bar{\delta} > 0$ and continuous functions $\underline{\gamma}, \Delta : \mathbb{R}_+ \to \mathbb{R}_{++}$ such that if $\delta \in (0, \bar{\delta})$, $\gamma_2 > \underline{\gamma}(\delta)$, and $\gamma_1 - \gamma_2 < \Delta(\delta)$, then $e_{1,0} = e_{2,0} = 1$ and $d_1 = 2$ with positive probability in any surplus-maximizing recursive equilibrium. Under these conditions, the surplus-maximizing promotion policy is essentially unique:*[7] *there exist $\beta_1, \beta_2 \geq 0$ such that $d_1 = 1$ if and only if*

$$1 + \beta_1 \max\left\{0, 1 - \frac{1}{l(y_{1,0})}\right\} \geq \beta_2 \left(1 - \frac{1}{l(y_{2,0})}\right). \tag{1}$$

On the equilibrium path, continuation play following $d_1 = 1$ generates strictly higher total expected continuation surplus than continuation play following $d_1 = 2$.

Proof: See Appendix A.

Under the conditions of Proposition 1, any surplus-maximizing equilibrium entails a biased promotion tournament. Consequently, agent 2 is sometimes promoted even if promoting agent 1 would generate higher continuation surplus. Importantly, inequality (1) shows that the promotion decision is based on different criteria if agent 1 produces "high" versus "low" output. If $l(y_{1,0}) < 1$, then agent 2 is promoted based on his *absolute* performance: $d_1 = 2$ if $y_{2,0}$ exceeds some *fixed* threshold. In contrast, if $l(y_{1,0}) \geq 1$, then agent 2 must both

[6] Here, 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.

[7] Other surplus-maximizing promotion policies differ only after probability-0 events.

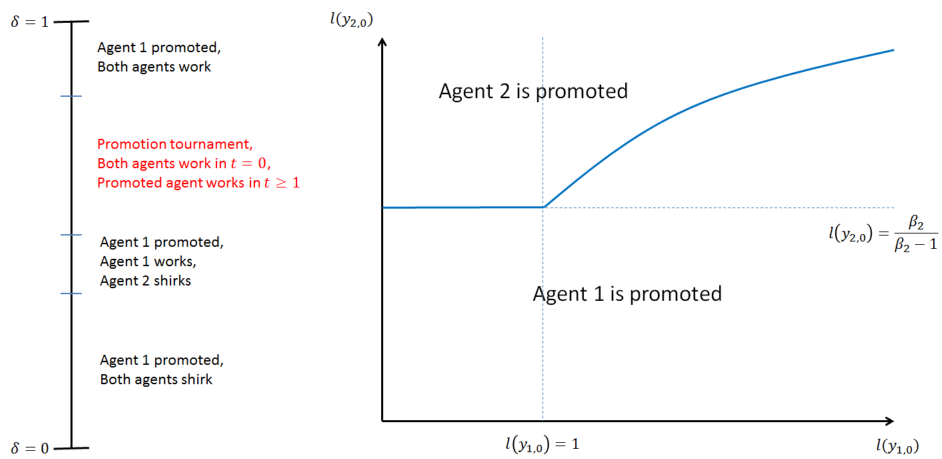

Figure 1: The left figure describes the surplus-maximizing promotion policy for different discount factors, holding all other parameters fixed. The right figure illustrates the optimal promotion scheme if the conditions of Proposition 1 are satisfied.

outperform this fixed threshold and produce high output *relative to* agent 1. See Figure 1 for an illustration.

If $l(y_{1,0}) \geq 1$, then the optimal promotion policy resembles a biased promotion tournament as in Lazear and Rosen (1981). However, the promotion does not affect *feasible* transfers in our setting and does not serve as a direct reward. Instead, we argue below that promotion determines the maximum reward that can be offered *in equilibrium*. As highlighted above, this mechanism generates stark predictions about which outputs are likely to lead to a biased promotion.

The proof of Proposition 1 first identifies necessary and sufficient conditions for participation decisions, efforts, and a promotion decision to be implemented in a recursive equilibrium. For example, agent i's incentive to exert effort in $t = 0$ depends on her expected continuation payoff. Define

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

as agent i's **reward scheme**.[8] Agent i is willing to choose $a^*_{i,0}$ and $e^*_{i,0}$ if and

[8] Abusing notation, we do not write the reward scheme conditional on $a_{i,0}$, which can be (almost) perfectly inferred from $y_{i,0}$.

only if

$$a_{i,0}^*, e_{i,0}^* \in \arg \max_{(a_{i,0}, e_{i,0})} \left\{ E[B_i(y_{i,0}) | a_{i,0}, e_{i,0}] - c a_{i,0} e_{i,0} \right\}. \tag{2}$$

Both the principal and agent i are able to walk away from their bilateral relationship, which constrains $B_i(\cdot)$ in equilibrium. In particular, agent i can guarantee himself a payoff of 0 by not participating and making no payments, so $B_i(y_{i,0}) \geq 0$. Similarly, the principal can deviate from an equilibrium payment to i. If she does so, then agent i can punish her by not participating in every future period. However, the principal can choose d_1 and make payments to the *other* agent as if she had not deviated, in which case the other agent never becomes aware of the deviation. Therefore, the principal stands to lose no more than agent i's future production if she reneges on i, and so she will renege on paying any reward that exceeds i-dyad surplus: $B_i(y_{i,0}) \leq \delta E[S_{i,1} | w_{i,0}, y_{i,0}]$. In equilibrium, each agent's reward scheme is bounded by the resulting **dynamic enforcement** constraints:

$$0 \leq B_i(y_{i,0}) \leq \delta E[S_{i,1} | w_{i,0}, y_{i,0}]. \tag{3}$$

A similar argument applies in other periods, so (2) and (3), and their analogues for $t > 0$, are necessary conditions for equilibrium. It turns out that they are also sufficient: we can implement a promotion policy, sequence of participation decisions, and sequence of efforts in equilibrium so long as exists a reward scheme that satisfies these conditions. We will return to the proof of sufficiency after discussing how these constraints shape the optimal promotion policy.

Promising to promote agent i increases i-dyad surplus and so relaxes the upper bound of (3). Suppose δ is small enough that no reward scheme satisfies (2) and (3) for $e_{i,t}^* = 1$ and $1_{i,t} = 0$. If γ_2 (and hence γ_1) are sufficiently large, then agent i *can* be motivated if d_1 is chosen uniformly at random. If $\gamma_1 - \gamma_2$ is not too large, then surplus is strictly larger when both agents work hard in $t = 0$, even if agent 2 is then promoted. Under those conditions, agent 2 is promoted with positive probability in any surplus-maximizing equilibrium.

Why is agent 2 promoted only if (1) is violated? Formally, (1) is the result of maximizing the probability that agent 1 is promoted, subject to the constraints that both agents exert effort in $t = 0$. The left-hand side of this constraint reflects the cost of promoting agent 2. Doing so has a direct cost because it decreases continuation surplus, which is reflected in the first term of (1). Promoting agent 2 also tightens the upper bound of (3) for agent 1. This upper bound does not bind if $l(y_{1,0}) < 1$, since agent 1 is optimally not rewarded for low output. However, it does bind if $l(y_{1,0}) > 1$, in which case promoting agent 2 has an incentive cost by making it more difficult to motivate agent 1 in $t = 0$. This incentive cost is represented by the second term in (1) and reflects how much an increase in agent 1's reward relaxes his incentive constraint, which depends on $l(y_{1,0})$. By the same logic, promoting agent 2 when $l(y_{2,0}) > 1$ motivates him to work harder; the right-hand side of (1) measures the size of this incentive benefit. Agent 2 is promoted exactly when this incentive benefit exceeds both the direct and incentive costs.

Why are (2) and (3), and their analogues for $t > 0$, sufficient conditions for equilibrium? Suppose for the moment that agent i earns his dyad-surplus at the start of $t = 1$ and consider the following transfers in $t = 0$: following output $y_{i,0}$, agent i pays the principal $-(1-\delta)\tau_{i,0} = B_i(y_{i,0}) - \delta E[S_{i,1}|w_{i,0}, y_{i,0}]$, while the principal pays $w_{i,0} = E[y_{i,0} - \tau_{i,0}]$ to each agent i at the start of $t = 0$. These transfers guarantee that agent i earns exactly $B_i(y_{i,0})$ in expectation after he produces $y_{i,0}$. Furthermore, given that $B_i(\cdot) \geq 0$ and agent i earns $E[S_{i,1}|w_{i,0}, y_{i,0}]$ in the continuation game, he would rather pay $-(1-\delta)\tau_{i,0}$ than renege and earn 0 continuation surplus as a punishment. The principal is willing to pay $w_{i,0}$ because she earns 0 from doing so and no more than 0 from deviating. So these payments are consistent with equilibrium so long as agents actually earn their dyad surpluses in the continuation equilibrium. We can guarantee that they do by constructing similar payments in future periods. Under these payments, the principal is willing to promote either agent because she earns 0 continuation surplus regardless of d_1.[9]

[9]This construction both makes the principal indifferent among decisions and holds her to 0 continuation surplus. We discuss the implications of relaxing these features in Section 5.

Finally, we compare Proposition 1 to a setting in which agents can perfectly coordinate punishments. In the **promotion game with public monitoring**, all variables except efforts are publicly observed. We show that surplus-maximizing promotions are never biased in this game.

Proposition 2 *In the promotion game with public monitoring, $d_1 = 1$ with probability 1 in any surplus-maximizing equilibrium.*

Proof: This result is a special case of Proposition 7 in Appendix D.

