Online Appendix: "The Allocation of Future Business," Isaiah Andrews and Daniel Barron[11]

Omitted Proofs

Proof of Lemma 1

We begin this proof with a lemma that bounds each player's punishment payoff from below.

Lemma A.1

For relational contract σ^* and on-path history h^t, consider $h^{t+1}, \tilde{h}^{t+1} \in \mathcal{H}_0^{t+1}$ such that $h^{t+1} \in \text{supp}\{\sigma^*|h^t\}$ but $\tilde{h}^{t+1} \notin \text{supp}\{\sigma^*|h^t\}$, for $h^{t+1} \in \text{supp}\{\sigma^*|h^t\}$ the support of h^{t+1} conditional on h^t under σ^*. Suppose $I_j(h^{t+1}) = I_j(\tilde{h}^{t+1})$ $\forall j \notin \{0, i\}$. Then

$$E_{\sigma^*}\left[U_{i,t+1}|I_i(\tilde{h}^{t+1})\right] \geq 0 \tag{7}$$

$$E_{\sigma^*}\left[U_{0,t+1}|\tilde{h}^{t+1}\right] \geq E_{\sigma^*}\left[\sum_{j\neq i}\sum_{t'=t+1}^{\infty}\delta^{t'-t-1}(1-\delta)(1_{j,t'}y_{t'} - w_{j,t'} - \tau_{j,t'})|h^{t+1}\right]. \tag{8}$$

Proof of Lemma A.1

If (7) were not satisfied, then agent i could profitably deviate by paying no transfers and choosing $d_t = 0$.

If (8) is not satisfied, then we claim the principal has a profitable deviation at $\tilde{h}^{t+1}$. Consider the following recursively-defined deviation following $\tilde{h}^{t+1}$. For any $h^{t+t'}$, $t' \geq 1$, the principal plays $\sigma_0^*(h^{t+t'})$, except he pays no transfers to agent i. Let $\tilde{h}^{t+t'+1}$ be the observed history at the beginning of $t + t' + 1$. The principal chooses $h^{t+t'+1}$ according to the distribution of length $t + t' + 1$

[11] We frequently refer to "all histories on the equilibrium path" such that some condition holds. Formally, interpret "all histories on the equilibrium path" as "almost surely on the equilibrium path."

histories induced by $\sigma^*(h^{t+t'})$, conditional on the event $I_j(\tilde{h}^{t+t'+1}) = I_j(h^{t+t'+1})$ $\forall j \neq i$. Under this strategy, agents $j \neq i$ cannot distinguish $\tilde{h}^{t+t'}$ and $h^{t+t'}$ for any $t' \geq 1$, so the principal earns at least the right-hand side of (8). $\square$

Proof of Lemma 1, Statement 1

Fix two histories $h_H^t, h_L^t \in \mathcal{H}_y^t$ whose sole difference is that $y_t = v_{x_t,t}$ in h_H^t and $y_t = 0$ in h_L^t. Define $R_i(j) = E_{\sigma^*}\left[(1-\delta)\tau_{i,t} + \delta U_{i,t+1} | h_j^t\right]$ for $j \in \{L, H\}$. On the equilibrium path, agent i chooses $e_t = 1$ only if

$$p_1 R_i(H) + (1-p_1) R_i(L) - (1-\delta)c \geq p_0 R_i(H) + (1-p_0) R_i(L). \qquad (9)$$

By Lemma A.1, $\tau_{i,t}$ is paid at $h^t \in \mathcal{H}_y^t$ in equilibrium only if

$$\begin{gathered}(1-\delta)E_{\sigma^*}\left[\tau_{i,t}|h^t\right] \leq \delta E_{\sigma}\left[\sum_{t'=t+1}^{\infty} \delta^{t'-t-1}(1-\delta)1_{i,t'}(y_{t'} - ce_{t'})|h^t\right] \; . \\ -(1-\delta)E_{\sigma^*}\left[\tau_{i,t}|I_i(h^t)\right] \leq \delta E_{\sigma^*}\left[U_{i,t+1}|I_i(h^t)\right]\end{gathered}$$

Otherwise, either the principal or agent i would deviate by not paying $\tau_{i,t}$. Plugging these constraints into (9) yields (1). $\square$

Proof of Lemma 1, Statement 2

We construct relational contract σ^* given a strategy σ that satisfies (1). For any strategy profile $\hat{\sigma}$, define $\mathcal{H}(\hat{\sigma}) \subseteq \mathcal{H}$ as the set of on-path histories. Define an *augmented history* as an element of $\mathcal{H} \times \mathcal{H}$. We will denote histories corresponding to σ or σ^* by h_a^t or $h_a^{t,*}$, respectively.

Constructing equilibrium strategies: We recursively construct a candidate equilibrium σ^*. The construction begins with an augmented history consisting of two histories at the start of the game, $(h_0^1, h_0^{1,*})$, where $h_0^1 = h_0^{1,*}$.

