Dynamic Managerial Compensation: A Variational Approach Supplementary Material

Daniel Garrett
Toulouse School of Economics

Alessandro Pavan
Northwestern University

April 16, 2015

This document contains proofs for Example 1, Propositions 6, and Proposition 7 omitted in the main document. It also contains a brief description of the numerical analysis at the end of Section 4 in the main document.

All numbered items in this document contain the prefix S. Any numbered reference without a prefix refers to an item in the main document. Please refer to the main document for notation and definitions.

- **Proof of Example 1.** Observe that the inverse hazard rate of the period-1 distribution is

$$\frac{1-F_1(\theta_1)}{f_1(\theta_1)} = \begin{cases} \frac{1-\frac{32}{5}(\theta_1-3\theta_1^2)}{\frac{32}{5}(1-6\theta_1)} & 0 \leq \theta_1 \leq \frac{1}{8} \\ \frac{1-\frac{32}{5}(\frac{3}{32}+3\theta_1^2-\frac{1}{2}\theta_1)}{\frac{32}{5}(6\theta_1-\frac{1}{2})} & \frac{1}{8} < \theta_1 \leq \frac{1}{4} \end{cases}$$

  Next, observe that, for the effort policy ξ^R to satisfy the integral monotonicity constraints in (5), it must be that the function $q(\theta_1)$ defined by $q(\theta_1) \equiv \theta_1 - (1+\gamma^2)\frac{1-F_1(\theta_1)}{f_1(\theta_1)}$ is everywhere non-decreasing. It is easy to verify that the above condition fails under the proposed distribution, for any $\gamma \in [0,1]$.[1]

Finally, observe that the inverse hazard rate of the period-1 distribution reaches a maximum at $\theta = 1/8$ where it is equal to 5/16. Clearly,

$$5/16 = \sup\left\{[1-F_1(\theta_1)]/f_1(\theta_1)]\right\} < (1+\gamma)/(1+\gamma^2) - \left(\bar{\theta}_1 - \underline{\theta}_1\right) = (1+\gamma)/(1+\gamma^2) - 1/4,$$

for any $\gamma \in [0,1]$. Hence this distribution satisfies the condition in Part d(ii) of Proposition 5. That ε is drawn from a Uniform distribution in turn implies that the conditional distribution $F_2(\cdot|\cdot)$ satisfies the monotone-likelihood-ratio property, which is the other condition in Part d(ii) of Proposition 5. That expected effort necessarily increases over time when $\gamma < 1$ then follows from Proposition 5. Q.E.D.

[1] For example $q(0) = -\frac{5}{32}(1+\gamma^2) > q(\frac{1}{8}) = \frac{1}{8} - (1+\gamma^2)\frac{10}{32}$, for any γ.

Proof of Proposition 6. Let $\xi^*_\rho \equiv \left(\xi^*_{\rho,1}, \xi^*_{\rho,2}\right)$ be the effort policies sustained under any optimal contract, when the manager's preferences over consumption are represented by the function v_ρ, with the function family $(v_\rho)_{\rho \geq 0}$ satisfying the properties described in the main text. Recall, from Proposition 2, that such policies are essentially unique. Next, let K^*_ρ be the expected payoff of the lowest period-1 type (i.e., $\mathbb{E}^{\tilde{\theta}|\underline{\theta}_1}[V(\tilde{\theta})]$) under any optimal contract, when the manager's risk-aversion index is ρ. Finally, let $c^*_\rho \equiv \left(c^*_{\rho,1}, c^*_{\rho,2}\right)$ be a compensation policy sustained under an optimal contract and recall that, again by virtue of Proposition 2, such a policy is also essentially unique when $\rho > 0$, i.e., when the manager is strictly risk averse. When $\rho = 0$, instead, the distribution of consumption over the two periods is indeterminate, in which case let $c^*_{0,1}(\theta_1) = 0$ for all θ_1 and then let $c^*_{0,2}$ be given by (7).

Proposition 5 implies that the expected power of incentives is strictly higher in period 2 than in period 1, when $\rho = 0$. Our goal is to show the existence of $\bar{\rho} > 0$ such that, for any $\rho \in [0, \bar{\rho}]$, the expected power of incentives under any optimal contract continues to be higher in period two than in period one.

Suppose, for a contradiction, that no such $\bar{\rho}$ exists. Letting w_ρ denote the inverse of the function v_ρ, we then have that the following is instead true.

Claim S-A. *For any $n \in \mathbb{N}$, there exists $\rho_n \in \left(0, \frac{1}{n}\right)$ such that*

$$\mathbb{E}\left[\psi'\left(\xi^*_{\rho_n,1}(\tilde{\theta}_1)\right) w'_{\rho_n}\left(v_{\rho_n}(c^*_{\rho_n,1}(\tilde{\theta}_1))\right)\right] \geq \mathbb{E}\left[\psi'\left(\xi^*_{\rho_n,2}(\tilde{\theta})\right) w'_{\rho_n}\left(v_{\rho_n}(c^*_{\rho_n,2}(\tilde{\theta}))\right)\right]. \tag{S1}$$

On the other hand, we have that, when $\rho = 0$ (that is, when the manager is risk neutral), $\mathbb{E}\left[\psi'\left(\xi^*_{0,1}(\tilde{\theta}_1)\right)\right] < \mathbb{E}\left[\psi'\left(\xi^*_{0,2}(\tilde{\theta})\right)\right]$, as established in Proposition 5. Given these observations, below we establish a series of three properties that together imply that Claim S-A above is false.