In the promotion game with public monitoring, players can jointly punish any deviation by the principal. Therefore, the principal stands to lose her entire continuation surplus following any deviation, so she is willing to reward the agents so long as the *sum* of those rewards does not exceed *total* continuation surplus. A biased promotion would decrease total continuation surplus and undermine the principal's incentive to follow through on payments, so is never optimal. That is, biased decisions are surplus-maximizing only because agents cannot coordinate to punish the principal.

3 Biased Decisions in Relational Contracts

This section generalizes Section 2 to show why biased decisions can maximize surplus in settings with uncoordinated punishments. Section 3.1 introduces a flexible model of decisions in relational contracts. Section 3.2 develops necessary and sufficient conditions for recursive equilibria. Section 3.3 applies these conditions to a class of games that yields a clean intuition for the costs and benefits of biased decisions.

3.1 The Framework

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, they play the following stage game:

1. A state of the world $\theta_t \in \Theta$ and feasible decision set $D_t \in \mathcal{D}$ 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 send to each agent i, where M is a large message space. Only the principal and agent i observe $m_{i,t}$.[10]

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 a 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.

Define the **net cost** of $(a_{i,t}, e_{i,t})$ as $C_{i,t} = a_{i,t}c(e_{i,t}) - (1 - a_{i,t})\bar{u}_i(d_t, \theta_t)$. Then agent i's and the principal's stage-game payoffs in each period t are

$$u_{i,t} = w_{i,t} + \tau_{i,t} - C_{i,t},$$
$$\pi_t = \sum_{i=1}^{N}(y_{i,t} - \tau_{i,t} - w_{i,t}),$$

[10] Formally, M is 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.

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, define **i-dyad surplus** in period t as the total continuation surplus produced by that agent,

$$S_{i,t} = \sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)(y_{i,t'} - C_{i,t'}). \tag{4}$$

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

Histories and Strategies Recall that $h_0^t \in \mathcal{H}_0^t$ is a history at the start of period 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 and $\mathcal{H}$ is the set of all 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 the principal observes all variables except effort, while agent i observes 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 $\sigma = \sigma_0 \times \cdots \times \sigma_N$, 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 to the distribution over decisions taken at that history.

Equilibrium We say a recursive equilibrium σ^* is **sequentially surplus-maximizing** if, at each on-path $h_0^t \in \mathcal{H}_0^t$, $\sigma^*|h_0^t$ 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. We give conditions under which backward-looking policies arise in surplus-maximizing recursive equilibria.

Since $E[y_{i,t}|\theta_t, d_t, e_{i,t}] \geq 0$, the harshest punishment agent i can impose on the principal is to choose $a_{i,t} = 0$ in each period t. If the principal chooses d_t to minimize $\bar{u}_i(\theta_t, d_t) \geq 0$, then $a_{i,t} = 0$ also attains agent i's min-max payoff. Given h_x^t and $i \in \{1, ..., N\}$, define agent i's **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 Section 5 discusses the fact that agents cannot coordinate punishments. Here, we highlight three other features of the model. First, $F(\cdot)$ depends only on publicly observed variables. Consequently, agents have common information about the continuation game at the start of each period, which rules out adverse selection problems. Second, $w_{i,t}$ 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, the principal sends a private message to each agent in each period, which allows the principal to reveal information about the true history to each agent. Following a deviation in one relationship, the principal can choose messages so that the other agents do not learn of that deviation. Requiring these messages to be public would not change our main results.

3.2 Necessary and Sufficient Conditions for Equilibrium

As in Section 2, agent i's incentive to exert effort in period t can be summarized in a **reward scheme** B_i that maps i's output to his expected payoff following that output. Let $\xi_{i,t} \equiv (m_{i,t}, w_{i,t}) \in \Xi_i \equiv M \times \mathbb{R}$ be the set of period-t variables that are realized before agent i chooses $(a_{i,t}, e_{i,t})$ and are not publicly observed.

Our first goal is to characterize a set of necessary and sufficient conditions for equilibrium in terms of the reward scheme B_i.

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

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

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

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

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

A credible reward scheme satisfies two sets of constraints that depend on continuation play. First, if σ specifies that the agent chooses $(a_{i,t}, e_{i,t})$ after observing $\xi_{i,t}$ at history h_d^t, then (IC) requires that B_i give him the incentive to do so. Second, (DE) bounds B_i for each output realization $y_{i,t}$. The lower bound of (DE) equals the agent's punishment payoff. The upper bound equals agent i's discounted dyad surplus, which depends on the principal's future decisions. Recall that recursive equilibrium requires that on the equilibrium path, players best-respond given the true h_0^t at the start of each period t but form Bayesian expectations given their private histories within a given period. Consequently, the expectations in (IC) and (DE) condition on the true history h_0^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 there exists a credible reward scheme for each agent i.

Lemma 1 1. *If σ^* is a recursive equilibrium, 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 recursive equilibrium σ^* that induces the same joint distribution over states of the world, decisions, efforts, and outputs as σ.*

Proof: See Appendix A.

The proof of Lemma 1 follows the intuition from Section 2. In particular, agent i is willing to choose $(a^*_{i,t}, e^*_{i,t})$ only if

$$B^*_i(h^t_d, \xi_{i,t}, y_{i,t}) = E_{\sigma^*}[(1-\delta)\tau_{i,t} + \delta U_{i,t+1} | h^t_d, \xi_{i,t}, y_{i,t}]$$

satisfies (IC). Agent i would rather renege and be punished than earn less than his punishment payoff, implying the lower bound of (DE). 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 would rather renege than pay agent i more than the total surplus he expects to produce in the future, which implies the upper bound of (DE). These arguments prove part 1.

The proof of part 2 constructs 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. 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 that corresponds to his realized output. A deviation is punished by the breakdown of the corresponding relationship.

On the equilibrium path, each agent can perfectly infer the principal's stage-game payoff from his wage, effort, and expected fines. 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.[11] Agent i earns his entire i-dyad surplus in each period, but he pays fines following low

[11]See Appendix F for more. Informally, this construction requires the agent to be able to infer when the principal would earn a strictly positive payoff from her interaction *with that agent*. In our model, this requirement is not particularly onerous; it requires only that each agent observe the impact that *his own* effort has on his expected output *today*.

output. He is willing to exert effort and make the specified payments because these fines are derived from a credible reward scheme.

3.3 Biased Decisions in Smooth Mean-Shifting Games

Biased decisions can affect equilibrium surplus in three ways. First, they have a *direct cost* because they reduce total continuation surplus. However, if they are biased toward agent i in the sense that they increase $E_{\sigma^*}\left[S_{i,t+1}|h_d^t, \xi_{i,t}, y_{i,t}\right]$, then they relax (DE) for i. So biased decisions can have an *incentive benefit*: agent i can earn larger rewards in equilibrium following $y_{i,t}$, 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 incentive cost and incentive benefit of a biased decision vary history-by-history because agent i's dynamic enforcement constraint (DE) might bind at some outputs but not others. The upper bound of (DE) 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 (DE) is unlikely to bind if agent i "performs poorly." Tightening i'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, at the expense of those who have performed poorly.

This intuition is particularly clear if decisions do not affect the precision of output as a signal of effort, and if surplus varies smoothly in decisions and effort. Our next result focuses on these games.

Definition 3 *A* ***smooth mean-shifting game*** *satisfies:*

1. *For every $t \geq 0$, $F(\cdot)$ is independent of history, with*

$$D_t = \left\{ (d_1, ..., d_N) | d_i \in \mathbb{R}_+, \sum_{i=1}^{N} d_i \leq 1 \right\}.$$

2. *For every $i \in \{1, ..., N\}$, $\mathcal{E}_i = [0,1]$ and $\bar{u}_i(d_t, \theta_t)$ is constant in d_t. $c_i(\cdot)$ is smooth, strictly increasing, and strictly convex, with $c_i'(0) = 0$.*

3. *For every $i \in \{1, ..., N\}$, there exists a function $\gamma_i : \Theta \times \mathbb{R} \to \mathbb{R}$ such that $y_{i,t} = x_{i,t} + \gamma_i(\theta_t, d_{i,t})$, where $x_{i,t}$ is a random variable with distribution*

$$(1 - e_i)\tilde{P}_i^L(x_i) + e_i \tilde{P}_i^H(x_i).$$

   Here, $\tilde{P}_i^L, \tilde{P}_i^H$ are smooth distributions with densities $\tilde{p}_i^L$ and $\tilde{p}_i^H$ such that $\frac{\tilde{p}_i^H(\cdot)}{\tilde{p}_i^L(\cdot)}$ is strictly increasing. For all $\theta \in \Theta$, $\gamma_i(\theta, \cdot)$ is smooth, concave, strictly increasing, and satisfies $\lim_{d_i \to 0} \frac{\partial \gamma_i}{\partial d_i}(\theta, d_i) = \infty$.

In a smooth mean-shifting game, the principal's decision in each period assigns a weight $d_{i,t}$ to each agent i. This weight increases agent i's output according to an effort-independent, strictly concave, and smooth function $\gamma_i(\theta, \cdot)$. In particular, the principal's decisions do not affect the precision of output as a signal of effort. Agent i's output is drawn from a mixture distribution, which ensures that we can replace the incentive-compatibility constraint (IC) with its first-order condition.[12]

Both θ_t and $d_{i,t}$ shift the mean of $y_{i,t}$ without otherwise affecting its distribution. Given that the distribution of $x_{i,t} \equiv y_{i,t} - \gamma_i(\theta_t, d_{i,t})$ depends only on $e_{i,t}$, first-best effort for agent i satisfies

$$e_i^{FB} \equiv \arg\max_{e_i} \left\{ E[x_i | e_i] - c(e_i) \right\}.$$

[12]Since $\frac{\tilde{p}_i^H}{\tilde{p}_i^L}$ is strictly increasing, such distributions satisfy CDFC and strict MLRP. See Rogerson (1985).