1. For any $t' \geq t$, let $(h_0^t, h_0^{t,*}) \in \mathcal{H}^{Aug}$ with $h_0^t, h_0^{t,*} \in \mathcal{H}_0^t$.
2. For each $h_v^{t,*} \in \mathcal{H}_v^t$ in the support of $\sigma^*|h_0^{t,*}$, let h_v^t be the corresponding successor to h_0^t with the same productivity realizations. The principal

chooses an h_e^t according to $\sigma|h_v^t$. The principal chooses x_t^* as in h_e^t, and pays $w_{i,t}^* = 0\ \forall i \neq x_t^*$, with

$$w_{x_t,t}^* = \left[(1-e_t)v_{x_t,t}p_0 + e_t\left(v_{x_t,t}p_1 + (1-p_1)\frac{\delta}{1-\delta}E_\sigma\left[S_{i,t+1}|I_i(h_e^t), y_t = 0\right]\right)\right]. \tag{10}$$

3. Agent x_t^* chooses $d_t^* = 1$ iff $w_{x_t,t}^* \geq v_{x_t,t}p_0$ and $e_t^* = 1$ iff $w_{x_t,t}^* \geq v_{x_t,t}p_1$.

4. If $e_t = 1$ and $y_t = 0$, then $\tau_{x_t,t}^* = -\frac{w_{x_t,t}^* - v_{x_t,t}p_1}{1-p_1}$. Otherwise, $\tau_{x_t,t}^* = 0$. For all agents $i \neq x_t^*$, $\tau_{i,t}^* = 0$.

5. Let $h_0^{t+1,*} \in \mathcal{H}_0^{t+1}$ be the realized history after following these steps. Let $h_y^t \in \mathcal{H}_y^t$ be the successor to h_e^t with the same y_t^* as in $h_0^{t+1,*}$. The principal chooses $h_0^{t+1} \in \mathcal{H}_0^{t+1}$ according to the conditional distribution $\sigma|h_y^t$. Repeat with the new histories $(h_0^{t+1}, h_0^{t+1,*})$.

6. If a deviation occurs in any variable except e_t, the principal thereafter chooses $x_t = 1$ and pays no transfers. Agents who observe a deviation choose $d_t = e_t = 0$ and pay no transfers.

Payoff Equivalence of σ and σ^*: We claim that σ and σ^* generate the same distribution over $\{\mathbf{v}_s, x_s, y_s\}_{s=1}^t$, $\forall t \geq 1$. By induction: the result is immediate for $t = 1$. If it holds for t, let $(h_0^t, h_0^{t,*})$ be the augmented history at the start of period $t+1$ in the recursive construction. F is exogenous. Conditional on $\mathbf{v}_t$, for $p_0 > 0$ actions (x_t^*, d_t^*, e_t^*) are drawn as in $\sigma|h_v^t$, while for $p_0 = 0$ the analogous statement holds for $(x_t^*, d_t^* e_t^*)$. In either case, the distribution over output is the same in σ and σ^*. Finally, the augmented history $(h_0^{t+1}, h_0^{t+1,*})$ is drawn from the distribution $\sigma|h_y^t$ over h_0^{t+1}. So the conditional distribution given by $\sigma^*|(h_0^t, h_0^{t,*})$ over period $t+1$ augmented histories is the same as the conditional distribution $\sigma|h_0^t$ over $\mathcal{H}_0^{t+1}$. The distribution over (x_t^*, d_t^*, e_t^*) is also identical in σ and σ^*, which proves the claim.

We immediately conclude that for any $(h_0^t, h_0^{t,*})$ in the construction, $\sigma|h_0^t$ and $\sigma^*|h_0^t$ have the same expected total continuation surplus.

The principal has no profitable deviation: Under σ^*, the principal earns 0 in each period. In period t, the principal could deviate from σ^* in x_t^*, $\{w_{i,t}^*\}_{i=1}^N$, or $\{\tau_{i,t}^*\}_{i=1}^N$. Following the deviation, the principal earns 0 in all future periods. Since $\tau_{i,t}^* \leq 0$, the principal has no profitable deviation in $\{\tau_{i,t}^*\}_{i=1}^N$. A deviation from $w_{i,t}^* = 0$ for $i \neq x_t$ would be similarly unprofitable. Following a deviation in $w_{x_t,t}^*$, the principal earns $y_t - w_{x_t,t} - \tau_{x_t,t}$. If $w_{x_t,t} < v_{x_t,t}p_0$, then $d_t^* = 0$ and the principal's payoff equals 0. If $w_{x_t,t} = v_{x_t,t}p_0$, then either agent x_t detects a deviation ($y_t = \tau_{x_t,t} = 0$), or he doesn't and chooses $d_t^* = 1$, $e_t^* = 0$. In either case, the principal's payoff is weakly negative. If $w_{x_t,t}$ satisfies (10) and is not detected as a deviation, then agent x_t chooses $d_t^* = e_t^* = 1$ and pays $\tau_{x_t,t}^*$, so the principal earns 0. Any other $w_{x_t,t}$ is detected as a deviation and so is unprofitable. Since the principal can never earn more than 0 for any allocation decision, she has no profitable deviation from x_t^*. So the principal has no profitable deviation.

Each agent has no profitable deviation: We must show that an agent cannot profitably deviate from d_t^*, e_t^*, and $\tau_{i,t}^* < 0$.