Property S-(i). *For any ρ,*

$$\int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left\{\left|\xi^*_{\rho,1}(\theta_1)\right| + \left|\xi^*_{\rho,2}(\theta_1,\theta_2)\right|\right\} d\theta_2 d\theta_1 \leq 2\left(\frac{b}{a}\right)^2. \tag{S2}$$

Proof of Property S-(i). Let

$$L_\rho \equiv \int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left\{\xi^*_{\rho,1}(\theta_1)^2 + \xi^*_{\rho,2}(\theta_1,\theta_2)^2\right\} d\theta_2 d\theta_1. \tag{S3}$$

For arbitrary ρ, consider the gain in expected profits from using an optimal policy rather than simply paying the manager a constant wage equal to his outside option (assumed equal to zero), thus eliciting no effort. Given that w_ρ lies nowhere below the identity function, and given the bounds on the densities over Θ, it is easy to see that this gain must be no greater than

$$b^2\int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left\{\xi^*_{\rho,1}(\theta_1) + \xi^*_{\rho,2}(\theta_1,\theta_2)\right\} d\theta_2 d\theta_1 - \frac{1}{2}a^2 L_\rho.$$

Then note that

$$\int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left\{\xi^*_{\rho,1}(\theta_1) + \xi^*_{\rho,2}(\theta_1,\theta_2)\right\} d\theta_2 d\theta_1 \leq \int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left\{\left|\xi^*_{\rho,1}(\theta_1)\right| + \left|\xi^*_{\rho,2}(\theta_1,\theta_2)\right|\right\} d\theta_2 d\theta_1 \leq L_\rho^{1/2}, \tag{S4}$$

where the second inequality follows from Hölder's inequality. Hence, the expected gain from using the optimal policy is no greater than $b^2 L_\rho^{1/2} - \frac{1}{2}a^2 L_\rho$, which is non-negative only if $L_\rho \leq 4\left(\frac{b}{a}\right)^4$. The result then follows from (S4). ■

Property S-(ii). *Let $(\rho_n)_{n=1}^{\infty}$ be any sequence of real numbers satisfying the property in Claim S-A (that is, for all $n \in \mathbb{N}$, $\rho_n \in \left(0, \frac{1}{n}\right)$ is such that Condition (S1) holds). There exists $\epsilon > 0$ and $N \in \mathbb{N}$ such that, for all $n \geq N$,*

$$\int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left\{\left|\xi^*_{\rho_n,1}(\theta_1) - \xi^*_{0,1}(\theta_1)\right| + \left|\xi^*_{\rho_n,2}(\theta_1,\theta_2) - \xi^*_{0,2}(\theta_1,\theta_2)\right|\right\} d\theta_2 d\theta_1 > \epsilon, \tag{S5}$$

and

$$\left(\int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left\{\left(\xi^*_{\rho_n,1}(\theta_1) - \xi^*_{0,1}(\theta_1)\right)^2 + \left(\xi^*_{\rho_n,2}(\theta_1,\theta_2) - \xi^*_{0,2}(\theta_1,\theta_2)\right)^2\right\} d\theta_2 d\theta_1\right)^{1/2} > \epsilon. \tag{S6}$$

Proof of Property S-(ii). For each ρ, let

$$l_\rho \equiv \inf\left\{l \in \mathbb{R}_+ \cup \{+\infty\} : \left|w'_\rho(u) - 1\right| \leq l, \text{ all } u \in \mathbb{R}\right\}$$

and note that the assumptions on the function family $(v_\rho)_{\rho \geq 0}$ imply that $l_\rho \to 0$ as $\rho \to 0$. We start by proving the existence of ϵ and N such that, for all $n \geq N$, (S5) holds. Suppose for a contradiction that there are no such ϵ and N. Then there exists a subsequence (ρ_{n_k}) such that

$$\int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left|\xi^*_{\rho_{n_k},1}(\theta_1) - \xi^*_{0,1}(\theta_1)\right| + \left|\xi^*_{\rho_{n_k},2}(\theta_1,\theta_2) - \xi^*_{0,2}(\theta_1,\theta_2)\right| d\theta_2 d\theta_1 \to 0 \tag{S7}$$