For each e_i, define

$$l_i(x_i|e_i) = \frac{\tilde{p}_i^H(x_i) - \tilde{p}_i^L(x_i)}{(1-e_i)\tilde{p}_i^L(x_i) + e_i\tilde{p}_i^H(x_i)}.$$

Because $\frac{\tilde{p}_i^H(\cdot)}{\tilde{p}_i^L(\cdot)}$ is strictly increasing, there exists a unique $x_i^*(e_i) \in \mathbb{R}$ that satisfies $l_i(x_i^*(e_i)|e_i) = 0$. Loosely, $x_i > x_i^*(e_i)$ statistically suggests that agent i's effort was no less than e_i.

We prove that in a smooth mean-shifting game, every surplus-maximizing relational contract entails biased decisions so long as players are neither too patient nor too impatient.

Proposition 3 *Consider a smooth mean-shifting game. Then:*

1. *In any surplus-maximizing recursive equilibrium σ^* and any $t \geq 0$, $E_{\sigma^*}\left[\sum_{i=1}^N d_{i,t}\right] = 1$.*

2. *There exist $\underline{\delta} < \bar{\delta}$ such that if $\delta \in [\underline{\delta}, \bar{\delta}]$, no surplus-maximizing recursive equilibrium is sequentially surplus-maximizing.*

3. *Consider a surplus-maximizing recursive equilibrium, and suppose h_0^{t+1} is an on-path history such that $e_{i,t}^* \in (0, e_i^{FB})$, $x_{i,t} > x_i^*(e_{i,t})$, and $x_{j,t'} < x_j^*(e_{j,t'})$ for all $t' \leq t$. The continuation equilibrium at almost every such h_0^{t+1} is* not *surplus-maximizing.*

Proof: See Appendix A.

Proposition 3 is an implication of a more general result found in Appendix B. The first part of this result says that any surplus-maximizing relational contract will use the full decision "budget." Suppose $\sum_{i=1}^N d_{i,t} < 1$ at some on-path history in a recursive equilibrium σ^*. Because $d_{i,t}$ does not affect the precision of output as a signal of effort, we can construct a perturbed equilibrium by increasing $d_{i,t}$ for some agent i while holding all other actions constant. This perturbation strictly increases expected total surplus because $\gamma_i(\theta, d_i)$ is strictly increasing in d_i, so σ^* cannot be surplus-maximizing.

The second and third parts of Proposition 3 give conditions under which biased decisions are surplus-maximizing. Consider the first period of a surplus-maximizing equilibrium σ^*. We can perturb this equilibrium by increasing $d_{i,0}$ and decreasing $d_{j,0}$ by the same amount while holding all other actions fixed. Since $\gamma_i(\cdot)$ and $\gamma_j(\cdot)$ are both smooth, this perturbation smoothly increases $E_{\sigma^*}[S_{i,0}]$ and smoothly decreases $E_{\sigma^*}[S_{j,0}]$. Therefore, the equilibrium dyad surplus frontier is differentiable in this and so every period. In particular, starting from a surplus-maximizing continuation equilibrium, slightly increasing i-dyad surplus and slightly decreasing j-dyad surplus has a second-order effect on expected total continuation surplus.

Now, suppose h_0^{t+1} satisfies the conditions given in part three of Proposition 3. That is, agent i has just exerted positive effort that is less than first-best and produced high output given that effort, while agent j has *never* produced high output. We can find $\underline{\delta} < \bar{\delta}$ such that these histories occur with positive probability in any surplus-maximizing equilibrium for $\delta \in [\underline{\delta}, \bar{\delta}]$. If the continuation equilibrium at h_0^{t+1} was surplus-maximizing, then we could slightly increase $E_{\sigma^*}\left[S_{i,t+1}|h_0^{t+1}\right]$ and slightly decrease $E_{\sigma^*}\left[S_{j,t+1}|h_0^{t+1}\right]$ at a second-order direct cost. By (DE), increasing $E_{\sigma^*}\left[S_{i,t+1}|h_0^{t+1}\right]$ means that agent i can be given a strictly higher reward, which motivates i to exert strictly more effort because $x_{i,t} > x_i^*(e_{i,t}^*)$. Since $e_{i,t}^* < e_i^{FB}$, this perturbation has a first-order incentive benefit. While decreasing $E_{\sigma^*}\left[S_{j,t+1}|h_0^{t+1}\right]$ tightens the upper bound of (DE) for agent j in each $t' \leq t$, this upper bound does not bind in any of these periods because $x_{j,t'} < x_j^*(e_{j,t'}^*)$ for all $t' \leq t$. Therefore, this perturbation does not affect agent j's incentives and so has no incentive cost. 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, and so improves *ex ante* total surplus.

One subtlety complicates this intuition: increasing $e_{i,t}^*$ changes the distribution over $y_{i,t}$, which potentially changes the distribution over continuation play and hence the payments that can be promised to *other* agents in period t. Our proof constructs a mapping from the perturbed distribution over $y_{i,t}$ to the original distribution over continuation play to ensure that the distribution over all other agents' dyad surpluses, and hence their incentives to exert effort, remain unchanged as $e_{i,t}$ increases.

As Lemma 1 suggests, biased decisions can also be optimal in games that are neither smooth nor mean-shifting. Indeed, neither the promotions example in Section 2 nor the hiring example in the next section are smooth. Moreover, Appendix B extends the logic of Proposition 3 to a class of smooth games that are not mean-shifting. This extension requires a more complicated argument, but the tradeoff between incentive costs, direct costs, and incentive benefits still holds.

4 Biased Hiring Decisions

Consider an owner of an up-and-coming business who must decide how quickly to expand. 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 through performance bonuses or in the future through, say, equity. But promises to pay bonuses today and not to dilute equity in the future are only credible if early employees know they will remain indispensable in the future. We argue that the owner might ensure these employees remain essential by being slow to hire additional workers as demand increases.

Definition 4 *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.*[13]
- *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* $d_t = 2$.

[13] This assumption is without loss of generality for our result.

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. If a hired agent works hard, he produces output that is increasing in demand but exhibits diminishing returns—represented by $\alpha < 1$—in the number of workers hired.

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

Proposition 4 *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 recursive equilibrium 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 $Pr_{\sigma^*}\{d_{t'} = 1, \theta_{t'} = R\} > 0$.*

One surplus-maximizing recursive equilibrium has the following hiring policy: 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 $t' > t$.

Proof: See Appendix A.

The two conditions in Proposition 4 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 sequentially efficient decision in each period. However, if demand is initially weak, then the firm might continue to hire only one worker even after demand becomes robust. Under the conditions of Proposition 4, agent 1 is willing to exert effort while $\theta_t = W$ only if decisions are biased towards him once demand becomes robust. The principal does so by refraining from hiring agent 2, which decreases total surplus but increases the surplus produced by agent 1.

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, but it illustrates that the surplus-maximizing relational contract can entail substantial and long-lasting distortions.

5 Discussion and Conclusion

Extensions: While our main results restrict to recursive equilibria, our central intuition does not depend on this refinement. Appendix C proves an analogue to Proposition 3 for the full, non-recursive set of PBE. In a PBE, total expected continuation surplus at a history depends on players' information at that history. Therefore, we cannot define sequential surplus-maximization in terms of maximizing expected continuation surplus *at a history*. Instead, we prove that *ex ante* expected continuation surplus from period t onward is independent of t, so we can define sequentially surplus-maximizing PBE in terms of *ex ante* expected total continuation surplus in each period. We show that in smooth mean-shifting games, decisions cannot depend on the history in a sequentially surplus-maximizing PBE, which consequently cannot perform better than sequentially surplus-maximizing recursive equilibria. Hence, surplus-maximizing PBE entail biased decisions under the conditions of Proposition 3.

Appendix D considers surplus-maximizing equilibria if all variables except effort are publicly observed in the model from Section 3.1. The resulting game has product monitoring, so agents do not need to condition on private histories in surplus-maximizing PBE.[14] As a result, agents can perfectly coordinate to punish the principal, which means that biased decisions are never surplus-maximizing in the game with public monitoring. As in Proposition 2, biased decisions decrease total continuation surplus and so make the principal less

[14] See Fudenberg and Levine (1994).

willing to follow through on payments. Appendix D also shows that if agents can communicate with one another, then they can be given the incentive to truthfully report their private histories. Consequently, communication also leads to coordinated punishments and so obviates the need for biased decisions in surplus-maximizing equilibria.

Even if agents can coordinate punishments, they might not be able to do so perfectly. Appendix E explores imperfectly coordinated punishments in the context of our hiring example. We look at a highly stylized monitoring structure in which agents can coordinate to punish the principal with a positive probability that is strictly less than 1. If they fail to coordinate, then only the betrayed agent punishes the principal. We give conditions under which the principal delays hiring in any surplus-maximizing equilibrium, which suggests that our intuition extends to at least some settings with imperfectly coordinated punishments.

