Supplemental Appendix: Dynamic Screening and the Dual Roles of Monitoring

Dhruva Bhaskar*

January 28, 2025

Proof of Observation 1

Proof. To prove the first part, note that δ_2 is defined as a set of δ that solve a quadratic equation in δ, so has at most two solutions. A solution exists by the intermediate value theorem since w_L is continuous in δ: $w_L \to -\infty$ as $\delta \to 0$, and $w_L(1) = \mu_0 + (1-\mu_0)(1-u_L) > \overline{w}$. Moreover, there is only solution in $(0,1)$, since for $\delta \geq 0$

$$w_L'(\delta) > 0 > -(1-\overline{w}),$$

where $-(1-\overline{w})$ is the derivative of $1-\delta+\delta\overline{w}$ with respect to δ.

For i)-iii), note first that $w_1(1-u_L) = w_L(1-u_L)$ for any u_L. Therefore if $u_L = 1-\delta_1$, we have $\delta_1 = 1-u_L$, and $1-\delta_1+\delta_1\overline{w} = w_1(\delta_1) = w_L(\delta_1)$, so $\delta_2 = \delta_1$.

If $u_L < 1-\delta_1$, the facts that $w_1(\delta_1) = 1-\delta_1+\delta_1\overline{w}$ and $w_1'(\delta) > 0$ imply that $w_L(1-u_L) = w_1(1-u_L) > u_L + (1-u_L)\overline{w}$. As $w_L'(\delta) > 0$, this implies $\delta_2 < 1-u_L$. An analogous argument shows that if $u_L > 1-\delta_1$, then $\delta_2 > 1-u_L$. This completes the proof of Observation 1. ■

The role of the lower bound on μ_0 when $u_L > 0$

I will also show here that the assumption that when $u_L > 0$ we have

$$\mu_0 > \frac{u_L\overline{w}}{u_L\overline{w}+1-\overline{w}}$$

*Baruch College, CUNY

is necessary for the construction I use. In particular, if the inequality does not hold, we have $\delta_1 > 1 - u_L$ and δ_2 is not well-defined (only $\delta_2 \geq 1$ will satisfy the condition), which means neither payoff w_1 nor w_L is feasible. To see this, substitute $\mu_0 = \frac{u_L\overline{w}}{u_L\overline{w}+1-\overline{w}}$ into $\delta_1 = \frac{1+c-\mu_0}{1+c-\mu_0(c+\overline{w}}$. The resulting expression is strictly bigger than $1 - u_L$:

$$\frac{(1-\overline{w})(1+c)+cu_L\overline{w}}{(1-\overline{w})(1+c+u_L\overline{w})} > 1 - u_L,$$

which holds if and only if

$$c\overline{w} + (1-\overline{w})(1+c-\overline{w}(1-u_L)) > 0,$$

which is true for $u_L > 0$. Therefore, $\delta_1 > 1 - u_L$ for $\mu_0 \leq \frac{u_L\overline{w}}{u_L\overline{w}+1-\overline{w}}$ since δ_1 is decreasing in μ_0. Therefore, if $\mu_0 \leq \frac{u_L\overline{w}}{u_L\overline{w}+1-\overline{w}}$ the construction requires $\delta > \delta_2$. However, $w_L(1) = \overline{w}$ when $\mu_0 = \frac{u_L\overline{w}}{u_L\overline{w}+1-\overline{w}}$, which means that for $\delta < 1$, $w_L(\delta) < 1 - \delta + \delta\overline{w}$. Therefore, for any $\mu_0 < \frac{u_L\overline{w}}{u_L\overline{w}+1-\overline{w}}$, we have $w_L(1) < \overline{w}$ since $w_L(1)$ increasing in μ_0. This implies $w_L(\delta) < 1 - \delta + \delta\overline{w}$ for all $\delta \in (0,1)$, so δ_2 is not well-defined, and the construction fails.