as $n_k \to \infty$. However, then note that

$$\begin{aligned}
&\mathbb{E}\left[\begin{array}{c} \psi'\left(\xi^*_{\rho_{n_k},1}(\tilde\theta_1)\right) w'_{\rho_{n_k}}\left(v_{\rho_{n_k}}(c^*_{\rho_{n_k},1}(\tilde\theta_1))\right) - \psi'\left(\xi^*_{\rho_{n_k},2}(\tilde\theta)\right) w'_{\rho_{n_k}}\left(v_{\rho_{n_k}}(c^*_{\rho_{n_k},2}(\tilde\theta))\right) \\ -\left(\psi'\left(\xi^*_{0,1}(\tilde\theta_1)\right) - \psi'\left(\xi^*_{0,2}(\tilde\theta)\right)\right) \end{array}\right] \\
&= \mathbb{E}\left[\begin{array}{c} \left(\xi^*_{\rho_{n_k},1}(\tilde\theta_1) - \xi^*_{0,1}(\tilde\theta_1)\right) - \left(\xi^*_{\rho_{n_k},2}(\tilde\theta) - \xi^*_{0,2}(\tilde\theta)\right) \\ + \left(w'_{\rho_{n_k}}\left(v_{\rho_{n_k}}(c^*_{\rho_{n_k},1}(\tilde\theta_1))\right) - 1\right)\xi^*_{\rho_{n_k},1}(\tilde\theta_1) \\ - \left(w'_{\rho_{n_k}}\left(v_{\rho_{n_k}}(c^*_{\rho_{n_k},2}(\tilde\theta))\right) - 1\right)\xi^*_{\rho_{n_k},2}(\tilde\theta) \end{array}\right] \\
&\leq \mathbb{E}\left[\begin{array}{c} \left|\xi^*_{\rho_{n_k},1}(\tilde\theta_1) - \xi^*_{0,1}(\tilde\theta_1)\right| + \left|\xi^*_{\rho_{n_k},2}(\tilde\theta) - \xi^*_{0,2}(\tilde\theta)\right| \\ + \left|w'_{\rho_{n_k}}\left(v_{\rho_{n_k}}(c^*_{\rho_{n_k},1}(\tilde\theta_1))\right) - 1\right|\left|\xi^*_{\rho_{n_k},1}(\tilde\theta_1)\right| \\ + \left|w'_{\rho_{n_k}}\left(v_{\rho_{n_k}}(c^*_{\rho_{n_k},2}(\tilde\theta))\right) - 1\right|\left|\xi^*_{\rho_{n_k},2}(\tilde\theta)\right| \end{array}\right] \\
&\leq b^2 \int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left\{\left|\xi^*_{\rho_{n_k},1}(\theta_1) - \xi^*_{0,1}(\theta_1)\right| + \left|\xi^*_{\rho_{n_k},2}(\theta_1,\theta_2) - \xi^*_{0,2}(\theta_1,\theta_2)\right|\right\} d\theta_2 d\theta_1 \\
&+ b^2 l_{\rho_{n_k}} \int_{\Theta_1}\int_{\Theta_2(\theta_1)} \left\{\left|\xi^*_{\rho_{n_k},1}(\theta_1)\right| + \left|\xi^*_{\rho_{n_k},2}(\theta_1,\theta_2)\right|\right\} d\theta_2 d\theta_1.
\end{aligned}$$

The final expression converges to 0 as $n_k \to +\infty$, the first integral by (S7), and the second by Property S-1 above along with the fact that $l_\rho \to 0$ as $\rho \to 0$. It follows that, for any $\nu > 0$, there exists N such that, for any $n_k > N$,

$$\begin{aligned}&\mathbb{E}\left[\psi'\left(\xi^*_{\rho_{n_k},1}(\tilde{\theta}_1)\right) w'_{\rho_{n_k}}\left(v_{\rho_{n_k}}(c^*_{\rho_{n_k},1}(\tilde{\theta}_1))\right) - \psi'\left(\xi^*_{\rho_{n_k},2}(\tilde{\theta})\right) w'_{\rho_{n_k}}\left(v_{\rho_{n_k}}(c^*_{\rho_{n_k},2}(\tilde{\theta}))\right)\right] \\ &\leq \mathbb{E}\left[\psi'\left(\xi^*_{0,1}(\tilde{\theta}_1)\right) - \psi'\left(\xi^*_{0,2}(\tilde{\theta})\right)\right] + \nu.\end{aligned}$$

The right-hand side is negative whenever ν is taken sufficiently small, since, as noted above,

$$\mathbb{E}\left[\psi'\left(\xi^*_{0,1}(\tilde{\theta}_1)\right)\right] < \mathbb{E}\left[\psi'\left(\xi^*_{0,2}(\tilde{\theta})\right)\right].$$

However, this contradicts the assumption that the original sequence (ρ_n) satisfies Condition (S1).

Finally, that (S6) is also true follows simply from (S5) using Hölder's inequality. ∎

Now, for any $\rho \geq 0$, effort policy ξ, first-period consumption policy c_1, and constant $K \geq 0$, let

$$h_\rho(\xi, c_1, K) \equiv \mathbb{E}\left[\tilde{\theta}_1 + \xi_1(\tilde{\theta}_1) + \tilde{\theta}_2 + \xi_2(\tilde{\theta}) - c_1(\tilde{\theta}_1) - w_\rho\left(W(\tilde{\theta}; \xi) + K - v_\rho(c_1(\tilde{\theta}_1))\right)\right],$$

which is simply the principal's payoffs, as in (8). The final property we need to prove is the following.

Property S-(iii). *Take any sequence $(\rho_n)_{n=1}^{\infty}$ such that, for all $n \in \mathbb{N}$, $\rho_n \in \left(0, \frac{1}{n}\right)$. Then $h_0\left(\xi^*_{\rho_n}, c^*_{\rho_n,1}, K^*_{\rho_n}\right) \to h_0\left(\xi^*_0, c^*_{0,1}, K^*_0\right)$ as $n \to +\infty$.*

Proof of Property S-(iii). Suppose not. Then there must exist a subsequence (ρ_{n_k}) and $\eta > 0$ such that

$$h_0\left(\xi^*_{\rho_{n_k}}, c^*_{\rho_{n_k},1}, K^*_{\rho_{n_k}}\right) < h_0\left(\xi^*_0, c^*_{0,1}, K^*_0\right) - \eta$$

for all n_k. Then note that, for any (ξ, c_1, K), any $\rho \geq 0$, $h_\rho(\xi, c_1, K) \leq h_0(\xi, c_1, K)$, since w_ρ lies everywhere above w_0. Hence, for all n_k,

$$h_{\rho_{n_k}}\left(\xi^*_{\rho_{n_k}}, c^*_{\rho_{n_k},1}, K^*_{\rho_{n_k}}\right) < h_0\left(\xi^*_0, c^*_{0,1}, K^*_0\right) - \eta.$$

However, below we show that $h_{\rho_{n_k}}\left(\xi^*_0, c^*_{0,1}, K^*_0\right) \to h_0\left(\xi^*_0, c^*_{0,1}, K^*_0\right)$ as $n_k \to \infty$, which contradicts the optimality of $\left(\xi^*_{\rho_{n_k}}, c^*_{\rho_{n_k},1}, K^*_{\rho_{n_k}}\right)$ for some n_k sufficiently large.

