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 \textit{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

²Appendix 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 Malcolmson (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)). 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)).

3Many repeated games with public monitoring rule out transfers entirely, leading to equilibrium dynamics. See Goldlücke 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 $\delta$. 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 \(\frac{p(1)}{p(0)}\) 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

\[
\pi_{i,t} = (1 - \delta) \sum_{i=1}^{2} (y_{i,t} + a_{i,t}1_{i,t}\gamma_i - w_{i,t} - \tau_{i,t}) - (1 - \delta)(w_{i,t} + \tau_{i,t} - ca_{i,t}e_{i,t}) 
\]

respectively, where \(\gamma_i\) 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 \(i\),

\[
S_{i,t} = \sum_{t'=t}^{\infty} \delta^{t'-t} (1 - \delta) (y_{i,t} + a_{i,t}1_{i,t}\gamma_i - ca_{i,t}e_{i,t}) .
\]

\(^4\)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 $h^t_0$ 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 $h^t_0 \in \mathcal{H}_0^t$, 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.\footnote{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.} The goal of this section is to give conditions under which every surplus-maximizing equilibrium 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 $\gamma, \Delta : \mathbb{R}_+ \to \mathbb{R}_{++}$ such that if $\delta \in (0, \bar{\delta})$, $\gamma_2 > \gamma(\delta)$, 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: \footnote{Other surplus-maximizing promotion policies differ only after probability-0 events.} there exist $\beta_1, \beta_2 \geq 0$ such that $d_1 = 1$ if and only if

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

(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...
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.\footnote{Abusing notation, we do not write the reward scheme conditional on $a_{i,0}$, which can be (almost) perfectly inferred from $y_{i,0}$.
} Agent $i$ is willing to choose $a_{i,0}^*$ and $e_{i,0}^*$ if and...
only if
\[ a^*_i, e^*_i \in \arg \max_{(a_i, e_i)} \{ E[B_i(y_{i,0})|a_{i,0}, e_{i,0}] - ca_{i,0}e_{i,0} \}. \] \( (2) \)

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

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 \( \delta \) is small enough that no reward scheme satisfies (2) and (3) for \( e^*_{i,t} = 1 \) and \( 1_{i,t} = 0 \). If \( \gamma_2 \) (and hence \( \gamma_1 \)) 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 $i \in \{1, \ldots, N\}$, each “he”) interact repeatedly. Time is discrete and indexed by $t \in \{0, 1, \ldots\}$. 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 \( m_{i,t} \).

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

\[
\begin{align*}
 u_{i,t} &= w_{i,t} + \tau_{i,t} - C_{i,t}, \\
 \pi_t &= \sum_{i=1}^N (g_{i,t} - \tau_{i,t} - w_{i,t}),
\end{align*}
\]

\textsuperscript{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 $\Pi_t$. For each agent $i$, define $i$-dyad surplus in period $t$ as the total continuation surplus produced by that agent,

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

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

**Histories and Strategies** Recall that $h^t_0 \in H^t_0$ is a history at the start of period $t$. For any variable $x$ realized during a period, let $h^t_x$ be a within-period history immediately after that variable is realized, so for example $h^t_y = h^t_0 \cup \{\theta_t, D_t, d_t, w_t, m_t, a_t, e_t, y_t\}$. Then $H^t_x$ is the set of such histories and $H$ is the set of all histories. For every agent $i$, let $\phi_i : H \rightarrow 2^H$ denote agent $i$’s information set, so $\phi_i(h^t_x)$ is the set of histories that $i$ cannot distinguish from $h^t_x$. Similarly, $\phi_0(h^t_x)$ is the principal’s information set. Recall that the principal observes all variables except effort, while agent $i$ observes only $\theta_t, D_t, d_t, y_t$, and variables with subscript $i$. Let $\phi_i(H)$ 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 $\sigma_i$ maps $\phi_i(H)$ 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 $\theta_t$ and $D_t$ to the distribution over decisions taken at that history.