Finally, the construction in the proof of Lemma 1 has two stark features. First, the principal is indifferent to her on-path decisions, which means that we can implement any policy so long as (IC) and (DE) are satisfied. However, our basic intuition does not rely on principal indifference. Using the promotions example from Section 2, Appendix F considers equilibria in which the principal *strictly* prefers her on-path promotion decisions. We prove that equilibria with biased promotion policies can still outperform sequentially surplus-maximizing equilibria, even with this additional constraint. Second, the principal earns 0 continuation surplus in our construction. While this feature is convenient, such extreme transfers are not required in many applications. If agents' outside options $(\bar{u}_i)_{i=1}^N$ do not depend on d and $E[S_{i,t}|h_0^t]$ is strictly positive at every on-path history h_0^t, we can construct an equilibrium in which the principal earns strictly positive continuation surplus from agent i in every period.

Conclusion: We have argued that biased decisions increase the future surplus produced by some agents and so complement monetary rewards in equilibrium. Consequently, employees are rewarded with both higher compensation and greater responsibilities, divisions are promised both monetary incentives and non-monetary investments, and firms encourage effort today by promising not to expand too quickly in the future.

Proposition 2 and Appendix D imply that coordinating punishments strictly improves total surplus. In practice, the principal might facilitate communication among agents, commit to a public bonus pool, or take other steps to help agents coordinate. Such attempts will be successful only if the principal can commit not to distort messages or divert funds, and if the agents actually follow through on joint punishments. We view these requirements as key stumbling blocks that could undermine such attempts in practice.

An important extension would be to settings in which agents' efforts are complements or substitutes. In such settings, each agent's dyad surplus depends on the other agents' private efforts, so a straightforward extension of our techniques would not work. We believe that conditions similar to (IC) and (DE) are necessary but not sufficient if efforts are substitutes. If efforts are complements, then equilibria must deter the principal from simultaneously reneging on multiple agents.

References

Aghion, P. and J. Tirole (1997). Formal and real authority in organizations. *Journal of Political Economy 105*(1), 1–29.

Ali, N. and D. Miller (2016). Ostracism and forgiveness. *American Economic Review 106*(8), 2329–2348.

Ali, N., D. Miller, and D. Yang (2016). Renegotiation-proof multilateral enforcement.

Andrews, I. and D. Barron (2016). The allocation of future business: Dynamic relational contracts with multiple agents. *American Economic Review 106*(9), 2742–2759.

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

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

Baker, G., M. Jensen, and K. Murphy (1988). Compensation and incentives: Practice vs. theory. *The Journal of Finance 43*(3), 593–616.

Benson, A., D. Li, and K. Shue (2017). Promotions and the "peter principle".

Bewley, T. (1999). *Why Wages Don't Fall During a Recession.* Cambridge, MA: Harvard University Press.

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. *The Quarterly Journal of Economics 102*(1), 147–159.

Dessein, W. (2002). Authority and communication in organizations. *The Review of Economic Studies 69*(4), 811–838.

Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica 61*(5), 1047–1071.

Fong, Y.-F. and J. Li (2017a). Information revelation in relational contracts. *The Review of Economic Studies 84*(1), 277–299.

Fong, Y.-F. and J. Li (2017b). Relational contracts, limited liability, and employment dynamics.

Fuchs, W. (2007). Contracting with repeated moral hazard and private evaluations. *The American Economic Review 97*(4), 1432–1448.

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.

Gertner, R. and D. Scharfstein (2013). Internal capital markets. In R. Gibbons and J. Roberts (Eds.), *Handbook of Organizational Economics*, pp. 655–679.

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.

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

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics 13*(2), 324–340.

Kandori, M. (1992). Social norms and community enforcement. *The Review of Economic Studies 59*(1), 63–80.

Kandori, M. (2002). Introduction to repeated games with private monitoring. *Journal of Economic Theory 102*(1), 1–15.

Krueger, A. and A. Mas (2004). Strikes, scabs, and tread separations: Labor strife and the production of defective bridgestone/firestone tires. *Journal of Political Economy 112*(2), 253–289.

Lazear, E. and P. Oyer (2013). Personnel economics. In R. Gibbons and J. Roberts (Eds.), *Handbook of Organizational Economics*, pp. 479–519.

Lazear, E. and S. Rosen (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy 89*(5), 841–864.

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

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

Li, J., N. Matouschek, and M. Powell (2017). Power dynamics in organizations. *American Economic Journal: Microeconomics*. Forthcoming.

Lipnowski, E. and J. Ramos (2017). Repeated delegation.

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. (2016). Relational contracts with private information. *Econometrica 84*(1), 317–346.

Mas, A. (2008). Labour unrest and the quality of production: Evidence from the construction equipment resale market. *The Review of Economic Studies 75*(1), 229–258.

Rayo, L. (2007). Relational incentives and moral hazard in teams. *The Review of Economic Studies 74*(3), 937–963.

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

Waldman, M. (2013). Theory and evidence in internal labor markets. In R. Gibbons and J. Roberts (Eds.), *Handbook of Organizational Economics*, pp. 520–571.

Watson, J. (1999). Starting small and renegotiation. *Journal of Economic Theory 85*(1), 52–90.

Watson, J. (2002). Starting small and commitment. *Games and Economic Behavior 38*(1), 176–199.

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

A For Online Publication: Proofs

A.1 Proof of Proposition 1

Define $d_0 = 0$ and $d_t = d_1$ for all $t > 1$. This proof involves two steps. First, we show that under the conditions of the result, any surplus-maximizing equilibrium entails promoting agent 2 with positive probability. Second, we characterize the optimal promotion policy by maximizing the probability that agent 1 is promoted, subject to the constraints that both agents are willing to exert effort in $t = 0$.

Lemma 1 identifies necessary and sufficient conditions for a recursive equilibrium. In particular, efforts, participation decisions, and promotion decisions are consistent with recursive equilibrium if and only if there exists a reward scheme $B_i(\cdot) : \mathcal{H}_0^t \times \{0,1,2\} \times \mathbb{R}^2$ such that for each $i \in \{1,2\}$ and on-path h_0^t,

$$(a_{i,t}, e_{i,t}) \in \arg\max_{\tilde{a}_{i,t},\tilde{e}_{i,t}} \left\{ E_y \left[B_i(h_0^t, d_t, w_{i,t}, y_{i,t}) | h_0^t, d_t, w_{i,t}, \tilde{a}_{i,t}, \tilde{e}_{i,t} \right] - (1-\delta) c \tilde{a}_{i,t} \tilde{e}_{i,t} \right\} \quad (5)$$

and

$$0 \le B_i(h_0^t, d_t, w_{i,t}, y_{i,t}) \le \delta E[S_{i,t+1} | h_0^t, d_t, w_{i,t}, y_{i,t}] \quad (6)$$

for each $y_{i,t} \in \mathbb{R}$.[15]

Define $S^B = E[y_i | e_i = 1] - c$ and $\underline{S} = E[y_i | e_i = 0]$. Let y^* be the unique output for which $l(y^*) = 1$. Then it is straightforward to show that agent i has the strongest incentive to choose $e_{i,t} = 1$ if the lower bound of (6) binds following $y_{i,t} < y^*$, and the upper bound binds otherwise. Consequently, an

[15] Unlike the model in Section 3, the principal cannot send messages to her agents when wages are paid in this example. In the proof of Lemma 1, (5) and (6) are necessary conditions for equilibrium, regardless of whether messages are available or not. The proof that these conditions are sufficient, uses messages to inform agents about (i) the history at the start of that period, (ii) that agent's equilibrium participation and effort decisions, and (iii) the equilibrium penalty schedule that that agent is supposed to pay after output is realized. In all the constructions used here, agents can infer this information from what they observe, and so messages are not needed for these constructions to be recursive equilibria.

equilibrium exists in which agent i exerts effort in periods $t \geq 1$ if and only if

$$c \leq \frac{\delta}{1-\delta} \int_{y^*}^{\infty} (S^B + 1_{i,t}\gamma_i)(p(y_i|1) - p(y_i|0))dy_i. \tag{7}$$

Let $\bar{\delta}$ satisfy (7) with equality for $1_{i,t} = 0$. If $\delta \in (0, \bar{\delta})$, then $e_{i,t} = 0$ in $t \geq 1$ if $1_{i,t} = 0$. For any such δ, define $\underline{\gamma}$ so that (7) holds with equality for $1_{i,t} = \frac{1}{2}$. For any $\gamma_1 > \gamma_2 > \underline{\gamma}$, both agents can be motivated to work hard in $t = 0$ if each is promoted with probability $\frac{1}{2}$, independent of realized output. Finally, define $\bar{\Delta} \equiv \frac{1-\delta}{\delta} S^B$. Then for any $\gamma_1 - \gamma_2 < \bar{\Delta}$, setting $e_{1,0} = e_{2,0} = 1$ and allocating the promotion at random generates more surplus than setting $e_{1,0} = 1$, $e_{2,0} = 0$, and always promoting agent 1.