As a first step, suppose $(h^t, h^{t,*})$ and $(\hat{h}^t, \hat{h}^{t,*})$ are two augmented histories from the construction. Suppose $I_i(h^t) = I_i(\hat{h}^t)$. Then we claim $I_i(h^{t,*}) = I_i(\hat{h}^{t,*})$. For $t = 1$ the result trivially holds. Suppose it holds for all histories of length $t-1$, and suppose towards contradiction that $I_i(h^{t,*}) \neq I_i(\hat{h}_0^{t,*})$. Then there exists some variable in period $t-1$ that differs between $h^{t,*}$ and $\hat{h}^{t,*}$ and is observed by agent i. All non-transfer variables are the same in $h^{t,*}$ and $\hat{h}^{t,*}$ because they are the same in h^t and $\hat{h}^t$. Transfer $\tau_{i,t}^*$ depends only on $w_{i,t}^*$, and $w_{i,t}^*$ depends only on non-transfer variables and $I_i(h_e^{t-1})$, where $I_i(h_e^{t-1}) = I_i(\hat{h}_e^{t-1})$ by assumption. Therefore, $I_i(h^{t,*}) = I_i(\hat{h}^{t,*})$ as desired.

Now, suppose the agent deviates from $\tau_{i,t}^*$. If $\tau_{i,t}^* = 0$, then such a deviation cannot be profitable. Suppose $\tau_{i,t}^* < 0$. By the proof that σ and σ^* are payoff equivalent,

$$E_\sigma\left[S_{i,t+1}|h_0^{t+1}\right] = E_{\sigma^*}\left[\sum_{t'=t+1}^{\infty} \delta^{t'-t-1}(1-\delta)1_{i,t'}(y_{t'} - ce_{t'})|(h_0^{t+1}, h_0^{t+1,*})\right]$$

for an augmented $(h_0^{t+1}, h_0^{t+1,*})$ from the construction of σ^*. Since the prinicpal earns 0 in each period of σ^*,

$$E_{\sigma^*}\left[\sum_{t'=t+1}^{\infty} \delta^{t'-t-1}(1-\delta)u_i^{t'}|(h_0^{t+1}, h_0^{t+1,*})\right] = E_\sigma\left[S_{i,t+1}|h_0^{t+1}\right]. \qquad (11)$$

Consider a history $h_y^{t,*}$ such that $\tau_{i,t}^* < 0$ (so $e_t^* = 1$ and $y_t^* = 0$). Then

$$E_{\sigma^*}\left[U_{i,t+1}|I_i(h_y^{t,*})\right] = E_{\sigma^*}\left[S_{i,t+1}|(\hat{h}_y^t, \hat{h}_y^{t,*}) \text{ s.t. } I_i(\hat{h}_y^{t,*}) = I_i(h_y^{t,*})\right] = .$$
$$E_{\sigma^*}\left[E_\sigma\left[S_{i,t+1}|I_i(\hat{h}_y^t)\right]|(\hat{h}_y^t, \hat{h}_y^{t,*}) \text{ s.t. } I_i(\hat{h}_y^{t,*}) = I_i(h_y^{t,*})\right] = E_\sigma\left[S_{i,t+1}|I_i(h_y^t)\right]$$

The first of these equalities holds by (11). The second equality holds for two reasons: (a) σ^* induces a coarser partition over agent information sets than σ by the previous argument, and (b) by construction, the distribution induced by σ^* over $(h^t, h^{t,*})$ is the same as the distribution induced by σ over h^t.[12] The final equality holds because if $I_i(\hat{h}_y^{t,*}) = I_i(h_y^{t,*})$, then $\hat{w}_{i,t}^* = w_{i,t}^*$ in these two histories. But then $E_\sigma\left[S_{i,t+1}|I_i(h_e^t), y_t = 0\right] = E_\sigma\left[S_{i,t+1}|I_i(\hat{h}_e^t), y_t = 0\right]$ by the definition of $w_{i,t}^*$.

Following a deviation, $U_{i,t+1} = 0$. Therefore, agent i has no profitable deviation if

$$-(1-\delta)\tau_{i,t}^* \leq \delta E_\sigma\left[S_{i,t+1}|I_i(h_e^t), y_t = 0\right],$$

which holds (with equality) by construction.

Finally, we argue that agent i has no profitable deviation from d_t^* or e_t^*. If $d_t^* = 0$ or $e_t^* = 0$, the result follows immediately. Agent i has no profitable deviation from $e_t^* = 1$ if his IC constraint (9) holds. Given $\tau_{i,t}^*$, this constraint may be written

$$\delta E_{\sigma^*}\left[U_{i,t+1}|I_i(h_e^{t,*}), y_t^* > 0\right] \geq (1-\delta)\frac{c}{p_1 - p_0}.$$

But $E_{\sigma^*}\left[U_{i,t+1}|I_i(h_e^{t,*}), y_t > 0\right] = E_\sigma\left[S_{i,t+1}|I_i(h_e^t), y_t > 0\right]$ by the argument above.

[12] Note that every on-path h^t in σ corresponds to a unique augmented history $(h^t, h^{t,*})$ in the construction.

So this inequality holds by (1).