**Equilibrium** We say a recursive equilibrium $\sigma^*$ is **sequentially surplus-maximizing** if, at each on-path $h^t_0 \in H^t_0$, $\sigma^*|_{h^t_0}$ is surplus-maximizing in the continuation game starting at $h^t_0$. If $\sigma^*|_{h^t_0}$ is not surplus-maximizing, then we say that the decisions following $h^t_0$ 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 \( h^t_x \) and \( i \in \{1, ..., N\} \), define agent \( i \)'s punishment payoff as

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

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

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

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

   (IC)

2. **Dynamic enforcement**: for each on-path $h^t_y$,

   $$\delta E_\sigma \left[U_i(h^t_{0}+1) | h^t_{d} \right] \leq B_i(h^t_d, \xi_{i,t}, y_{i,t}) \leq \delta E_\sigma \left[S_{i,t+1} | h^t_d, \xi_{i,t}, y_{i,t} \right].$$

   (DE)

A credible reward scheme satisfies two sets of constraints that depend on continuation play. First, if $\sigma$ specifies that the agent chooses $(a_{i,t}, e_{i,t})$ after observing $\xi_{i,t}$ at history $h^t_d$, 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 $h^t_d$ 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 $h^t_d$ 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 $\sigma^*$ is a recursive equilibrium, then for all $i \in \{1, \ldots, N\}$, there exists a reward scheme $B^*_i$ that is credible in $\sigma^*$.

2. If $\sigma$ is a strategy with a credible reward scheme $B_i$ for each $i \in \{1, \ldots, N\}$, then there exists a recursive equilibrium $\sigma^*$ that induces the same joint distribution over states of the world, decisions, efforts, and outputs as $\sigma$. 

19
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 $\sigma$ and credible reward scheme $B_i$. In each period of our construction, the principal chooses the same decision as in $\sigma$. 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.\footnote{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.} Agent $i$ earns his entire $i$-dyad surplus in each period, but he pays fines following low
output. He is willing to exert effort and make the specified payments because these fines are derived from a credible reward scheme.

3.3 Biased Decisions in Smooth Mean-Shifting Games

Biased decisions can affect equilibrium surplus in three ways. First, they have a \textit{direct cost} because they reduce total continuation surplus. However, if they are biased toward agent $i$ in the sense that they increase $E_{\sigma^*}[S_{i,t+1}|h^t_d, \xi_{i,t}, y_{i,t}]$, then they relax (DE) for $i$. So biased decisions can have an \textit{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 \textit{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.

\textbf{Definition 3} A \textit{smooth mean-shifting game} satisfies:
1. For every $t \geq 0$, $F(\cdot)$ is independent of history, with

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

2. For every $i \in \{1, \ldots, N\}$, $E_i = [0, 1]$ 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 $i \in \{1, \ldots, N\}$, there exists a function $\gamma_i : \Theta \times \mathbb{R} \rightarrow \mathbb{R}$ 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, $\bar{P}_i^L, \bar{P}_i^H$ are smooth distributions with densities $\tilde{p}_i^L$ and $\tilde{p}_i^H$ such that $\tilde{p}_i^H(\cdot)$ 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.\footnote{Since $\frac{\tilde{p}_i^H}{\tilde{p}_i^L}$ is strictly increasing, such distributions satisfy CDPC and strict MLRP. See Rogerson (1985).}

Both $\theta_t$ 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

$$e_{i}^{FB} \equiv \arg \max_{e_i} \{E[x_i|e_i] - c(e_i)\}.$$
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 $\frac{\hat{p}_i^H(\cdot)}{\hat{p}_i^L(\cdot)}$ 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 $\sigma^*$ and any $t \geq 0$, $E_{\sigma^*} \left[ \sum_{i=1}^N d_{i,t} \right] = 1$.
2. There exist $\delta < \delta$ such that if $\delta \in [\delta, \delta]$, 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 $e_i^* \in (0, e_i^{FB})$, $x_i > x_i^*(e_{i,t})$, and $x_{i,t'} < x_j^*(e_{j,t'})$ 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 $\sigma^*$. 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 $\sigma^*$ 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 $\sigma^*$. 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 $h_{t+1}^0$ 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 $\delta < \bar{\delta}$ 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 $h_{t+1}^0$ was surplus-maximizing, then we could slightly increase $E_{\sigma^*}[S_{i,t+1}|h_{t+1}^0]$ and slightly decrease $E_{\sigma^*}[S_{j,t+1}|h_{t+1}^0]$ at a second-order direct cost. By (DE), increasing $E_{\sigma^*}[S_{i,t+1}|h_{t+1}^0]$ 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 $E_{\sigma^*}[S_{j,t+1}|h_{t+1}^0]$ 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.\textsuperscript{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 $d_t = 2$.

\textsuperscript{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 $\delta < \hat{\delta}$ such that if $\delta \in (\hat{\delta}, \delta)$, any surplus-maximizing recursive equilibrium $\sigma^*$ 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

---

14See 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 \( E[S_{i,t}|h_{t0}] \) is strictly positive at every on-path history \( h_{t0} \), 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


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


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