To this end, for each ρ, we find a lower bound on $h_\rho\left(\xi^*_0, c^*_{0,1}, K^*_0\right) - h_0\left(\xi^*_0, c^*_{0,1}, K^*_0\right)$. Recall that, as a consequence of the result in Lemma 2 in the proof of Proposition 1, the function $\psi'\left(\xi^*_{0,1}(\theta_1)\right) + \gamma\mathbb{E}^{\tilde{\theta}_2|\theta_1}\left[\psi'(\xi^*_{0,2}(\theta_1, \tilde{\theta}_2))\right]$ is bounded uniformly over θ_1, and hence integrable; similarly, for each θ_1,

$\xi^*_{0,2}(\theta_1,\cdot)$ is uniformly bounded over $\Theta_2(\theta_1)$, and hence integrable. We then have that, for any ρ,

$$
\begin{aligned}
& h_\rho\left(\xi^*_0, c^*_{0,1}, K^*_0\right) - h_0\left(\xi^*_0, c^*_{0,1}, K^*_0\right) \\
& \geq -l_\rho \mathbb{E}\left[\left|\begin{array}{c} \psi(\xi^*_{0,1}(\tilde\theta_1)) + \psi\left(\xi^*_{0,2}(\tilde\theta)\right) \\ +K^*_0 + \int_{\underline\theta_1}^{\tilde\theta_1}\left\{\psi'\left(\xi^*_{0,1}(s)\right) + \gamma \mathbb{E}^{\tilde\theta_2|s}\left[\psi'(\xi^*_{0,2}(s,\tilde\theta_2))\right]\right\} ds \\ +\int_{\underline\theta_2}^{\tilde\theta_2}\psi'(\xi^*_{0,2}(\tilde\theta_1,s))ds - \mathbb{E}^{\tilde\theta_2|\tilde\theta_1}\left[\int_{\underline\theta_2}^{\tilde\theta_2}\psi'(\xi^*_{0,2}(\tilde\theta_1,s))ds\right] \end{array}\right|\right] \\
& \geq -l_\rho \mathbb{E}\left[\begin{array}{c} \psi(\xi^*_{0,1}(\tilde\theta_1)) + \psi\left(\xi^*_{0,2}(\tilde\theta)\right) \\ +K^*_0 + \int_{\underline\theta_1}^{\tilde\theta_1}\left\{\psi'\left(\xi^*_{0,1}(s)\right) + \gamma \mathbb{E}^{\tilde\theta_2|s}\left[\psi'(\xi^*_{0,2}(s,\tilde\theta_2))\right]\right\} ds \\ +\int_{\underline\theta_2(\tilde\theta_1)}^{\tilde\theta_2}\left|\psi'(\xi^*_{0,2}(\tilde\theta_1,s))\right|ds + \mathbb{E}^{\tilde\theta_2|\tilde\theta_1}\left[\int_{\underline\theta_2(\tilde\theta_1)}^{\tilde\theta_2}\left|\psi'(\xi^*_{0,2}(\tilde\theta_1,s))\right|ds\right] \end{array}\right] \\
& = -l_\rho \mathbb{E}\left[\begin{array}{c} \psi(\xi^*_{0,1}(\tilde\theta_1)) + \psi\left(\xi^*_{0,2}(\tilde\theta)\right) \\ +K^*_0 + \frac{1-F_1(\tilde\theta_1)}{f_1(\tilde\theta_1)}\left[\psi'\left(\xi^*_{0,1}\left(\tilde\theta_1\right)\right) + \gamma\psi'(\xi^*_{0,2}(\tilde\theta))\right] \\ +2\frac{1-F_2(\tilde\theta_2|\tilde\theta_1)}{f_2(\tilde\theta_2|\tilde\theta_1)}\left|\psi'(\xi^*_{0,2}(\tilde\theta))\right| \end{array}\right].
\end{aligned}
$$

The first inequality follows because, for any $y \in \mathbb{R}$, any ρ, $|w_\rho(y) - y| \leq l_\rho |y|$. The equality follows from integration by parts after noting that

$$\int_{\underline\theta_1}^{\theta_1}\left\{\psi'\left(\xi^*_{0,1}(s)\right) + \gamma \mathbb{E}^{\tilde\theta_2|s}\left[\psi'(\xi^*_{0,2}(s,\tilde\theta_2))\right]\right\} ds$$

is absolutely continuous in θ_1, and that $\int_{\underline\theta_2(\theta_1)}^{\theta_2}\left|\psi'(\xi^*_{0,2}(\theta_1,s))\right| ds$ is absolutely continuous in θ_2 for each θ_1 (these observations follow in turn from the integrability of $\psi'\left(\xi^*_{0,1}(\theta_1)\right) + \gamma\mathbb{E}^{\tilde\theta_2|\theta_1}\left[\psi'(\xi^*_{0,2}(\theta_1,\tilde\theta_2))\right]$, and from the integrability of $\xi^*_{0,2}(\theta_1,\cdot)$ for each θ_1). Using the bounds on f_1 and f_2, we then have that