We conclude that σ^* is a relational contract that satisfies the conditions of Lemma 1. $\square$

Lemma A.2

Statement of Lemma A.2

For Ω_t^{FSA} and $\tilde{\Omega}_t$ as defined in the proof of Lemma 3, $\Omega_t^{FSA} \leq \tilde{\Omega}_t$ for all t.

Proof of Lemma A.2

Define $b_i^L(h^t, k) = 1_{i,t}(k) 1\{i \notin H(h^t)\}$, and let $\beta_{i,t}^L(k) = E_\sigma\left[b_i^L(h^t, k)\right]$. For any strategy which both implies first-best surplus and treats agents symmetrically (from an ex-ante perspective), however, $\beta_t^L(k) + \beta_t^H(k) \equiv \frac{1}{N} F^k$ for $F^k = Pr\left\{v_{\max,t} = v^k\right\}$ (where we again use symmetry to drop the i subscript). Thus we can thus re-write obligation as

$$\Omega_t = \sum_{s=1}^{t} \delta^{s-t} \sum_{k=1}^{K} \beta_s^L(k) \left(p_1 \delta \tilde{S} + (1-\delta) V_k^{FB}\right) - (1-\delta) \sum_{s=1}^{t} \delta^{s-t} \sum_{k=1}^{K} \frac{1}{N} F^k V_k^{FB}.$$

Note that (for fixed model parameters) obligation is entirely determined by the sequence $\left\{\left\{\beta_t^L(k)\right\}_{k=1}^{K}\right\}_{t=1}^{\infty}$, which we will denote by $\mathcal{B}$. Letting $\tilde{\mathcal{B}}$ and $\mathcal{B}^{FSA}$ be the sequences implied by $\tilde{\sigma}$ and σ^{FSA}, with elements $\tilde{\beta}_t^L(k)$ and $\beta_t^{L,FSA}(k)$ respectively (again using symmetry to drop the i subscript), we prove the lemma by constructing $\mathcal{B}^n$ for $n = 0, 1, 2, ...$ such that $\mathcal{B}^1 = \tilde{\mathcal{B}}$ and Ω_t^n, the obligation implied by $\mathcal{B}^n$, is decreasing in n with $\lim_{n\to\infty} \Omega_t^n = \Omega_t^{FSA}$. To formally define $\mathcal{B}^n$, for each $n \geq 1$ set $\beta_t^{L,n}(k) = \beta_t^{L,FSA}(k)$ for $t < n$, and $\beta_t^{L,n}(k) = \tilde{\beta}_t^L(k)$ for $t > n$. For $t = n$, select $\beta_n^{L,n}(k)$ such that

$$\beta_n^{L,n}(k) \geq \max\left\{\beta_n^{L,FSA}(k), \tilde{\beta}_n^L(k)\right\} \text{ for all } k$$

and

$$\sum_{t=1}^{\infty} \sum_{k=1}^{K} \beta_t^{L,n}(k) \equiv \frac{1}{p_1}. \qquad (12)$$

Lemma A.3 establishes that such selections are always feasible. The constraint (12) is a summing-up constraint on the allocation probabilities, and one can show that any sequence $\mathcal{B}$ generated by a strategy profile σ yielding first-best surplus (including $\tilde{\sigma}$ and σ^{FB}) must satisfy this constraint. Note, however, that there are other constraints on the set of sequences $\mathcal{B}$ which may be generated by first-best strategy profiles which we do not require of $\mathcal{B}^n$, so it will not in general be the case that $\mathcal{B}^n$ corresponds to any strategy profile, though we can still calculate the implied obligation.

We need to show that for any n, t pair, $\Omega_t^n \geq \Omega_t^{n+1}$. Note that $\beta_s^{L,n}(k) = \beta_s^{L,n+1}(k)$ for $s < n$, and $\beta_n^{L,n}(k) \geq \beta_n^{L,n+1}(k)$ by construction, so since Ω_t is increasing in $\beta_s^L(k)$ for all s and k the result is immediate for $t \leq n$. Hence, it remains only to address the case with $t > n$. In this case, note that $\beta_s^{L,n}(k)$ and $\beta_s^{L,n+1}(k)$ still coincide for all $s \notin \{n, n+1\}$, so the difference in obligations is

$$\Omega_t^n - \Omega_t^{n+1}$$
$$= \delta^{n-t}\left(\sum_{k=1}^{K}\left(p_1\delta\tilde{S} + (1-\delta)V_k^{FB}\right)\left(\beta_n^{L,n}(k) - \beta_n^{L,n+1}(k) + \delta\beta_{n+1}^{L,n}(k) - \delta\beta_{n+1}^{L,n+1}(k)\right)\right).$$