Now, suppose $\delta < \bar{\delta}$, $\gamma_2 > \underline{\gamma}$, and $\gamma_1 - \gamma_2 < \bar{\Delta}$. By the argument above, no surplus-maximizing equilibrium entails $d_1 = 1$ with probability 1. So it suffices to find the surplus-maximizing promotion tournament that induces both agents to work hard in $t = 0$. Following output $y_0 \in \mathbb{R}^2$ in $t = 0$, let $\rho(y_0)$ denote the probability that $d_1 = 1$. Then the surplus-maximizing equilibrium must maximize the expected probability that $d_1 = 1$, conditional on motivating both agents to work hard.

$$\max_{\rho:\mathbb{R}^2 \to [0,1]} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho(y) p(y_1|1) p(y_2|1) dy_1 dy_2$$

subject to both agents choosing $e_{i,0} = 1$:

$$\begin{gathered} c \leq \tfrac{\delta}{1-\delta} \int_{-\infty}^{\infty} \int_{y^*}^{\infty} \left(\underline{S} + \rho(y) \left[S^B - \underline{S} + \gamma_1\right]\right) \left[p(y_1|1) - p(y_1|0)\right] p(y_2|1) dy_1 dy_2 \\ c \leq \tfrac{\delta}{1-\delta} \int_{-\infty}^{\infty} \int_{y^*}^{\infty} \left(\underline{S} + (1-\rho(y)) \left[S^B - \underline{S} + \gamma_2\right]\right) \left[p(y_2|1) - p(y_2|0)\right] p(y_1|1) dy_2 dy_1 \end{gathered}$$

for agents 1 and 2, respectively.

This constrained maximization problem in linear in $\rho(y)$ for each y, and its Lagrangian can be solved pointwise. If $l(y_2) < 1$, then clearly $\rho(y) = 1$. If $l(y_2) \geq 1$ and $l(y_1) < 1$, then $\rho(y) = 1$ whenever

$$1 > \lambda_2 \frac{\delta}{1-\delta} \left(S^B - \underline{S} + \gamma_2\right) \left(1 - \frac{1}{l(y_2)}\right).$$

If $l(y_2) \geq 1$ and $l(y_1) \geq 1$, then $\rho(y) = 1$ whenever

$$1+\lambda_1 \frac{\delta}{1-\delta}\left(S^B - \underline{S} + \gamma_1\right)\left(1 - \frac{1}{l(y_1)}\right) > \lambda_2 \frac{\delta}{1-\delta}\left(S^B - \underline{S} + \gamma_2\right)\left(1 - \frac{1}{l(y_2)}\right).$$

Defining $\beta_i = \lambda_i \frac{\delta}{1-\delta}(S^B - \underline{S} + \gamma_i)$, we can combine these conditions inequalities to yield (1). ■

A.2 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 (IC) 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 (DE). 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 (DE). $\square$

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 $E_\sigma\left[y_{i,t}|h_d^t, \xi_{i,t}\right] \geq 0$ by assumption and (DE) 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 (DE).

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 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 $t+1$ onward, 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, (IC) 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 (DE).

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

A.3 Proof of Proposition 3

This proof builds on Proposition 5, which covers a more general class of games and may be found in Appendix B.

Suppose continuation equilibrium $\sigma^*|h_0^t$ is surplus-maximizing at h_0^t. Claim 6 of Proposition 5 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 each $\gamma_i(\theta, \cdot)$ is 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 $e_i^* \in [0, e_i^{FB}]$. 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 (10), e_i^* is defined implicitly by

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

where $S_i^* = E[y_i - c(e_i^*)|e_i^*]$ is a strictly concave function of e_i^*. Because $c'(0) = 0$, $e_i^{FB} > 0$ 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 $e_{i,t} = e_i^*$, $x_{i,t} > x_i^*$ with positive probability in each t. Similarly, $x_{j,t'} < x_j^*$ for all $t' \leq t$ with positive probability in each t. Therefore, the conditions of Proposition 5, part 1, hold for a set of histories Z_t that occur with positive probability in each $t > 0$ in any sequentially surplus-maximizing equilibrium. Proposition 5 then implies that continuation play at these histories cannot be surplus-maximizing. This contradicts our assumption that σ^* is surplus-maximizing. ■

A.4 Proof of Proposition 4

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 4, where $\chi > 0$ is chosen so that agent 1's constraint (DE) 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 (DE) binds, proving the claim. ■

B For Online Publication: Smooth Games that are not Mean-Shifting

B.1 Statement of Result

This appendix extends the analysis in Section 3.3 to a broader class of **smooth** games. The key difference is that the principal's decision potentially affects the informativeness of output as a function of effort in smooth games that are not mean-shifting. This added generality substantially complicates both the statement of the result and the proof. In particular, since each agent's weight $d_{i,t}$ potentially affects their equilibrium efforts *in period* t, we must ensure that a higher weight $d_{i,t}$ actually leads to a higher i-dyad surplus in period t.

Definition 5 *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*,*

$$p_i(y_i|\theta, \tilde{d}_i, e_i) = \int_{-\infty}^{\infty} R_i(y_i|x) p_i(x|\theta, d_i, e_i) dx. \tag{8}$$

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. Proposition 3 is a straightforward implication of this result.

Proposition 5 *Let σ^* be a surplus-maximizing recursive equilibrium of a smooth game. Then:*

1. ***Backward-looking policies:*** *For any agents i and j, let Z_{t+1} be the set of on-path 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 $h_0^{t+1} \in Z_{t+1}$, $\sigma^*|h_0^{t+1}$ is not surplus-maximizing.*

2. *For all $t \geq 0$, $E_{\sigma^*}\left[\sum_{i=1}^{N} d_{i,t}\right] = 1$.*

B.2 Proof of Proposition 5

B.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 $h_0^{t+1} \in Z_{t+1}$ 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 $h_0^{t+1} \in Z_{t+1}$ 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 $t+1$ onward, 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 $\Pr\{Z_{t+1}\} > 0$ 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 (IC) and (DE) 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 (IC) with its first-order condition. To maximize i's effort, the lower bound of (DE) 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 $t+2$ onward.

6. (Completes elements 2 and 3) We consider $h_0^{t+1} \in Z_{t+1}$. 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 5 implies that the *most efficient* e_i satisfying (IC) and (DE), 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 (DE) 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.

B.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 B.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

$$0 = \begin{array}{c} \int_{-\infty}^{y_i^*(\theta,d_i,e_i)} \bar{U}_i(\theta)\frac{\partial p_i}{\partial e_i}(y_i|\theta, \tilde{d}_i, \hat{e}_i)dy_i + \\ \int_{y_i^*(\theta,d_i,e_i)}^{\infty} S_i\left(G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \hat{e}_i)\right)\frac{\partial p_i}{\partial e_i}(y_i|\theta, \tilde{d}_i, \hat{e}_i)dy_i - c'(\hat{e}_i) \end{array}. \tag{9}$$

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 (9) 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 $\frac{\partial H}{\partial \hat{e}_i} \neq 0$.[16]

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 (IC) and (DE) if and only if either (i) $e_{i,t}^* = \min \mathcal{E}_i$, or (ii)

$$c'(e_{i,t}^*) \leq \begin{array}{c} \int_{-\infty}^{y_i^*(\theta_t,d_{i,t},e_{i,t}^*)} \bar{U}_i(\theta_t) \frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) dy_i + \\ \int_{y_i^*(\theta_t,d_{i,t},e_{i,t}^*)}^{\infty} E_{\sigma^*}[S_i|h_d^t, \xi_{i,t}^*, y_i] \frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) dy_i \end{array}. \tag{10}$$

Proof of Claim 3: Suppose $e_{i,t}^* > \min \mathcal{E}_i$ does not satisfy (10). Because p_i satisfies MLRP and CDFC, we can replace (IC) with its first-order condition

[16] 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 $x = G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \hat{e}_i)$.

as in Rogerson (1985):

$$c'(e_{i,t}^*) = \int_{-\infty}^{\infty} B_i(h_d^t, \xi_{i,t}^*, y_i) \frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) dy_i. \quad (11)$$

Consider choosing B_i to maximize the right-hand side of this equality, subject to the constraint (DE). 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 (10). 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 (10). Clearly, the right-hand side of (11) 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 (11) 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 (11) is satisfied. □

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 (10) 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 (10) 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 (10) for agent i without affecting it for $j \neq i$. So we can assume S_i is increasing without loss.

Suppose (10) 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 (10) 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 $h^t_0 \in \mathcal{H}^t_0$, suppose $\sigma^*|h^t_0$ is surplus-maximizing with $d_{i,t}, d_{j,t} \in (0,1)$. Define $\mathbb{I}_{i,t} = 1$ if (10) holds with equality at a successor history h^t_d, 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 $i, j \in \{1, ..., N\}$,

$$\frac{\partial s_i}{\partial d_i} + \mathbb{I}_{i,t}\frac{\partial s_i}{\partial e_i}\frac{\partial \hat{e}_i}{\partial \tilde{d}_i} = \frac{\partial s_j}{\partial d_j} + \mathbb{I}_{j,t}\frac{\partial s_j}{\partial e_j}\frac{\partial \hat{e}_j}{\partial \tilde{d}_j} \tag{12}$$

with probability 1 following h^t_0.

Proof of Claim 5: Suppose towards contradiction that the left-hand side of (12) 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 (10). For each agent $k \in \{1, ..., N\}$, this perturbation induces an identical marginal distribution over continuation play from $t + 1$ onward. 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 (10) was slack in the original equilibrium. But (10) 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 (11) 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 $t+1$ onward 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

$$s_i(\theta_t, d_{i,t}+\epsilon, \tilde{e}_{i,t})+s_j(\theta_t, d_{j,t}-\epsilon, \tilde{e}_{j,t})-(s_i(\theta_t, d_{i,t}, e_{i,t}) + s_j(\theta_t, d_{j,t}, e_{j,t})) \leq 0 \quad (13)$$

Dividing by $\epsilon > 0$, and taking the limit as $\epsilon \to 0$ results in (12) with a weak inequality $\leq$. Contradiction; we assumed $>$. □

Completing the proof of Statement 1 Let $h_0^{t+1} \in Z_{t+1}$. If $\sigma^*|h_0^{t+1}$ is surplus-maximizing, then (12) 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 5. 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 (10) under $\tilde{d}_{i,t+1}$ because it satisfied this inequality under $d^*_{i,t+1}$. We conclude that continuation play from period $t+1$ onward 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 (13). But then (12) 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 $t' \leq t$, (10) implies that it is without loss to assume that the upper bound of (DE) 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 5. Furthermore, $\tilde{e}_{i,t+1}$ is weakly more efficient than $e^*_{i,t+1}$ by construction. Hence, $E_{\tilde{\sigma}}\left[S_{i,t+1} | h_d^t, \xi_{i,t}, y_{i,t}\right] > E_{\sigma^*}\left[S_{i,t+1} | h_d^t, \xi_{i,t}, y_{i,t}\right]$ as desired.