$$
\begin{aligned}
& \mathbb{E}\left[\begin{array}{c} \psi(\xi^*_{0,1}(\tilde\theta_1)) + \psi\left(\xi^*_{0,2}(\tilde\theta)\right) \\ +K^*_0 + \frac{1-F_1(\tilde\theta_1)}{f_1(\tilde\theta_1)}\left[\psi'\left(\xi^*_{0,1}\left(\tilde\theta_1\right)\right) + \gamma\psi'(\xi^*_{0,2}(\tilde\theta))\right] \\ +2\frac{1-F_2(\tilde\theta_2|\tilde\theta_1)}{f_2(\tilde\theta_2|\tilde\theta_1)}\left|\psi'(\xi^*_{0,2}(\tilde\theta))\right| \end{array}\right] \\
& \leq b^2 \int_{\Theta_1}\int_{\Theta_2(\theta_1)}\left[\begin{array}{c} \frac{1}{2}\left[\xi^*_{0,1}(\theta_1)^2 + \xi^*_{0,2}(\theta_1,\theta_2)^2\right] + K^*_0 \\ \frac{1}{a}\left(\left|\xi^*_{0,1}(\theta_1)\right| + \gamma\left|\xi^*_{0,2}(\theta_1,\theta_2)\right|\right) + \frac{2}{a}\left|\xi^*_{0,2}(\theta_1,\theta_2)\right| \end{array}\right] d\theta_2 d\theta_1
\end{aligned}
$$

which, using the boundedness of (S3) together with (S2), is finite. Thus the fact that $l_{\rho_{n_k}} \to 0$ as $n_k \to +\infty$ implies $h_{\rho_{n_k}}\left(\xi^*_0, c^*_{0,1}, K^*_0\right) - h_0\left(\xi^*_0, c^*_{0,1}, K^*_0\right) \to 0$, which is what we wanted to show. ■

Now let N and ϵ be the values defined in Property S-(ii). Note that, for any $n > N$, there exists an incentive-compatible mechanism implementing the effort policy $\frac{1}{2}\xi^*_{\rho_n} + \frac{1}{2}\xi^*_0$, under which the manager's period-1 compensation is given by an arbitrary function $c^{\#}_{\rho_n,1}(\cdot)$, say $c^{\#}_{\rho_n,1}(\theta_1) = 0$ for all θ_1, and under which the lowest period-1 type's expected payoff is $\frac{1}{2}K^*_{\rho_n} + \frac{1}{2}K^*_0$ (the existence of such a mechanism follows from Proposition 1; in particular, because ψ is quadratic, $\frac{1}{2}\xi^*_{\rho_n} + \frac{1}{2}\xi^*_0$ satisfies Condition B(i) of this result). For such a mechanism, we have that

$$h_0\left(\frac{1}{2}\xi^*_{\rho_n} + \frac{1}{2}\xi^*_0, c^{\#}_{\rho_n,1}, \frac{1}{2}K^*_{\rho_n} + \frac{1}{2}K^*_0\right) - \left[\frac{1}{2}h_0\left(\xi^*_{\rho_n}, c^*_{\rho_n,1}, K^*_{\rho_n}\right) + \frac{1}{2}h_0\left(\xi^*_0, c^*_{0,1}, K^*_0\right)\right] \tag{S8}$$

$$= -\frac{1}{2}\mathbb{E}\left[\begin{array}{c} \left(\frac{1}{2}\xi^*_{\rho_n,1}(\tilde{\theta}_1) + \frac{1}{2}\xi^*_{0,1}(\tilde{\theta}_1)\right)^2 - \frac{1}{2}\xi^*_{\rho_n,1}(\tilde{\theta}_1)^2 - \frac{1}{2}\xi^*_{0,1}(\tilde{\theta}_1)^2 \\ + \left(\frac{1}{2}\xi^*_{\rho_n,2}(\tilde{\theta}) + \frac{1}{2}\xi^*_{0,2}(\tilde{\theta})\right)^2 - \frac{1}{2}\xi^*_{\rho_n,2}(\tilde{\theta})^2 - \frac{1}{2}\xi^*_{0,2}(\tilde{\theta})^2 \end{array}\right]$$

$$= \frac{1}{8}\mathbb{E}\left[\left(\xi^*_{\rho_n,1}(\tilde{\theta}_1) - \xi^*_{0,1}(\tilde{\theta}_1)\right)^2 + \left(\xi^*_{\rho_n,2}(\tilde{\theta}) - \xi^*_{0,2}(\tilde{\theta})\right)^2\right] \geq \frac{a^2}{8}\epsilon^2.$$

where the inequality follows from (S6) along with the fact that $f_1(\theta_1) f_2(\theta_2|\theta_1) > a^2$ for all $\theta_1 \in \Theta_1$, $\theta_2 \in \Theta_2(\theta_1)$.