We know by construction of $\mathcal{B}^n$ that

$$\beta_n^{L,n}(k) - \beta_n^{L,n+1}(k) = \beta_n^{L,n}(k) - \beta_n^{L,FSA}(k) \geq 0$$

while

$$\beta_{n+1}^{L,n}(k) - \beta_{n+1}^{L,n+1}(k) = \tilde{\beta}_{n+1}^{L}(k) - \beta_{n+1}^{L,n+1}(k) \leq 0$$

and by (12) and the fact that $\beta_s^{L,n}(k)$ and $\beta_s^{L,n+1}(k)$ coincide for $s \notin \{n, n+1\}$

$$\sum_{k=1}^{K}\left(\beta_n^{L,n}(k) - \beta_n^{L,n+1}(k)\right) = -\sum_{k=1}^{K}\left(\beta_{n+1}^{L,n}(k) - \beta_{n+1}^{L,n+1}(k)\right) = \Delta_n$$

for some value $\Delta_n \geq 0$. Thus, we can bound $\Omega_t^n - \Omega_t^{n+1}$ from below by assuming that $\beta_n^{L,n}(1) - \beta_n^{L,n+1}(1) = \Delta_n$, $\beta_{n+1}^{L,n}(K) - \beta_{n+1}^{L,n+1}(K) = -\Delta_n$, and all other differences are equal to zero. This yields

$$\Omega_t^n - \Omega_t^{n+1} \geq \delta^{n-t}\left(\left(p_1\delta\tilde{S} + (1-\delta)V_1^{FB}\right)\Delta_n - \Delta_n\delta\left(p_1\delta\tilde{S} + (1-\delta)V_K^{FB}\right)\right)$$

which is greater than zero if and only if $\delta V_K^{FB} \leq V_1^{FB} + p_1\delta\tilde{S}$, which is precisely the condition assumed in the statement of Lemma 3.□

Proof of Lemma A.3

Statement of Lemma A.3

For all n, we can select $\mathcal{B}^n$ such that $\beta_t^{L,n}(k) = \beta_t^{L,FSA}(k)$ for $t < n$, $\beta_t^{L,n}(k) = \tilde{\beta}_t^L(k)$ for $t > n$, $\beta_n^{L,n}(k) \geq \max\left\{\beta_n^{L,FSA}(k), \tilde{\beta}_n^L(k)\right\}$, and $\sum_{t=1}^{\infty}\sum_{k=1}^{K}\beta_t^{L,n}(k) = \frac{1}{p_1}$.

Proof of Lemma A.3

Note that

$$\beta_n^{L,n}(k) \geq \max\left\{\beta_n^{L,FSA}(k), \tilde{\beta}_n^L(k)\right\}$$

if and only if

$$\beta_n^{L,n}(k) - \tilde{\beta}_n^L(k) \geq \max\left\{\beta_n^{L,FSA}(k) - \tilde{\beta}_n^L(k), 0\right\}.$$

Since our choice of $\beta_n^{L,n}(k)$ is restricted only by the adding-up constraint (12), we are able to choose such $\beta_n^{L,n}(k)$ for all k if and only if we can ensure that

$$\sum_{k=1}^{K}\left(\beta_n^{L,n}(k) - \tilde{\beta}_n^L(k)\right) \geq \sum_{k=1}^{K}\max\left\{\beta_t^{L,FSA}(k) - \tilde{\beta}_t^L(k), 0\right\}. \quad (13)$$

Since $\beta_t^{L,n}(k)$ and $\tilde{\beta}_t^L(k)$ coincide for $t > n$, (12) and its counterpart for $\tilde{\mathcal{B}}$ imply that

$$\sum_{t=1}^{n}\sum_{k=1}^{K}\left(\beta_t^{L,n}(k) - \tilde{\beta}_t^L(k)\right) = 0$$

which since $\beta_t^{L,n}(k)$ and $\beta_t^{L,FSA}(k)$ coincide for $t<n$, implies that

$$\sum_{k=1}^{K}\left(\beta_n^{L,n}(k)-\tilde{\beta}_n^{L}(k)\right)=\sum_{t=1}^{n-1}\sum_{k=1}^{K}\left(\tilde{\beta}_t^{L}(k)-\beta_t^{L,FSA}(k)\right).$$

Thus, we can re-write (13) as

$$\sum_{t=1}^{n-1}\sum_{k=1}^{K}\left(\tilde{\beta}_t^{L}(k)-\beta_t^{L,FSA}(k)\right)\geq\sum_{k=1}^{K}\max\left\{\hat{\beta}_n^{L}(k)-\tilde{\beta}_n^{L}(k),0\right\}.$$

This statement holds if and only if

$$p_1\sum_{t=1}^{n-1}\sum_{k=1}^{K}\left(\tilde{\beta}_t^{L}(k)-\hat{\beta}_t^{L,FSA}(k)\right)+p_1\sum_{k=1}^{K}\min\left\{\tilde{\beta}_n^{L}(k)-\hat{\beta}_n^{L,FSA}(k),0\right\}\geq 0. \tag{14}$$

Note that

$$p_1\beta_t^{L}(k)=Pr_{\sigma}\left\{\begin{array}{c}\text{agent } i \text{ produces high output for the}\\ \text{first time in round } t\text{, and the output type is } v^k\end{array}\right\},$$

where the events inside the probability are disjoint for different t and/or k. Thus, for $\mathcal{K}$ any subset of $\{1,...,K\}$ and

$$C(n,\mathcal{K})=\left\{\begin{array}{c}\text{agent } i \text{ produces high output in first } n-1 \text{ rounds,}\\ \text{or produces high output } v^k \text{ for } k\in\mathcal{K} \text{ in round } n\end{array}\right\}$$

we have that

$$Pr_{\sigma^{FSA}}\{C(n,\mathcal{K})\}=p_1\sum_{t=1}^{n-1}\sum_{k=1}^{K}\beta_t^{L,FSA}(k)+p_1\sum_{k\in\mathcal{K}}\beta_n^{L,FSA}(k).$$