By assumption, $e^*_{i,t} < e_i^{FB}(\theta_t, d^*_{i,t})$. Consequently, (10) 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 (10). 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 onward. 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 $\lim_{\epsilon \to 0} \tilde{e}_{i,t-1} = e^*_{i,t-1} < e_i^{FB}(\theta_{t-1}, d_{i,t-1})$. The second term is positive by the argument above. Hence, this perturbation increases total continuation surplus in period $t-1$ onward. 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^*. ■

B.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 (DE) in all previous periods. But σ^* is surplus-maximizing; contradiction. ■

C For Online Publication: Biased Decisions in PBE

This appendix shows that an analogue of Proposition 3 holds for the full set of PBE in smooth mean-shifting games. The central difficulty in extending Proposition 3 is that different players potentially form different beliefs about the true history in each period. In particular, in a recursive equilibrium, both (IC) and (DE) condition on the *true* history at the start of period t, h_0^t. In a PBE, however, these constraints would condition only on agent i's information set, $\phi_i(h_0^t)$. Consequently, play at a given history is not necessarily an equilibrium of the continuation game.

This complication means that our definition of sequentially surplus-maximizing equilibria does not immediately extend to PBE, since the set of expected *continuation* payoffs might vary with the history and so not be easily comparable to equilibrium payoffs in the first period. However, it turns out that the set of *ex ante* expected continuation payoffs attainable in a PBE is stationary over time. That is, define

$$\bar{V} = \max_{\sigma^* \in PBE} E_{\sigma^*}\left[\sum_{i=1}^{N} S_{i,0}\right]$$

as the maximum *ex ante* total surplus attainable in a PBE. We show that If (θ_t, D_t) is i.i.d., then *ex ante* expected continuation payoffs from any period t onward in a PBE cannot exceed $\bar{V}$.

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 C.1.

Lemma 2 shows that equilibrium *ex ante* expected continuation payoffs are recursive in t, even if continuation play is not. The proof of this result has two steps. First, establishes appropriate analogues of (IC) and (DE) for the full set of PBE. This argument 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, the principal earns 0 continuation surplus on the equilibrium path in our construction.

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 chooses $h_0^t \in \mathcal{H}_0^t$ 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 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^*.

With Lemma 2 in hand, we can define what it means for a PBE to be sequentially surplus-maximizing. Say a PBE is **PBE-sequentially surplus-maximizing** if in each $t \geq 0$, $E_{\sigma^*}\left[\sum_{i=1}^N S_{i,t}\right] = \bar{V}$. Lemma 2 implies that a PBE-sequentially surplus-maximizing equilibrium maximizes *ex ante* expected continuation surplus in each period.

Lemma 2 shows that PBE-sequentially surplus-maximizing do indeed attain the maximum *ex ante* expected continuation surplus in every period. Given this result, we can prove that in smooth mean-shifting games, there exists a range of discount factors for which no surplus-maximizing PBE is PBE-sequentially surplus-maximizing.

Proposition 6 *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 Proposition 3. Then no surplus-maximizing PBE is PBE-sequentially surplus-maximizing.*

Proof: See Appendix C.2.

As in Proposition 3, backward-looking policies are surplus-maximizing in Proposition 6 because they make strong effort incentives credible. In any

PBE-sequentially surplus-maximizing equilibrium, 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 any sequentially surplus-maximizing PBE as a function of θ_t, which implies that on-path decisions depend only on the public history. As a result, any PBE-sequentially surplus-maximizing equilibrium generates the same total surplus as a sequentially surplus-maximizing RE. But such equilibria cannot be surplus-maximizing under the conditions of Proposition 3. Hence, backward-looking policies remain surplus-maximizing, even in the full set of PBE.

C.1 Proof of Lemma 2

We first prove an extension of Lemma 1 to PBE.

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

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

$$(a_{i,t}, e_{i,t}) \in \arg\max_{a_i, e_i} E_\sigma \left[B_i(\phi_i(h^t_d), \xi_{i,t}, y_{i,t}) | \phi_i(h^t_d), \xi_{i,t}, a_i, e_i \right] - (1-\delta) C_i. \tag{14}$$

2. For each on-path h^t_y,

$$\delta E_\sigma \left[\bar{U}_i(h^{t+1}_0) | \phi_i(h^t_d) \right] \leq B_i(\phi_i(h^t_d), \xi_{i,t}, y_{i,t}) \leq \delta E_\sigma \left[S_{i,t+1} | \phi_i(h^t_d), \xi_{i,t}, y_{i,t} \right]. \tag{15}$$

C.1.1 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^*.[17]

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.

C.1.2 Proof of Claim 1

This proof is extended from Andrews and Barron (2016), 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 (14) and the first inequality of (15) or else the agent would deviate from $(a_{i,t}, e_{i,t})$ or $\tau_{i,t}$, respectively. The second inequality of (15) 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

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

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

[17] 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.

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 (14).

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 (15) holds. So σ^* is the desired PBE. ■

C.1.3 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 onward. This is clearly an equilibrium that attains continuation surplus $\bar{V}$ from period $t > 0$ onward. ■

C.2 Proof of Proposition 6

Let σ^* be a PBE-sequentially surplus-maximizing equilibrium. By definition, for any $t \geq 0$,

$$E_{\sigma^*}\left[\sum_{i=1}^{N} S_{i,t}\right] = \bar{V}.$$

Suppose $h_\theta^t \in \mathcal{H}_\theta^t$ is a history that occurs on the equilibrium path such that there exist $i, j \in \{1, ..., N\}$ with

$$E_{\sigma^*}\left[\frac{\partial \gamma_i}{\partial d_i}(\theta, d_{i,t})|h_\theta^t\right] > E_{\sigma^*}\left[\frac{\partial \gamma_j}{\partial d_j}(\theta, d_{j,t})|h_\theta^t\right].$$

Define $\tilde{\sigma}$ as the following strategy: at the start of the game, the principal chooses a history h_0^t from the distribution over $\mathcal{H}_0^t$ induced by σ^*, and play continues as in $\sigma^*|h_0^t$. As argued in the proof of Lemma 2, the strategy $\tilde{\sigma}$ can be made a PBE.

Now, consider a strategy profile that is identical to $\tilde{\sigma}$, except in the first period. In that period, after $\theta_0 \in \Theta$ is observed, the principal chooses d_0 so that

$$E_{\sigma^*}\left[\frac{\partial \gamma_i}{\partial d_i}(\theta, d_{i,t})|h_\theta^t\right] = E_{\sigma^*}\left[\frac{\partial \gamma_j}{\partial d_j}(\theta, d_{j,t})|h_\theta^t\right]$$

for all $i, j \in \{1, ..., N\}$. The principal then privately draws a $\tilde{d}_0$ according to $\tilde{\sigma}$, and play continues as if the principal chose $\tilde{d}_0$ in $\tilde{\sigma}$. The decision d_0 only affects the terms $(\gamma_i)_{i=1}^N$ in period 0, so this strategy can also be made a PBE using techniques very similar to those in Lemma 2. But this PBE generates strictly larger surplus than $\tilde{\sigma}$ by construction. So $E_{\tilde{\sigma}}\left[\sum_{i=1}^N S_{i,t}\right] < \bar{V}$, which contradicts the assumption that σ^* is PBE-sequentially surplus-maximizing.

The previous argument proves that if σ^* is PBE-sequentially surplus-maximizing equilibrium, then for any $t \geq 0$ and h_θ^t that occurs on the equilibrium path,

$$E_{\sigma^*}\left[\frac{\partial \gamma_i}{\partial d_i}(\theta, d_{i,t})|h_\theta^t\right] = E_{\sigma^*}\left[\frac{\partial \gamma_j}{\partial d_j}(\theta, d_{j,t})|h_\theta^t\right].$$

In particular, the decision d_t depends only on the payoff-relevant history. In other words, the principal's relationship with each agent is independent of the choices made by other agents, so the problem reduces to a set of N bilateral relational contracts between the principal and each agent. Consequently, efforts in a PBE-sequentially surplus-maximizing equilibrium 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. ■

D For Online Publication: Public Monitoring and Communication

D.1 Statement of Result

This Appendix shows that biased decisions are never surplus-maximizing if monitoring is imperfect but public. We also show that we can replicate equilibrium outcomes from the game with public monitoring in the baseline model, provided that agents can costlessly and immediately communicate with one another.

The **game with public monitoring** is identical to the game in Section 3 with one exception: all variables except e_t are publicly observed, while e_t remains private.[18] Under this monitoring structure, all agents can punish any deviation by the principal, who is therefore willing to pay rewards only if

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

the *sum* of those rewards is smaller than *total* continuation surplus. Biased decisions decrease total continuation surplus 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 7 *In the game with public monitoring, every surplus-maximizing recursive equilibrium is sequentially surplus-maximizing.*

Proof: See Appendix D.2.