That the property in Claim S-A is false then follows from the combination of the result in (S8) along with Property S-(iii) above, which jointly imply that as $n \to +\infty$

$$h_0\left(\frac{1}{2}\xi^*_{\rho_n} + \frac{1}{2}\xi^*_0, c^{\#}_{\rho_n,1}, \frac{1}{2}K^*_{\rho_n} + \frac{1}{2}K^*_0\right) \geq \frac{1}{2}h_0\left(\xi^*_{\rho_n}, c^*_{\rho_n,1}, K^*_{\rho_n}\right) + \frac{1}{2}h_0\left(\xi^*_0, c^*_{0,1}, K^*_0\right) + \frac{a^2}{8}\epsilon^2$$
$$> h_0\left(\xi^*_0, c^*_{0,1}(\theta_1), K^*_0\right)$$

thus contradicting the optimality of $\left(\xi^*_0, c^*_{0,1}, K^*_0\right)$.

Q.E.D.

Proof of Proposition 7. Let $\xi^*_\gamma \equiv \left(\xi^*_{\gamma,1}, \xi^*_{\gamma,2}\right)$ be the (essentially unique) effort policy sustained under any optimal contract, when the persistence of the productivity process is γ. Let K^*_γ be the optimal choice of the expected payoff for the lowest period-1 type (i.e., $\mathbb{E}^{\tilde{\theta}|\underline{\theta}_1}\left[V(\tilde{\theta})\right]$) when the persistence of the process is γ. Finally, let $c^*_\gamma \equiv \left(c^*_{\gamma,1}, c^*_{\gamma,2}\right)$ be a compensation policy sustained under an optimal contract and recall that, by virtue of Proposition 2, such a policy is also essentially unique.

Part (a). Consider the case of $\gamma = 1$. From (13), note that, for almost all $\theta_1 \in \Theta_1$,

$$\mathbb{E}^{\tilde{\theta}|\theta_1}\left[\psi'\left(\xi^*_{1,2}(\tilde{\theta})\right) w'\left(v(c^*_{1,2}(\tilde{\theta}))\right)\right] = \psi'\left(\xi^*_{1,1}(\theta_1)\right) w'\left(v\left(c^*_{1,1}(\theta_1)\right)\right) \tag{S9}$$
$$- \mathbb{E}^{\tilde{\theta}|\theta_1}\left[\frac{\psi''\left(\xi^*_{1,2}(\tilde{\theta})\right)}{f_2(\tilde{\theta}_2|\tilde{\theta}_1)} \int_{\underline{\theta}_2}^{\bar{\theta}_2} \left\{w'\left(v(c^*_{1,2}(\tilde{\theta}_1, r))\right) - w'\left(v(c^*_{1,1}(\tilde{\theta}_1))\right)\right\} f_2(r|\tilde{\theta}_1) dr\right].$$

We now establish that, whenever v is strictly concave, then with probability one (that is, for all but a zero-measure set of θ),

$$m(\theta_2; \theta_1) \equiv \int_{\theta_2}^{\bar{\theta}_2} \left\{w'\left(v\left(c^*_{1,2}(\theta_1, r)\right)\right) - w'\left(v\left(c^*_{1,1}(\theta_1)\right)\right)\right\} f_2(r|\theta_1)\, dr \geq 0. \tag{S10}$$

To see this, note that, for all θ_1, almost all θ_2,

$$\frac{\partial m(\theta_2;\theta_1)}{\partial \theta_2} = -\left[w'\left(v\left(c^*_{1,2}\left(\theta_1,\theta_2\right)\right)\right) - w'\left(v\left(c^*_{1,1}\left(\theta_1\right)\right)\right)\right] f_2\left(\theta_2|\theta_1\right). \tag{S11}$$

Next, recall from (12) that, with probability one, $w'\left(v\left(c^*_{1,1}(\theta_1)\right)\right) = \mathbb{E}^{\tilde{\theta}_2|\theta_1}\left[w'\left(v(c^*_{1,2}(\theta_1,\tilde{\theta}_2))\right)\right]$. Moreover, $c^*_{1,2}\left(\theta_1,\cdot\right)$ must be non-decreasing (this follows from the fact that incentive compatibility requires that $\pi_2(\theta_1,\cdot)$ be non-decreasing, as established in Proposition 1). Therefore, there exists $\hat{\theta}_2\left(\theta_1\right) \in \left[\underline{\theta}_2\left(\theta_1\right), \bar{\theta}_2\left(\theta_1\right)\right]$ such that $c^*_{1,2}\left(\theta_1,\theta_2\right) \leq c^*_{1,1}\left(\theta_1\right)$ for $\theta_2 \leq \hat{\theta}_2\left(\theta_1\right)$ and $c^*_{1,2}\left(\theta_1,\theta_2\right) > c^*_{1,1}\left(\theta_1\right)$ for $\theta_2 > \hat{\theta}_2\left(\theta_1\right)$. Using the property that $w'\left(v\left(\cdot\right)\right)$ is increasing, together with (S11), we then have that the function $m(\cdot;\theta_1)$ must be quasi-concave on $\left[\underline{\theta}_2\left(\theta_1\right), \bar{\theta}_2\left(\theta_1\right)\right]$. Finally, note that $w'\left(v\left(c^*_{1,1}(\theta_1)\right)\right) = \mathbb{E}^{\tilde{\theta}_2|\theta_1}\left[w'\left(v(c^*_{1,2}(\theta_1,\tilde{\theta}_2))\right)\right]$, implies that $m(\underline{\theta}_2\left(\theta_1\right);\theta_1) = 0$. That $m(\underline{\theta}_2\left(\theta_1\right);\theta_1) = m(\bar{\theta}_2\left(\theta_1\right);\theta_1) = 0$, along with the property that $m(\cdot;\theta_1)$ is quasi-concave, establish the claim in (S10). Similarly, it is easy to see that the inequality in (S10) is strict, unless $c^*_{1,2}\left(\theta_1,\cdot\right)$ is constant over $\left[\underline{\theta}_2\left(\theta_1\right), \bar{\theta}_2\left(\theta_1\right)\right]$.