Proof of Observation 2

Proof. For i): By definition, this holds for $k = 1$. For $k > 1$, we simply substitute the expressions for m, v^E and $\hat{H}_{k-1}$ in the definition of $\hat{H}_k$, and collect terms. For ii) note that by i), $B_2 - B_1 = 2 + \frac{c(1-\delta)}{\delta} - (\overline{w} + \overline{w} + (1-u_H)) > 0$, which follows from the fact that $\overline{w} + \overline{v} \leq 1 + u_H$. Now proceed by induction, and assume that $B_k > B_{k-1}$. Then

$$B_{k+1} - B_k = A_{k-1} - A_k + (1-u_H)(B_k - B_{k-1}) = \frac{c(1-\delta)}{\delta} + (1-u_H)(B_k - B_{k-1}) > 0.$$

Moreover, the difference is bounded below by a constant, so $B_k \to \infty$. ■

Proof of Observation 4

Proof. For i): For $k < n+1$, we proceed by induction. Note that $v_0 = \overline{v} > u_H$ by Claim 2. Assume $v_k > u_H$, then $v_{k+1} = \frac{v_k}{1+v_k-u_H} > \frac{u_H}{1+u_H-u_H} = u_H$, where the inequality comes from the fact that $\frac{v}{1+v-u_H}$ is strictly increasing in v. This proves that $v_k > u_H$ for $k \leq n+1$, and therefore, that $v_{k+1} = \frac{v_k}{1+v_k-u_H} < v_k$.

The same argument shows that $v_{k+1} < v_k$ and $v_k \geq u_L$, using the definition of the sequence for $k > n+1$. Therefore, the sequence is bounded below by u_L, and converges to some limit $l \geq u_L$. Taking limits, we have

$$\lim_{k\to\infty} v_k = l = \frac{l}{1 + l - u_L},$$

so l solves $l(l - u_L) = 0$. Therefore, $l = 0$ or $l = u_L$, and clearly the sequence must converge to the larger root, which proves ii). iii) For $k \leq n+1$, this is obvious. Using the fact that v_n and v_{n+1} are larger that u_H, and setting $v_n = v_{n+1} = u_H$ in the definition of the sequence gives us a lower bound for every point v_k for $k > n+1$ that is independent of δ. For δ sufficiently high, $1 - \delta$ is below this lower bound.

■

Proof of Proposition 5: last statement

Proof. Here we prove the last statement in the proposition, that for any k, there exists $\hat{\delta}$ such that if $\delta > \hat{\delta}$, we have $v^* < v_{n+k}$. It is easily checked from the definition of the sequence $\{v_k\}$, that v_k is increasing in δ for all k, and from the definition of $\hat{F}$, v_n is increasing in δ—note here that n itself depends on δ.

Fix $k \geq 5$. By Observation 4, for δ sufficiently high $v_{n+k} > 1 - \delta$. We will show that for δ sufficiently high, the left derivative of F at v_{n+k} can be made arbitrarily small, so that the optimiser of $\hat{W}$ lies strictly below v_{n+k}. In particular, we wish to show that $B_{n+k} < \frac{(1-\mu_0)}{\mu_0}(\delta\overline{w} - c(1-\delta))$. Start with δ sufficiently high that $\frac{(1-\mu_0)}{\mu_0}(\delta\overline{w} - c(1-\delta)) > 0$, and let $\underline{C}$ be the value at the initial δ. Since this expression is increasing in δ, $\underline{C}$ is a lower bound as we increase δ, so it is sufficient to show that $B_{n+k} < \underline{C}$.

Let $\epsilon > 0$ be sufficiently small that

$$\epsilon \sum_{i=0}^{\frac{k-1}{2}} (1 - u_L)^i < \underline{C}.$$

v_n is not explicitly given, but is implicitly defined as the point at which F switches from increasing to decreasing, and we do not know the actual slope of F in this range. Therefore, we need to uniformly bound the incremental change of the slope of F, moving from one interval to the next, and this will provide an upper bound on B_{n+k}.

By Observation 2,

$$B_2 - B_1 = 1 - \overline{w} - \overline{v} - (1 - u_H) + c(1-\delta) + 1,$$

which converges to 0 as $\delta \to 1$.[1] Therefore, there exists $\delta(\epsilon) < 1$ such that $\delta > \delta(\epsilon)$ implies that $2 - \overline{w} - \overline{v} - (1 - u_H) < \frac{\epsilon}{2}$ and $\frac{c(1-\delta)}{\delta} < \frac{\epsilon}{2}$, so $B_2 - B_1 < \epsilon$. Moreover, if for $i \leq n$, $B_i - B_{i-1} < \frac{\epsilon}{2}$, then by Observation 2,

$$B_{i+1} - B_i = (B_i - B_{i-1})(1 - u_H) + \frac{c(1-\delta)}{\delta} < \epsilon.$$