The proof of Proposition 7 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. The severity of this punishment depends on total continuation surplus rather than i-dyad surplus. Biased decisions decrease total continuation surplus and make the punishment for a deviation less severe, so they have no place in a surplus-maximizing equilibrium.

Finally, we show that public monitoring outcomes can be replicated in the game with bilateral monitoring so long as agents can communicate with one another. Define the **game with communication** as identical to the model from Section 3, except that agents simultaneously send costless public messages at the end of each period (chosen from a large message space). Then we prove the following result.

Corollary 1 *For any surplus-maximizing equilibrium of the game with public monitoring, there exists a recursive equilibrium of the game with communication that implements the same policy and efforts and generates the same* ex ante *total expected surplus.*

The equilibrium construction in Proposition 7 holds agents to their punishment payoff both on the equilibrium path and after that agent deviates. Therefore, agents are willing to truthfully report what they observe in each period. The collection of these reports reveals the true history, which can then be used to jointly punish the principal following a deviation. Consequently, surplus-maximizing equilibria in the game with communication can be no worse than those in the game with public monitoring. Indeed, they might be even better, since the true history is automatically revealed in the game with public monitoring but not necessarily in the game with communication.

D.2 Proof of Proposition 7

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 $B_i : \phi_0(\mathcal{H}_y^t) \to \mathbb{R}$ satisfying:

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

$$(a_{i,t}, e_{i,t}) \in \arg\max_{a_i, e_i} E_{\sigma^*}\left[B_i(\phi_0(h_y^t)) - (1-\delta)C_i | h_a^t, e_i\right]. \tag{16}$$

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

$$\delta \sum_{i \in I} E_{\sigma^*}\left[\bar{U}_i(h_0^{t+1}) | h_y^t\right] \leq \sum_{i \in I} B_i(\phi_0(h_y^t)) \leq \delta E_{\sigma^*}\left[\sum_{i \in I} U_{i,t+1} + \Pi_{t+1} | h_y^t\right]. \tag{17}$$

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

$$E_{\sigma^*}[U_{j,t+1} | h_d^t] \geq \bar{U}_j(h_d^t). \tag{18}$$

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

$$E_{\sigma^*}\left[\Pi_{t+1} | h_d^t\right] \geq \sum_{i \in I}\left(E_{\sigma^*}\left[B_i(\phi_0(h_y^t)) - (1-\delta)C_{i,t} | h_d^t\right] - E_{\sigma^*}\left[U_{i,t} | h_d^t\right]\right). \tag{19}$$

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 (16). 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 (17).

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 (18) and (19). ■

Completing Proof of Proposition 7

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 are chosen so that $E[U_{i,t}|h_d^t] = \bar{U}_i(h_d^t)$ at every h_d^t on the equilibrium path. Then it is easy to show that $\tilde{\sigma}$ satisfies (16) for the same B_i as σ. $\tilde{\sigma}$ satisfies (19) 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 satisfies $E[U_{i,t}|h_d^t]=\bar{U}_i(h_d^t)$ for every h_d^t that succeeds h_0^t. It is easy to see that $\tilde{\sigma}^*$ satisfies (16)-(19) 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 (18) and (19). Each agent i is willing to choose $a_{i,t}$ and $e_{i,t}$ because $\tilde{\sigma}^*$ satisfies (16). And the principal is willing to pay $\tau_{i,t}^*$ because $\tilde{\sigma}^*$ satisfies (17). 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. ■

D.3 Proof of Corollary 1

Consider augmenting the strategy profile σ^* constructed at the end of the proof of Proposition 7 with the following messages: at the end of each period, each agent reports the outcomes that he observed in that period, except for effort. Collectively, those messages identify a unique history h_0^t. If h_0^t is on the equilibrium path, then the continuation equilibrium is $\sigma^*|h_0^t$. If h_0^t is not on the equilibrium path, then $w_{i,t'} = \tau_{i,t'} = a_{i,t'} = 0$ for every $i \in \{1, ..., N\}$ in all $t' \geq t$, which is a continuation equilibrium because it is an equilibrium of the one-shot game. If exactly one agent lies, the principal can identify that agent because she observes the full history. So the principal chooses $d_{t'}$ to min-max that agent in all $t' \geq t$; if multiple agents lie, then she chooses $d_{t'}$ uniformly at random.

Each agent i earns $\bar{U}_i(h_0^t)$ by reporting truthfully at h_0^t. Suppose agent i lies. If the resulting history $\tilde{h}_0^t$ is on the equilibrium path, then players have the incentive to follow $\sigma^*|\tilde{h}_0^t$ because h_0^t and $\tilde{h}_0^t$ induce the same continuation game and $\sigma^*|\tilde{h}_0^t$ is an equilibrium of that game. So the agent earns $\bar{U}_i(h_0^t)$: he would earn continuation utility $\bar{U}_i(\tilde{h}_0^t)$ at $\tilde{h}_0^t$, and $\bar{U}_i(\tilde{h}_0^t) = \bar{U}_i(h_0^t)$ because $\tilde{h}_0^t$ and h_0^t have the same public history. If $\tilde{h}_0^t$ is off the equilibrium path, then agent i earns $\bar{U}_i(h_0^t)$ because $a_{i,t'} = 0$ in $t' \geq t$ and the principal chooses $d_{t'}$ to min-max him. So i's payoff is independent of his message and so he has no profitable deviation. Therefore, these messages along with the other actions in σ^* form a recursive equilibrium of the game with communication. ■

E For Online Publication: Imperfectly Coordinated Punishments

E.1 Statement of Result

In Appendix D, agents immediately and perfectly coordinate to punish the principal. 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 allows imperfect coordination among agents in the hiring example from Section 4 to argue that biased decisions might remain surplus-maximizing.

In the hiring game, suppose that deviations are $\epsilon-$**uncoordinated:** 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 coordinate their punishments after the principal deviates.

So long as $\epsilon > 0$, Proposition 8 shows that there exist parameter values for which any surplus-maximizing relational contract has a backward-looking policy.

Proposition 8 *Consider the hiring game with ϵ-uncoordinated monitoring. If $\epsilon > 0$, then there exists an open set of* other *parameters (not including ϵ) such that for those parameters, no surplus-maximizing recursive equilibrium is sequentially surplus-maximizing.*

Proof: See Appendix E.2.

Proposition 8 illustrates that, in our hiring example, backward-looking policies might remain surplus-maximizing so long as coordination among agents is not perfect. The intuition for this result 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 $\epsilon > 0$, 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 4, 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 fact that, unlike Lemma 1, the principal may not be willing to implement some policies in equilibrium. However, in the hiring game, the surplus-maximizing policy depends only on the public history. Therefore, deviations from this policy can be jointly punished by all agents.

E.2 Proof of Proposition 8

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 (IC) 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 4, 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. ■

F For Online Publication: Strict Principal Preferences over Decisions

F.1 Statement of Result and Discussion

Our equilibrium construction in Lemma 1 makes the principal indifferent to her on-path decisions. Consequently, (IC) and (DE) are not only necessary conditions for equilibrium, but sufficient as well. This appendix uses the promotions application from Section 2 to show that biased promotions can still be surplus-maximizing even if the principal is required to strictly prefer her equilibrium decision.

Consider the set of **recursive equilibria with strict decisions**, which are recursive equilibria that require the principal to strictly prefer not to deviate from her equilibrium choice of d_1 on the equilibrium path.[19] This additional requirement limits the types of policies that can be implemented in equilibrium. In particular, promotion policies that depend on the *entire vector* of realized outputs are more difficult to implement, since no single agent can observe whether the principal deviates from such a policy. Nevertheless, we prove that a backward-looking policy can be surplus-maximizing in this class of equilibria.

Proposition 9 *In the promotion game in Section 2, suppose $E[y_i|e_i = 0] > 0$, and define $\bar{\delta} < 1$ and $\underline{\gamma}(\cdot)$ as in Proposition 1. There exists a $\tilde{\Delta} : \mathbb{R}_+ \to \mathbb{R}_{++}$ such that if $\delta \in (0, \bar{\delta})$, $\gamma_2 > \underline{\gamma}(\delta)$, and $\gamma_1 - \gamma_2 < \tilde{\Delta}(\delta)$, then there exists a recursive equilibrium with strict decisions that generates strictly higher* ex ante *total expected surplus than any sequentially surplus-maximizing equilibrium.*

Proof: See Appendix F.2.

As in Proposition 1, any sequentially surplus-maximizing equilibrium entails $d_1 = 1$ on the equilibrium path. Both (IC) and (DE) are necessary conditions for equilibrium, so if players are not too patient, then an agent who is never promoted cannot be motivated to work hard in equilibrium. In that case, only agent 1 works hard in any sequentially surplus-maximizing equilibrium.

The proof of Proposition 9 constructs a recursive equilibrium with strict decisions that dominates any sequentially surplus-maximizing equilibrium under the conditions in the statement. The promotion policy in this equilibrium

[19] In general, this constraint means that a surplus-maximizing equilibrium might not exist. To sidestep this problem, Proposition 9 instead states that a recursive equilibrium with strict decisions can strictly improve upon any sequentially surplus-maximizing equilibrium.

depends only on agent 2's output: agent 2 is promoted if and only if y_2 exceeds some threshold. Therefore, agent 2 can observe whether or not the principal follows the equilibrium decision. Since $E[y_i|e_i = 0] > 0$, agent 2 can punish the principal for promoting the wrong agent by rejecting future production. So long as the principal earns strictly positive continuation surplus following her promotion decision, this threat of punishment is enough to give the principal a strict incentive to follow the equilibrium decision. This promotion policy motivates agent 2 to exert effort in $t = 0$, since she is promoted following sufficiently high output. If $\gamma_1 - \gamma_2$ is not too large, then the resulting equilibrium strictly dominates any sequential surplus-maximizing equilibrium.