Combining (S9) with (S10) permits us to conclude that, when $\gamma = 1$, the expected power of incentives is weakly lower in period 2 than in period 1 (strictly lower, unless, with probability one, $c^*_{1,2}\left(\theta_1,\cdot\right)$ is constant over $\left[\underline{\theta}_2\left(\theta_1\right), \bar{\theta}_2\left(\theta_1\right)\right]$).

Next consider Part (b). Suppose the result is not true. Then the following must be true.

Claim S-B. *There exists a sequence $(\gamma_n)_{n=1}^{\infty}$, with $\gamma_n \geq \gamma'$ all $n \in \mathbb{N}$, converging to 1 from below and such that, for all n,*

$$\mathbb{E}\left[\psi'\left(\xi^*_{\gamma_n,1}(\tilde{\theta}_1)\right) w'\left(v(c^*_{\gamma_n,1}(\tilde{\theta}_1))\right)\right] \leq \mathbb{E}\left[\psi'\left(\xi^*_{\gamma_n,2}(\tilde{\theta}_1,\gamma\tilde{\theta}_1+\tilde{\varepsilon})\right) w'\left(v(c^*_{\gamma_n,2}(\tilde{\theta}_1,\gamma\tilde{\theta}_1+\tilde{\varepsilon}))\right)\right]. \tag{S12}$$

Below we show that Claim S-B is inconsistent with the fact that, by assumption, when $\gamma = 1$,

$$\mathbb{E}\left[\psi'\left(\xi^*_{1,1}(\tilde{\theta}_1)\right) w'\left(v(c^*_{1,1}(\tilde{\theta}_1))\right)\right] > \mathbb{E}\left[\psi'\left(\xi^*_{1,2}(\tilde{\theta}_1,\gamma\tilde{\theta}_1+\tilde{\varepsilon})\right) w'\left(v(c^*_{1,2}(\tilde{\theta}_1,\gamma\tilde{\theta}_1+\tilde{\varepsilon}))\right)\right]. \tag{S13}$$

We establish the inconsistency by means of three properties that jointly lead to a contradiction of the claim.

First note that, by assumption, each $\xi^*_{\gamma_n}$ and ξ^*_1 are uniformly bounded, with the bound M uniform over n. This last property, along with (3) and (12), in turn imply that there exists $\bar{C} > 0$ such that $\left|c^*_{\gamma_n,1}\right|, \left|c^*_{1,1}\right|, \left|c^*_{\gamma_n,2}\right|, \left|c^*_{1,2}\right| \leq \bar{C}$ almost everywhere, and uniformly over n. Furthermore, from the optimality of the policies, one can easily see that there must exist $\bar{K} > 0$ such that $K^*_{\gamma_n} \leq \bar{K}$ for all n.

The following must then be true.

Property S-(iv). *Assume Claim S-B is true. Then there exist $\epsilon > 0$ and $N \in \mathbb{N}$ such that,* *for all $n \geq N$, at least one of the following holds:*

$$\Pr\left(\left|\xi^*_{\gamma_n,1}\left(\tilde{\theta}_1\right) - \xi^*_{1,1}\left(\tilde{\theta}_1\right)\right| > \epsilon\right) \geq \epsilon,$$
$$\Pr\left(\left|\xi^*_{\gamma_n,2}\left(\tilde{\theta}_1, \gamma_n\tilde{\theta}_1 + \tilde{\varepsilon}\right) - \xi^*_{1,2}\left(\tilde{\theta}_1, \tilde{\theta}_1 + \tilde{\varepsilon}\right)\right| > \epsilon\right) \geq \epsilon,$$
$$\Pr\left(\left|c^*_{\gamma_n,1}\left(\tilde{\theta}_1\right) - c^*_{1,1}\left(\tilde{\theta}_1\right)\right| > \epsilon\right) \geq \epsilon, \text{ or}$$
$$K^*_{\gamma_n} - K^*_1 \geq \epsilon.$$

Proof of Property S-(iv). Suppose Property S-(iv) is false. Then there exists a subsequence (γ_{n_k}) such that $\left|\xi^*_{\gamma_{n_k},1}(\theta_1) - \xi^*_{1,1}(\theta_1)\right|$, $\left|\xi^*_{\gamma_{n_k},2}(\theta_1, \gamma_n\theta_1 + \varepsilon) - \xi^*_{1,2}(\theta_1, \theta_1 + \varepsilon)\right|$, $\left|c^*_{\gamma_{n_k},1}(\theta_1) - c^*_{1,1}(\theta_1)\right|$, and $\left|c^*_{\gamma_{n_k},2}(\theta_1, \gamma_n\theta_1 + \varepsilon) - c^*_{1,2}(\theta_1, \theta_1 + \varepsilon)\right|$ all converge in probability to zero, which, given the boundedness of the policies, implies that (S12) and (S13) are mutually inconsistent. ■

Now, abusing the notation introduced in the proof of Proposition 6, we let

$$h_\gamma(\xi, c_1, K) \equiv \mathbb{E}\left[\tilde{\theta}_1 + \xi_1(\tilde{\theta}_1) + \tilde{\theta}_2 + \xi_2(\tilde{\theta}) - c_1(\tilde{\theta}_1) - w\left(W(\tilde{\theta};\xi) + K - v(c_1(\tilde{\theta}_1))\right)\right],$$

denote the firm's expected profits under the policies (ξ, c_1), when the lowest period-1 type's expected payoff is K. Note that the dependence on γ is both directly through the fact that $\tilde{\theta}_2 = \gamma\tilde{\theta}_1 + \tilde{\varepsilon}$ as well as through the function $W(\cdot;\cdot)$ that, along with (ξ, c_1, K), determines the period-2 compensation policy $c_2(\cdot)$ according to (7). Note that h_γ is strictly concave in ξ, $v(c_1)$ and K (this follows straightforwardly from the convexity of w and ψ). Strict concavity of $h_1(\xi, c_1, K)$, in particular, implies the following property (the result is obvious and hence the proof omitted).