Note, however, that

$$p_1\sum_{k}\min\left\{\tilde{\beta}_n^{L}(k)-\beta_n^{L,FSA}(k),0\right\}=\min_{\mathcal{K}\subseteq\{1,\ldots,K\}}p_1\sum_{k\in\mathcal{K}}\left(\tilde{\beta}_n^{L}(k)-\beta_n^{L,FSA}(k)\right).$$

Thus, (14) is equivalent to

$$Pr_{\sigma^{FSA}}\left\{C\left(n,\mathcal{K}\right)\right\} \leq Pr_{\tilde{\sigma}}\left\{C\left(n,\mathcal{K}\right)\right\} \text{ for all } \mathcal{K}\subseteq\{1,...,K\}$$

and it suffices to show that σ^{FSA} minimizes $Pr_{\sigma}\left\{C\left(n,\mathcal{K}\right)\right\}$ over strategies σ.

In particular, for symmetric strategy profile σ define $\gamma_t\left(m\right) = Pr_{\sigma}\left\{|H(h^t)| = m\right\}$ to be the probability that exactly m agents have produced high output before period t under σ, and let

$$\psi_t\left(m,k\right) = Pr_{\sigma}\left\{x_t \in H\left(h^t\right), v_{\max,t} = v^k \| H(h^t)| = m\right\}$$

to be the probability that the principal chooses an agent who has already produced high output conditional on this event, decomposed by productivity type. These probabilities have a direct connection to β_t^L under σ:

$$\beta_t^L = \frac{1}{N}\sum_{m=0}^{N-1}\gamma_t\left(m\right)\sum_{k=1}^{K}\left(F^k - \psi_t\left(m,k\right)\right).$$

Note, in addition, that for any strategy profile yielding first-best surplus,

$$\psi_t\left(m,k\right) \leq Pr_{\sigma}\left\{\mathcal{M}_t \cap H\left(h^t\right) \neq \emptyset, v_{\max,t} = v^k \| H\left(h^t\right)| = m\right\},$$

since otherwise σ must allocate production to inefficient agents with positive probability.[13] Finally, note that we can define $\gamma_t\left(m\right)$ recursively in terms of $\psi_t\left(m,k\right)$, since $\gamma_t\left(0\right) = \left(1-p_1\right)\gamma_{t-1}\left(0\right)$, while for $m \geq 1$

$$\gamma_t\left(m\right) = \sum_{k=1}^{K}\left(\begin{array}{c}\psi_{t-1}\left(m,k\right)\gamma_{t-1}\left(m\right) + (F^k - \psi_{t-1}\left(m,k\right))(1-p_1)\gamma_{t-1}\left(m\right) \\ +(F^k - \psi_{t-1}\left(m-1,k\right))p_1\gamma_{t-1}\left(m-1\right)\end{array}\right).$$

Using these new definitions, note that

$$\begin{array}{c}Pr_{\sigma}\left\{C\left(n,\mathcal{K}\right)\right\} = \frac{1}{N}\sum_{m=0}^{N} m\gamma_n\left(m\right) + p_1\sum_{k\in\mathcal{K}}\beta_{i,n}^L\left(k\right) \\ = \frac{1}{N}\sum_{m=0}^{N} m\gamma_n\left(m\right) + \frac{p_1}{N}\sum_{k\in\mathcal{K}}\sum_{m=0}^{N-1}\gamma_n\left(m\right)\left(F^k - \psi_n\left(m,k\right)\right).\end{array} \quad (15)$$

[13]Technically, this is not a problem if $\gamma_t\left(m\right) = 0$, but in this case $\psi_t\left(m,k\right)$ has no effect on obligation and may be chosen arbitrarily.

Note further that $\sum_{l=0}^{N}\gamma_n(l)\equiv 1$ by construction, so $\sum_{l=0}^{N}\frac{\partial\gamma_n(l)}{\partial\psi_s(m,k)}=0$, while for $s\leq n$,

$$\frac{\partial\gamma_n(l)}{\partial\psi_s(m,k)}=\begin{cases}0 & \text{if } l<m\\ >0 & \text{if } l=m\\ <0 & \text{if } l>m\end{cases}.$$

Since $0\leq\sum_k\left(F^k-\psi_t(m,k)\right)\leq 1$ for all m by definition, $m+p_1\sum_k\left(F^k-\psi_t(m,k)\right)$ is increasing in m, with the result that (15) is weakly decreasing in $\psi_s(m,k)$ for all s. Since σ^{FSA} implies $\psi_s(m,k)$ as large as possible this proves the lemma.□

Proof of Corollary 1

Any stationary first-best equilibrium must satisfy $x_t\in\mathcal{M}_t$ in each period $t\geq 1$. Agents are symmetric, so in any stationary equilibrium there exists some agent i such that

$$\Pr\{i\in\mathcal{M}_t,x_t=i\}E[v_{\max,t}|i\in\mathcal{M}_t,x_t=i]\leq\frac{1}{N}E[v_{\max,t}],$$

$\forall t\geq 1$. Then (1) implies that first-best is attainable by a stationary equilibrium only if$\tilde{S}\leq\frac{1}{N}\left(E[v_{\max,t}]p_1-c\right)$, which implies that $\delta\geq\delta^{Stat}$ is a necessary condition. It is sufficient because the allocation rule that chooses $x_t\in\mathcal{M}_t$ uniformly at random and efforts $e_t=1$, $\forall t$, satisfies (1) for any $\delta\geq\delta^{Stat}$.