We draw two conclusions from this analysis. First, requiring the principal to have strict preferences over her decision constrains the kinds of promotion policies that can be implemented in equilibrium. Second, and importantly, biased promotion decisions can remain surplus-maximizing even within this more constrained set of equilibria. Though the details of the surplus-maximizing policy might change, the basic logic of our argument remains the same because (IC) and (DE) are still necessary conditions for equilibrium.

F.2 Proof of Proposition 9

Define $\underline{S} = E[y_i|e_i = 0]$ and $S^B = E[y_i|e_i = 1] - c$. By assumption, $S^B > \underline{S} > 0$. Define $\tilde{\Delta}(\delta) = \min\{\underline{S}, \Delta(\delta)\}$, where $\Delta(\delta)$ is the function from Proposition 1.

Lemma 1 says that (IC) and (DE) are necessary conditions for equilibrium. Consequently, if $\delta < \bar{\delta}$, then the proof of Proposition 1 immediately implies that for any $t \geq 1$, $e_{i,t} = 0$ if $d_1 \neq i$, and moreover $e_{i,0} = 0$ if agent i is never promoted on-path. The requirement that decisions be strict does not constrain continuation equilibria after the principal chooses d_1. Since $\gamma_2 > \underline{\gamma}(\delta)$, whichever agent is promoted can be motivated to exert effort in $t \geq 1$, so the surplus-maximizing continuation equilibrium in $t = 1$ must entail $d_1 = 1$ because $\gamma_1 > \gamma_2$. Consequently, any sequentially surplus-maximizing equilibrium must satisfy $e_{2,t} = 0$ for all $t \geq 0$ and so can generate no more surplus than $S^B + \underline{S} + \delta\gamma_1$.

To prove the result, it suffices to construct an equilibrium with total surplus strictly larger than $S^B + \underline{S} + \delta\gamma_1$. Fix some $q \in (0,1)$ and $\tilde{y}$ satisfying $l(\tilde{y}) > 1$. Consider the following strategy profile:

1. In $t = 0$, $w_{i,0} = 0$ and $a_{i,0} = e_{i,0} = 1$ for $i \in \{1,2\}$.

2. Let $P(\tilde{y}) = \Pr\{y_{2,0} > \tilde{y} | e_{2,0} = 1\}$. Bonuses $\tau_{i,0}$ equal:

$$(1-\delta)\tau_{2,0} = \begin{cases} -\delta(1-q)\underline{S} & l(y_{2,0}) \leq 1 \\ 0 & l(y_{2,0}) > 1 \end{cases}$$

   and

$$(1-\delta)\tau_{1,0} = \begin{cases} -\delta(1-q)\left(P(\tilde{y})\underline{S} + (1-P(\tilde{y}))(S^B + \gamma_1)\right) & l(y_{1,0}) \leq 1 \\ 0 & l(y_{1,0}) > 1 \end{cases}$$

3. $d_1 = 1$ if $l(y_{2,0}) \leq l(\tilde{y})$ and $d_1 = 2$ otherwise.

4. Suppose $d_1 = i$. Then in all $t \geq 1$, play is stationary and satisfies:

   (a) $a_{i,t} = e_{i,t} = 1$, $\tau_{i,t} = -\frac{\delta}{1-\delta}(1-q)(S^B + \gamma_i)$ if $l(y_{i,t}) \leq 1$, and $\tau_{i,t} = \frac{\delta}{1-\delta}q(S^B + \gamma_i)$ if $l(y_{i,t}) > 1$, while $w_{i,t} = E[y_{i,t} - \tau_{i,t} | e_{i,t} = 1] - q(S^B + \gamma_i)$.

   (b) $a_{-i,t} = 1$, $e_{-i,t} = 0$, and $\tau_{-i,t} = 0$, while $w_{-i,t} = E[y_{-i,t} - \tau_{-i,t} | e_{-i,t} = 0] - q\underline{S}$.

5. If the principal and agent i observe a deviation in period $t \geq 0$, then $a_{i,t'} = e_{i,t'} = w_{i,t'} = \tau_{i,t'} = 0$ in $t' \geq t$. If agent 1 observes a deviation in $t = 0$, then the principal follows the equilibrium d_1. If agent 2 or both agents observe a deviation in $t = 0$, then $d_1 = 1$.

For appropriate q and $\tilde{y}$, we argue that this strategy is a recursive equilibrium with strict decisions. Consider $t \geq 1$, after $d_1 = i$ is chosen. The principal's continuation payoff from her relationship with agents i and $-i$ equal $q(S^B + \gamma_i)$

and $q\underline{S}$, respectively, while i's payoff is $(1-q)(S^B+\gamma_i)$ and $-i$'s payoff is $(1-q)\underline{S}$. Therefore, the principal and agent i prefer to pay their respective $\tau_{i,t}$ than renege and earn 0 from the continuation relationship. Since $\gamma_2 > \underline{\gamma}(\delta)$,

$$c \leq \frac{\delta}{1-\delta}\int_{y^*}^{\infty}(S^B+\gamma_i)(p(y_i|1)-p(y_i|0))dy_i$$

and so agent i is willing to choose $e_{i,t}=1$. For $q \in (0,1)$, the principal and agent j both earn strictly positive payoffs from their relationship if they follow $w_{j,t}$ and 0 otherwise. Therefore, players have no profitable deviation following the choice of d_1.

Now, consider the principal's choice of d_1. On the equilibrium path: suppose $y_2 \leq \tilde{y}$, so the equilibrium specifies $d_1 = 1$. If the principal deviates to $d_1 = 2$, then agent 2 observes that deviation, so the principal earns no more than $q\underline{S}$ in the continuation game. If $y_2 > \tilde{y}$ and the principal deviates to $d_1 = 1$, then she earns no more than $q(S^B+\gamma_1)$ in the continuation game. So this deviation is strictly unprofitable for $\gamma_1 - \gamma_2 < \tilde{\Delta}(\delta) \leq \underline{S}$. Under this condition, the principal has a strict incentive to follow the equilibrium decision for any $q > 0$.

Suppose agent 1 observed a deviation in $t = 0$. Agent 2 will observe if the principal deviates from d_1, so the principal earns 0 if she deviates in d_1 and no less than $\delta q\underline{S}$ if she does not. Therefore, the principal cannot profitably deviate in d_1.

If instead agent 2 observed a deviation in $t = 0$, then the principal earns $\delta q(S^B+\gamma_1)$ from choosing $d_1 = 1$ and $\delta\underline{S}$ from choosing $d_1 = 2$, so she has no incentive to deviate from $d_1 = 1$. If both agents observed a deviation in $t = 0$, then the principal earns 0 regardless of d_1, so again has no incentive to deviate from d_1.

Finally, consider actions in $t = 0$. Since $\tau_{1,0}, \tau_{2,0} \leq 0$, the principal cannot profitably deviate from either. If $l(y_{2,0}) \leq 1$, then agent 2 earns 0 by paying $\tau_{2,0}$ and no more than 0 from a deviation, so has no profitable deviation from $\tau_{2,0}$. Similarly, agent 1's expected continuation surplus is $\delta(1-q)\left(P(\tilde{y})\underline{S}+(1-P(\tilde{y}))(S^B+\gamma_1)\right)$ following any y_1, so he is willing to pay $\tau_{1,0}$ rather than earn 0.

Let y^* satisfy $l(y^*) = 1$. Agent 1 is willing to work hard if

$$c \leq \frac{\delta}{1-\delta} \int_{y^*}^{\infty} (1-q) \left(P(\tilde{y})\underline{S} + (1 - P(\tilde{y}))(S^B + \gamma_1) \right) (p(y_i|1) - p(y_i|0))dy_i.$$

Since $\gamma_1 > \underline{\gamma}(\delta)$, this incentive constraint is slack for $q = 0$ and $P(\tilde{y}) \leq \frac{1}{2}$. Agent 2 is willing to work hard if

$$c \leq \frac{\delta}{1-\delta} \left(\int_{y^*}^{\tilde{y}} (1-q)\underline{S}(p(y_2|1) - p(y_2|0))dy_2 + \int_{\tilde{y}}^{\infty} (1-q)(S^B + \gamma_2)(p(y_2|1) - p(y_2|0)) \right) dy_2.$$

Since $p(\cdot|e)$ is strictly MLRP-increasing in e and $\gamma_2 > \underline{\gamma}(\delta)$, this incentive constraint is slack for $q = 0$ and some $\tilde{y}$ such that $P(\tilde{y}) < \frac{1}{2}$. Therefore, for this $\tilde{y}$ and $q > 0$ sufficiently small, both agents are willing to choose $e_{i,0} = 1$. Under those conditions, this strategy profile is a recursive equilibrium with strict decisions.

It remains to show that this recursive equilibrium dominates any sequentially surplus-maximizing equilibrium. Since $\gamma_1 - \gamma_2 < \Delta(\delta)$, the recursive equilibrium with a 50-50 coin flip dominates the sequentially surplus-maximizing equilibrium. But $P(\tilde{y}) < \frac{1}{2}$, so this equilibrium induces the same effort as the coin flip recursive equilibrium while promoting agent 2 strictly less often. *A fortiori,* this equilibrium dominates any sequentially surplus-maximizing equilibrium. ■