Online Appendix: Policies in Relational Contracts

Daniel Barron and Michael Powell

For Online Publication: Proofs

A1. 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,

(A1)
$$(a_{i,t}, e_{i,t}) \in \arg\max_{\tilde{a}_{i,t}, \tilde{e}_{i,t}} \left\{ E_y \left[B_i(h_0^t, d_t, w_{i,t}, y_{i,t}) | h_0^t, d_t, w_{i,t}, \tilde{a}_{i,t}, \tilde{e}_{i,t} \right] - (1 - \delta) c \tilde{a}_{i,t} \tilde{e}_{i,t} \right\}$$

and

$$0 \le B_i(h_0^t, d_t, w_{i,t}, y_{i,t}) \le \delta E[S_{i,t+1} | h_0^t, d_t, w_{i,t}, y_{i,t}] \tag{A2}$$

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 (A2) binds following $y_{i,t} < y^*$, and the upper bound binds otherwise. Consequently, an equilibrium exists in which agent i exerts effort in periods $t \ge 1$ if and only if

$$c \le \frac{\delta}{1 - \delta} \int_{y^*}^{\infty} (S^B + 1_{i,t} \gamma_i)(p(y_i | 1) - p(y_i | 0)) dy_i. \tag{A3}$$

[15] Unlike the model in Section II, the principal cannot send messages to her agents when wages are paid in this example. In the proof of Lemma 1, (A1) and (A2) 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.

Let $\bar{\delta}$ satisfy (A3) 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 (A3) 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 $e_{i,0} = 1$:

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

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). $\square$

A2. 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). □

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 $i \in \{1, ..., N\}$,

$$(1-\delta)\tau^*_{i,t} = B_i(h^t_d, \xi_{i,t}, y^*_{i,t}) - \delta E_\sigma \left[S_{i,t+1} | h^t_d, \xi_{i,t}, y^*_{i,t}\right]$$

where agent i infers the right-hand side from $m^*_{i,t}$. Note $\tau^*_{i,t} \leq 0$ by (DE).

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

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^t_0 and $h^{t,*}_0$ 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 $\sigma^* | h^{t,*}_0$ and $\sigma | h^t_0$ by construction.

Deviations by the Principal. — For any on-path $h^{t,*}_d$ and agent $i \in \{1, ..., N\}$, the distribution over $y^*_{i,t}$ is identical to $\sigma | h^t_d$. So

$$E_{\sigma^*}\left[y^*_{i,t} - w^*_{i,t} - \tau^*_{i,t} | h^{t,*}_d\right] = 0$$

and hence $E_{\sigma^*}\left[\Pi_{i,t} | h^{t,*}_d\right] = 0$. 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 $\tilde{h}^{t,*}_d$ such that $E_{\sigma^*}[y_{i,t} - \tilde{w}_{i,t} - \tau_{i,t} | \tilde{h}^{t,*}_d] = 0$ 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 $E_{\sigma^*}\left[U_{i,t+1} | h^{t+1,*}_0\right] = \bar{U}_i(h^{t+1,*}_0) = \bar{U}_i(h^{t+1}_0)$. Off-path, i has no profitable deviation, because $\bar{u}_i(d_t, \theta_t) \geq 0$.

At each on-path $h^{t,*}_0$, 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

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

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

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

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

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

A3. 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 (B3), 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. $\square$

A4. 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.

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

B1. Statement of Result

This appendix extends the analysis in Section II.C 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,*

$$p_i(y_i|\theta, \tilde{d}_i, e_i) = \int_{-\infty}^{\infty} R_i(y_i|x) p_i(x|\theta, d_i, e_i) dx. \tag{B1}$$

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

B2. Proof of Proposition 5

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.

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 6: *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. □

DEFINITION 7: *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*

$$0 = \begin{array}{c} \int_{-\infty}^{y_i^*(\theta,d_i,e_i)} \bar{U}_i(\theta)\frac{\partial p_i}{\partial e_i}(y_i|\theta, \tilde{d}_i, \hat{e}_i)dy_i + \\ \int_{y_i^*(\theta,d_i,e_i)}^{\infty} S_i\left(G_i(y_i|\theta, d_i, \tilde{d}_i, e_i, \hat{e}_i)\right)\frac{\partial p_i}{\partial e_i}(y_i|\theta, \tilde{d}_i, \hat{e}_i)dy_i - c'(\hat{e}_i) \end{array}. \tag{B2}$$

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 (B2) 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)

$$c'(e_{i,t}^*) \leq \begin{array}{c} \int_{-\infty}^{y_i^*(\theta_t,d_{i,t},e_{i,t}^*)} \bar{U}_i(\theta_t) \frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) dy_i + \\ \int_{y_i^*(\theta_t,d_{i,t},e_{i,t}^*)}^{\infty} E_{\sigma^*}[S_i|h_d^t, \xi_{i,t}^*, y_i] \frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) dy_i \end{array}. \tag{B3}$$