This is a uniform bound on differences for $i \leq n+1$, so $B_{n+1} < B_{n+1} - B_n < \epsilon$, since $B_n < 0$. By Observation 3,

$$\begin{aligned} B_{n+3} - B_{n+1} &= B_{n+1}(1 - u_L) - B_n(1 - u_H) - \frac{c(1-\delta)}{\delta} \\ &< B_{n+1} - B_n - u_L B_{n+1} < \epsilon(1 - u_L), \end{aligned}$$

and

$$B_{n+k} - B_{n+k-2} = (B_{n+k-2} - B_{n+k-4})(1 - u_L).$$

Therefore $B_{n+5} - B_{n+3} < \epsilon(1 - u_L)^2$. Moreover, if $B_{n+k-2} - B_{n+k-4} < \epsilon(1 - u_L)^{\frac{k-3}{2}}$, then $B_{n+k} - B_{n+k-2} < \epsilon(1 - u_L)^{\frac{k-1}{2}}$. In this case,

$$\begin{aligned} B_{n+k} &< \epsilon(1 - u_L)^{\frac{k-1}{2}} + B_{n+k-2} \\ &< \epsilon(1 - u_L)^{\frac{k-1}{2}} + ... + \epsilon(1 - u_L) + B_{n+1} \\ &< \epsilon(1 - u_L)^{\frac{k-1}{2}} + ... + \epsilon(1 - u_L) + \epsilon < \underline{C}, \end{aligned}$$

which proves what we want. ■

Proof of Proposition 3

I prove the following claim, which implies Proposition 3.

Claim 1. *Let $\delta > \overline{\delta}$. Then*

(i) If $v^* \geq v_{n+1}$, then the optimal contract in LBM is the same as the optimal contract in the main problem.

[1] To see this, note that $\overline{v}$, which is the larger solution of the quadratic equation $y(\overline{v}) + u_H = \overline{w} + \overline{v}$, converges to $1 + u_H - \overline{w}$ as $\delta \to 1$.

(ii) If $v^* \leq v_{n+2}$, an optimal contract in LBM has probability one monitoring at time zero, with the high type exerting effort, and if effort is observed, delivers value v_n using the optimal policy from (AP), from $t = 1$. The principal's payoff is

$$\mu_0((1-\delta)(1-c) + \delta F(v_n)) + (1-\mu_0)(\delta\overline{w} - c(1-\delta)).$$

Proof. For i), note that if $v^* \geq v_{n+1}$, the low type's incentive constrain is slack at all histories in the optimal contract. Therefore, the optimal contracts in LBM and the main problem coincide.

For ii): note that it is immediate that if $\delta > \overline{\delta}$, properties i)-iii) of Proposition 4 hold for any optimal contract in LBM—the only difference between the main problem and LBM is that the low type always shirks and has no incentive constraints. Denote by v_H the high type's time zero value from this contract. The high type incentive constraint in phase one requires that the monitoring probability $m \geq \frac{(1-\delta)(1-u_H)}{\delta v_H^E}$, where v_H^E is the high type's value after E. Clearly it is never optimal to set continuation play below the Pareto frontier, so $v_H^E \geq v_n$, and the principal's payoff after E is $F(v_H^E)$ with a known high type, with the optimal policy specified in the auxiliary problem for such a value. Therefore, the principal's payoff in LBM from an optimal contract can be written as

$$w\big(m, v_H^E\big) := \frac{\mu_0\big((1-\delta)(1-mc) + \delta m F\big(v_H^E\big)\big) + (1-\mu_0)m(\delta\overline{w} - c(1-\delta))}{1-\delta+\delta m},$$

and an optimal contract in LBM chooses m and v_H^E, subject to the high type's incentives, to maximise this function.

Now suppose $v_H^E > v_n$. Then the high type's incentive constraint must bind—otherwise we could lower v_H^E for an improvement. By the construction of the solution to the auxiliary problem, the principal's payoff must then be

$$\hat{W}(v_H) = \mu_0 F(v_H) + (1-\mu_0)\frac{1-v_H}{\delta}(\delta\overline{w} - c(1-\delta)).$$

Since $v^* \leq v_{n+2}$, we have that $\hat{W}(v_H) < \hat{W}(v^*)$ for $v_H \geq v_{n+1}$. Moreover, $\hat{W}(v^*)$ is feasible in LBM, so it must be that $v_H < v_{n+1}$. However, this implies that the high type's incentive constraint is slack, a contradiction. Therefore it must be that $v_H^E = v_n$, and the principal's payoff from an optimal policy in LBM is $w(m, v_n)$, for some m.