Since (1) is continuous in δ, to prove $\delta^{FSA}<\delta^{Stat}$ it suffices to show that (1) holds with strict inequality for σ^{FSA} at $\delta=\delta^{Stat}$. By Assumptions 1 and 3, $S_{(j)}^{FSA}$ is strictly decreasing in j. But then $S_{(1)}^{FSA}>\frac{1}{N}E[v_{\max,t}p_1-c]$ because $\sum_{j=1}^{N}S_{(j)}^{FSA}=E[v_{\max,t}p_1-c]=V^{FB}$. Thus $\delta^{FSA}<\delta^{Stat}$ by definition of δ^{Stat}.■

For Online Publication: Supplemental Results

Alternative Transfer Schemes

Description of Alternative Transfer Schemes

The equilibrium we construct has the following features: in each period t, the principal chooses x_t as in FSA. The chosen agent picks $e_{x_t,t} = 1$. The following transfers are paid:

1. $w_{x_t,t}$ equals

$$w_{x_t,t} = (v_{x_t,t} p_1 - E[\tau_{x_t,t} | e_t = 1]) - \delta S_{(N)}^{FSA},$$

   where $\tau_{x_t,t}$ is defined below. Note that $w_{x_t,t}$ could be either positive or negative (paid by either the principal or agent x_t).

2. Each agent $i \neq x_t$ pays

$$w_{i,t} = -\delta S_{(N)}^{FSA}.$$

3. For agents $i \neq x_t$, $\tau_{i,t} = 0$. Agent x_t believes her dyad-surplus is $E[S_{x_t,t+1} | I_{x_t}(h_0^{t+1})]$ following low output. Then

$$\tau_{x_t,t} = \begin{cases} \frac{\delta}{1-\delta} \delta S_{(N)}^{FSA} & \text{if } y_t = v_{x_t,t} \\ -\frac{\delta}{1-\delta} (E[S_{x_t,t+1} | I_{x_t}(h_0^{t+1})] - \delta S_{(N)}^{FSA}) & \text{if } y_t = 0 \end{cases}.$$

Following any deviation that is observed by the principal and agent i, agent i thereafter chooses $d_t = e_t = 0$ and $w_{i,t} = \tau_{i,t} = 0$. The principal continues to allocate production as in the FSA. If agent i is allocated production after he has observed a deviation, then he is treated as if he produced $y_t > 0$ with probability p_1 and is otherwise treated as if $y_t = 0$.

Statement of Result

Suppose that there exists a PBE that attains first-best and that the conditions of Proposition 1, Part 2 hold. Then the strategies described above are a PBE that attains first-best.

Proof

By Proposition 1, it suffices to show that the strategies described above are a PBE whenever the FSA with transfers as described in the paper is a PBE. At each on- and off-path history, the principal earns a fixed continuation payoff from each agent (equal to 0 if the agent has observed a deviation and $\delta S^{FSA}_{(N)}$ otherwise), regardless of her allocation decision at that history. The allocation decision also determines each agent's beliefs about their current rankings. However, these beliefs have exactly offsetting effects on the fees and bonuses for each agent and so have no effect on the principal's payoff. So the principal is indifferent between all allocation decisions and hence is willing to follow the FSA allocation rule.

An agent who is not allocated production earns no less than $S^{FSA}_{(N)} - \delta S^{FSA}_{(N)}$ in continuation surplus in each period because $S^{FSA}_{(k)} \geq S^{FSA}_{(N)}$ for all $k \in \{1, ..., N\}$. An agent earns 0 following a deviation. So agent $i \neq x_t$ is willing to pay $w_{i,t}$ if

$$-(1-\delta)w_{i,t} \leq \delta(1-\delta)S^{FSA}_{(N)}$$

which holds by construction. If $w_{x_t,t} < 0$, then a sufficient condition for agent x_t to be willing to pay $w_{x_t,t}$ is

$$-(1-\delta)(v_{x_t,t}p_1 - c - \delta S^N) \leq \delta(1-\delta)S^{FSA}_{(N)},$$

which holds because $v_{x_t,t}p_1 - c \geq 0$. The principal earns $(N-1)\delta S^{FSA}_{(N)}$ following a deviation observed by agent i. So the principal is willing to pay $w_{x_t,t} > 0$ if

$$(1-\delta)\left(v_{x_t,t}p_1 - w_{x_t,t} - E[\tau_{x_t,t}|e_t = 1]\right) + \delta^2 S^{FSA}_{(N)} \geq 0.$$

Plugging in $w_{x_t,t}$ yields $\delta S^{FSA}_{(N)} \geq 0$, which always holds. So there are no profitable deviations in fees.