Proof of Claim 3. — Suppose $e_{i,t}^* > \min \mathcal{E}_i$ does not satisfy (B3). Because p_i satisfies MLRP and CDFC, we can replace (IC) with its first-order condition as in Rogerson (1985):

$$c'(e_{i,t}^*) = \int_{-\infty}^{\infty} B_i(h_d^t, \xi_{i,t}^*, y_i) \frac{\partial p_i}{\partial e_i}(y_i|\theta_t, d_{i,t}^*, e_{i,t}^*) dy_i. \tag{B4}$$

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

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 $B_i(h^t_d, \xi^*_{i,t}, y_i) = \bar{U}_i(\theta_t)$, and otherwise $B_i(h^t_d, \xi^*_{i,t}, y_i) = E_{\sigma^*}[S_i|h^t_d, \xi^*_{i,t}, y_i]$. But this is exactly the B_i implemented in (B3). Contradiction.

If $e^*_{i,t} = \min \mathcal{E}_i$, then the reward scheme $B_i(h^t_d, \xi^*_{i,t}, y_i) = \bar{U}_i(\theta_t)$ induces $e^*_{i,t}$ because $c(e_{i,t})$ is monotonically increasing. Suppose $e^*_{i,t} > \min \mathcal{E}_i$ satisfies (B3). Clearly, the right-hand side of (B4) is strictly smaller than the left-hand side if $B_i(h^t_d, \xi^*_{i,t}, y_i) = \bar{U}_i(\theta_t)$. The right-hand side of (B4) 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 (B4) is satisfied. □

Claim 4. — Let σ^* be a surplus-maximizing equilibrium, and fix some $(h^t_d, \xi^*_{i,t})$ on the equilibrium path. Define $S_i(y_{i,t}) = E_{\sigma^*}\left[S_{i,t+1}|h^t_d, \xi^*_{i,t}, y_{i,t}\right]$. Without loss, $S_i(y_{i,t})$ is increasing in $y_{i,t}$. Moreover, if (B3) 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 (B3) 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 (B3) for agent i without affecting it for $j \neq i$. So we can assume S_i is increasing without loss.

Suppose (B3) 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 (B3) holding with equality. □

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 (B3) 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 $i, j \in \{1, ..., N\}$,

$$\frac{\partial s_i}{\partial d_i} + \mathbb{I}_{i,t}\frac{\partial s_i}{\partial e_i}\frac{\partial \hat{e}_i}{\partial \tilde{d}_i} = \frac{\partial s_j}{\partial d_j} + \mathbb{I}_{j,t}\frac{\partial s_j}{\partial e_j}\frac{\partial \hat{e}_j}{\partial \tilde{d}_j} \tag{B5}$$

with probability 1 following h_0^t.

PROOF OF CLAIM 5. — Suppose towards contradiction that the left-hand side of (B5) 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 (B3). 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 (B3) was slack in the original equilibrium. But (B3) 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 (B4) 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

$$\text{(B6)} \quad s_i(\theta_t, d_{i,t}+\epsilon, \tilde{e}_{i,t}) + s_j(\theta_t, d_{j,t}-\epsilon, \tilde{e}_{j,t}) - (s_i(\theta_t, d_{i,t}, e_{i,t}) + s_j(\theta_t, d_{j,t}, e_{j,t})) \leq 0$$

Dividing by $\epsilon > 0$, and taking the limit as $\epsilon \to 0$ results in (B5) with a weak inequality $\leq$. Contradiction; we assumed $>$. $\square$

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 (B5) 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 (B3) 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 (B6). But then (B5) 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 $t' \leq t$, (B3) 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, (B3) 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 (B3). 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^*. $\square$

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. $\square$

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^t_0. In a PBE, however, these constraints would condition only on agent i's information set, $\phi_i(h^t_0)$. 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.C1.

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.C2.

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.

C1. Proof of Lemma 2

We first prove an extension of Lemma 1 to PBE.

DEFINITION 8: *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:*

PBE Incentive Constraint: *For each h^t_d, $\xi_{i,t}$, and $(a_{i,t}, e_{i,t})$ on the equilibrium path,*

$$\text{(C1)} \quad (a_{i,t}, e_{i,t}) \in \arg\max_{a_i, e_i} E_\sigma \left[B_i(\phi_i(h^t_d), \xi_{i,t}, y_{i,t}) | \phi_i(h^t_d), \xi_{i,t}, a_i, e_i \right] - (1-\delta) C_i.$$

PBE Dynamic Enforcement: *For each on-path h^t_y,*

$$\text{(C2)} \quad \delta E_\sigma \left[\bar{U}_i(h^{t+1}_0) | \phi_i(h^t_d) \right] \leq B_i(\phi_i(h^t_d), \xi_{i,t}, y_{i,t}) \leq \delta E_\sigma \left[S_{i,t+1} | \phi_i(h^t_d), \xi_{i,t}, y_{i,t} \right].$$

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.