Now re-write the principal's payoff from a policy that optimally delivers $v \in (v_{n+2}, v_{n+1})$ to the low type in the main problem, with the implied phase one monitoring probability $m(v)$ as an argument:[2]

$$\tilde{W}(v, m(v)) := \frac{\mu_0((1-\delta)(1-m(v)c) + \delta m(v) F(v_n)) + (1-\mu_0) m(v)(\delta \overline{w} - c(1-\delta))}{1-\delta+\delta m(v)}.$$

Since $v^* < v_{n+1}$, we have that $\hat{W}_-(v_{n+1}) < 0$, or equivalently, that the total derivative of the function above is negative in the interval (v_{n+2}, v_{n+1}):

$$\hat{W}'(v) = \frac{d\tilde{W}}{dv}(v, m(v)) = \frac{\partial \tilde{W}}{\partial v} + \frac{\partial \tilde{W}}{\partial m}\frac{dm}{dv} = \frac{\partial \tilde{W}}{\partial m}\frac{dm}{dv} < 0.$$

Since $m(v)$ is decreasing in v, this means that $\frac{\partial \tilde{W}}{\partial m} > 0$. Now note that this partial derivative is equal to the partial derivative of $w(m, v_n)$:

$$\frac{\partial \tilde{W}}{\partial m} = \frac{(1-\delta)(\mu_0 \delta F(v_n) + (1-\mu_0)\delta \overline{w} - c(1-\delta) - \delta \mu_0)}{(1-\delta+\delta m)^2} = w_m(m, v_n).$$

The numerator of this derivative is independent of m so is positive for all m, and it is optimal to set $m = 1$ in LBM. This contract screens the low type in the first period, and from the second period delivers v_n using the policy in the auxiliary problem to the high type, and the principal's payoff is

$$w(1, v_n) = \mu_0((1-\delta)(1-c) + \delta F(v_n)) + (1-\mu_0)(\delta \overline{w} - c(1-\delta)).$$

■

Low Type Auxiliary Problem 2

I define a second auxiliary problem here. The solution to this problem defines an upper bound on P's payoff after a relevant history in the proof of Proposition 4. Consider the problem in which P knows that the agent is the low type, is constrained to deliver a value v to the low type, and has full commitment power. Note that the high type plays no role in this problem. Define the program as

$$\begin{aligned} G(v) := \max_{\sigma} \ & W(\sigma) && \text{(AP2)} \\ \text{s.t. } & V_L(\sigma) = v && (PK) \\ & ICL \ \forall h && (IC) \end{aligned}$$

[2] This is the same function as $\hat{W}$ but without substituting for $m(v)$. Note that in the interval specified, $v^E = v_n$ is constant since neither ICH nor ICL binds.

There are no principal constraints because P has full commitment, and to make sure that the problem is well-defined, assume that payments are bounded, so $p(h) \in [0, \overline{p}]$, where we take the upper bound sufficiently large, and P can deliver any value $v \in [0, M]$ to the low type.

Lemma 1.

$$G(v) = \begin{cases} (1-v)\overline{w} & \textit{if } v \in [0,1], \\ 1-v & \textit{if } v \in (1, M]. \end{cases}$$

I omit the proof as it follows a similar structure to the proof of the dynamic program in the main auxiliary problem, (AP) but is much simpler and standard. As $y_L + u_L < \overline{w}$, it is not efficient to incentivise effort from the agent. Moreover, since $\overline{w} < 1$, paying the agent by letting him shirk is more efficient than direct transfers.[3] If $v \leq 1$, the optimal outcome can be implemented by firing the agent with probability $1 - v$ at time zero, or with probability v, employing him forever and letting him shirk, with no monitoring. If $v > 1$, the agent is employed forever and allowed to shirk, and a payment of $v - 1$ is made in every period.

[3] If this wasn't the case, it would instead be optimal to pay the agent and fire him.