The agent earns 0 continuation surplus following low output and $\delta S^{FSA}_{(1)}$

surplus following high output, so is willing to work hard if

$$\frac{c}{p_1 - p_0} \leq \frac{\delta}{1-\delta} S_{(1)}^{FSA}.$$

This is exactly the condition in Proposition 1.

The principal is willing to pay $\tau_{x_t,t} > 0$ because she earns $\delta S_{(N)}^{FSA}$ from agent x_t from period $t+1$ on, and

$$(1-\delta)\tau_{x_t,t} = (1-\delta)\frac{\delta}{1-\delta}\delta S_{(N)}^{FSA} \leq \delta^2 S_{(N)}^{FSA}$$

by construction. Agent x_t believes his continuation surplus is $E[S_{x_t,t+1}|I_{x_t}(h_0^{t+1})] - \delta S_{(N)}^{FSA}$ if $y_t = 0$, so he is willing to pay $\tau_{x_t,t} < 0$ because

$$-(1-\delta)\tau_{x_t,t} = \delta(E[S_{x_t+1}|I_{x_t}(h_0^{t+1})] - \delta S_{(N)}^{FSA}) \leq \delta(E[S_{x_t,t+1}|I_{x_t}(h_0^{t+1})] - \delta S_{(N)}^{FSA}).$$

Thus, this strategy is an equilibrium that attains first-best whenever FSA attains first-best.

Following a deviation that is observed only by agent i, the principal earns $(N-1)\delta S_{(N)}^{FSA}$ regardless of the allocation rule and so is willing to follow the equilibrium allocation. The principal's relationship with agents $j \neq i$ is unchanged, so there is no profitable deviation in these relationships. In the relationship with agent i, $w_{i,t} = \tau_{i,t} = 0$ and $d_t = e_t = 0$ whenever $x_t = i$ are myopic best responses. A similar argument applies if any subset of agents has observed a deviation. So there is no profitable deviation off the equilibrium path.■

The Equations for $S_{(k)}^{FSA}$

As in Section II.A, let $F_{(j)}^k$ be the probability that $v_{\max,t} = v^k$ and that the $j-1$ most recently productive agents are not in $\mathcal{M}_t$ at time t:

$$F_{(j)}^k = Pr\left\{\{1,...,j-1\} \cap \mathcal{M}_t = \emptyset, v_{\max,t} = v^k\right\}$$

and let

$$F_{(j)} = \sum_k F^k_{(j)} = Pr\left\{\{1, ..., j-1\} \cap \mathcal{M}_t = \emptyset\right\}.$$

FSA continuation surplus $S^{FSA}_{(j)}$ is defined recursively by

$$\begin{gathered} S^{FSA}_{(j)} = \\ \left(1 - F_{(j)}\right) \delta S^{FSA}_{(j)} + \textstyle\sum_{k=1}^{N} \left(F^k_{(j)} - F^k_{(j+1)}\right)(1-\delta)(v^k p_1 - c) + \\ \left(F_{(j)} - F_{(j+1)}\right) \delta \left(p_1 S^{FSA}_{(1)} + (1-p_1) S^{FSA}_{(j)}\right) + F_{(j+1)} \delta \left(p_1 S^{FSA}_{(j+1)} + (1-p_1) S^{FSA}_{(j)}\right) \end{gathered} \tag{16}$$

where we define $F_{(N+1)} = 0$ and have used the fact that

$$\begin{aligned} &Pr\left\{\{1, ..., j-1\} \cap \mathcal{M}_t = \emptyset, j \in \mathcal{M}_t, v_{\max,t} = v^k\right\} \\ &= \left(1 - F^k_{(j+1)}\right) - \left(1 - F^k_{(j)}\right) = F^k_{(j)} - F^k_{(j+1)}. \end{aligned}$$

Stacking these equations, we can write

$$\begin{pmatrix} S^{FSA}_{(1)} \\ S^{FSA}_{(2)} \\ S^{FSA}_{(3)} \\ \vdots \\ S^{FSA}_{(N)} \end{pmatrix} = \mathbf{A} \begin{pmatrix} v^1 p_1 - c \\ v^2 p_1 - c \\ v^3 p_1 - c \\ \vdots \\ v^K p_1 - c \end{pmatrix} + \mathbf{B} \begin{pmatrix} S^{FSA}_{(1)} \\ S^{FSA}_{(2)} \\ S^{FSA}_{(3)} \\ \vdots \\ S^{FSA}_{(N)} \end{pmatrix}$$

or more compactly,

$$\mathbf{S^{FSA}} = \mathbf{A}\left(\mathbf{v} p_1 - \iota_{\mathbf{K}} c\right) + \mathbf{B}\mathbf{S^{FSA}}$$

for $N \times K$ and $N \times N$ matrices $\mathbf{A}$ and $\mathbf{B}$ which collect the coefficients from (16), and $\iota_{\mathbf{K}}$ a $K \times 1$ vector of ones. Solving for $\mathbf{S^{FSA}}$ finally yields

$$\mathbf{S^{FSA}} = \left(\mathbf{I_N} - \mathbf{B}\right)^{-1} \mathbf{A}\left(\mathbf{v} p_1 - \iota_{\mathbf{K}} c\right)$$

which allows us to easily calculate FSA surplus for all parameter values.