Proof of Claim 1. — This proof is extended from Andrews and Barron (2016), who provide more detail.

[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.

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 (C1) and the first inequality of (C2) or else the agent would deviate from $(a_{i,t}, e_{i,t})$ or $\tau_{i,t}$, respectively. The second inequality of (C2) must hold history-by-history or else the principal would deviate from $\tau_{i,t}$, so *a fortiori* must hold in expectation. □

Part 2:

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

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

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

and the transfer after output y_t^* equals

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

By construction, σ^* implements the same joint distribution over θ_t, d_t, e_t, and y_t as σ. We claim σ^* is a PBE. As in the proof of Lemma 1, the principal earns 0 from each agent i at each history h_0^t on and off the equilibrium path. So the principal has no deviation from σ^*.

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

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

As in Lemma 1, $E_{\sigma^*}\left[U_{i,t+1} | \phi_i(h_d^{t,*}), a_{i,t}, e_{i,t}\right] = E_{\sigma^*}\left[S_{i,t+1} | \phi_i(h_d^{t,*}), a_{i,t}, e_{i,t}\right]$. Furthermore, it can be shown that for every agent i, σ^* induces a coarser information partition over histories than σ: if h_0^t, $h_0^{t,*}$ and $\tilde{h}_0^t, \tilde{h}_0^{t,*}$ are two pairs of histories from the construction of σ^*, then $\phi_i(h_0^{t,*}) = \phi_i(\tilde{h}_0^{t,*})$ whenever $\phi_i(h_0^t) = \phi_i(\tilde{h}_0^t)$. Therefore, $E_{\sigma^*}\left[S_{i,t+1} | \phi_i(h_d^{t,*}), a_{i,t}, e_{i,t}\right] = E_\sigma\left[S_{i,t+1} | \phi_i(h_d^t), a_{i,t}, e_{i,t}\right]$. Plugging these expressions into agent i's IC constraint yields (C1).

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

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

This constraint is satisfied because (C2) holds. So σ^* is the desired PBE. $\square$

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. $\square$

C2. *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. □

For Online Publication: Public Monitoring and Communication

D1. 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 II 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 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.D2.

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

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

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 II, 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.

D2. 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,

(D1) $$(a_{i,t}, e_{i,t}) \in \arg\max_{a_i, e_i} E_{\sigma^*}\left[B_i(\phi_0(h_y^t)) - (1-\delta)C_i | h_a^t, e_i\right].$$

2) **Public Dynamic Enforcement:** for any $I \subseteq \{1, ..., N\}$ and h_y^t,
(D2)
$$\delta \sum_{i\in I} E_{\sigma^*}\left[\bar{U}_i(h_0^{t+1}) | h_y^t\right] \leq \sum_{i\in I} B_i(\phi_0(h_y^t)) \leq \delta E_{\sigma^*}\left[\sum_{i\in I} U_{i,t+1} + \Pi_{t+1} | h_y^t\right].$$

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

(D3)
$$E_{\sigma^*}[U_{j,t+1}|h_d^t] \geq \bar{U}_j(h_d^t).$$

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

(D4)
$$E_{\sigma^*}\left[\Pi_{t+1}|h_d^t\right] \geq \sum_{i \in I}\left(E_{\sigma^*}\left[B_i(\phi_0(h_y^t)) - (1-\delta)C_{i,t}|h_d^t\right] - E_{\sigma^*}\left[U_{i,t}|h_d^t\right]\right).$$

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 (D1). 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 (D2).

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 (D3) and (D4). □

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 (D1) for the same B_i as σ. $\tilde{\sigma}$ satisfies (D4) 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 (D1)-(D4) 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 (D3) and (D4). Each agent i is willing to choose $a_{i,t}$ and $e_{i,t}$ because $\tilde{\sigma}^*$ satisfies (D1). And the principal is willing to pay $\tau_{i,t}^*$ because $\tilde{\sigma}^*$ satisfies (D2). 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. $\square$

D3. 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. □

For Online Publication: Imperfectly Coordinated Punishments

E1. 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 III 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.E2.

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 III, 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.

E2. 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 $d_t = 2$.

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

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. □

For Online Publication: Strict Principal Preferences over Decisions

F1. 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 I 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.

[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.

PROPOSITION 9: *In the promotion game in Section I, 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.F2.

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 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.

F2. 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. $\square$