Property S-(v). *There exists a function $\kappa : \mathbb{R}_{++} \to \mathbb{R}_{++}$ satisfying the following property. Take any $\epsilon > 0$ and any pair (ξ', c'_1, K') and (ξ'', c''_1, K'') satisfying at least one of the following*

$$\Pr\left(\left|\xi'_1\left(\tilde{\theta}_1\right) - \xi''_1\left(\tilde{\theta}_1\right)\right| > \epsilon\right) \geq \epsilon,$$
$$\Pr\left(\left|\xi'_2\left(\tilde{\theta}_1, \tilde{\theta}_1 + \tilde{\varepsilon}\right) - \xi''_2\left(\tilde{\theta}_1, \tilde{\theta}_1 + \tilde{\varepsilon}\right)\right| > \epsilon\right) \geq \epsilon,$$
$$\Pr\left(\left|c'_1\left(\tilde{\theta}_1\right) - c''_1\left(\tilde{\theta}_1\right)\right| > \epsilon\right) \geq \epsilon, \text{ or}$$
$$K' - K'' \geq \epsilon.$$

Let (ξ''', c'''_1, K''') be defined by $\xi''' = \frac{1}{2}\xi' + \frac{1}{2}\xi''$, $c'''_1 = w\left(\frac{1}{2}v(c'_1) + \frac{1}{2}v(c''_1)\right)$, and $K''' = \frac{1}{2}K' + \frac{1}{2}K''$. Then $h_1(\xi''', c'''_1, K''') \geq \kappa(\epsilon)$.

Next, we use the boundedness of the optimal policies to establish the following property.

Property S-(vi). *Assume Claim S-B is true. For all $\theta \in \Theta$, all $n \in \mathbb{N}$, let $\xi'_{\gamma_n,1}(\theta_1) = \xi^*_{\gamma_n,1}(\theta_1)$, $\xi'_{\gamma_n,2}(\theta) = \xi^*_{\gamma_n,2}(\theta_1, \gamma_n\theta_1 + \theta_2 - \theta_1)$, $c'_{\gamma_n,1}(\theta_1) = c^*_{\gamma_n,1}(\theta_1)$, and $K'_{\gamma_n} = K^*_{\gamma_n}$. Then $h_1\left(\xi'_{\gamma_n}, c'_{\gamma_n,1}, K'_{\gamma_n}\right) \to h_1\left(\xi^*_1, c^*_{1,1}, K^*_1\right)$ as $n \to +\infty$.*

Proof of Property S-(vi). Our approach to the proof is as follows. We construct, for each n, a policy $\left(\xi^{\#}_{\gamma_n}, c^{\#}_{\gamma_n,1}, K^{\#}_{\gamma_n}\right)$ which (together with $c^{\#}_{\gamma_n,2}$ defined by (7)) is implementable when $\gamma = \gamma_n$. We choose $\left(\xi^{\#}_{\gamma_n}, c^{\#}_{\gamma_n,1}, K^{\#}_{\gamma_n}\right)$ in particular so that

$$h_{\gamma_n}\left(\xi^{\#}_{\gamma_n}, c^{\#}_{\gamma_n,1}, K^{\#}_{\gamma_n}\right) \to h_1\left(\xi^*_1, c^*_{1,1}, K^*_1\right) \tag{S14}$$

as $n \to \infty$. Similarly, for each n, we construct a policy $\left(\xi^{\#\#}_{\gamma_n}, c^{\#\#}_{\gamma_n,1}, K^{\#\#}_{\gamma_n}\right)$ which (together with $c^{\#\#}_{\gamma_n,2}$ defined by (7)) is implementable for $\gamma = 1$. Moreover, $\left(\xi^{\#\#}_{\gamma_n}, c^{\#\#}_{\gamma_n,1}, K^{\#\#}_{\gamma_n}\right)$ is chosen so that

$$h_1\left(\xi^{\#\#}_{\gamma_n}, c^{\#\#}_{\gamma_n,1}, K^{\#\#}_{\gamma_n}\right) \to h_{\gamma_n}\left(\xi^*_{\gamma_n}, c^*_{\gamma_n,1}, K^*_{\gamma_n}\right) \tag{S-15}$$

as $n \to \infty$. These observations, together with the fact that $h_{\gamma_n}\left(\xi^{\#}_{\gamma_n}, c^{\#}_{\gamma_n,1}, K^{\#}_{\gamma_n}\right) \le h_{\gamma_n}\left(\xi^*_{\gamma_n}, c^*_{\gamma_n,1}, K^*_{\gamma_n}\right)$ and $h_1\left(\xi^{\#\#}_{\gamma_n}, c^{\#\#}_{\gamma_n,1}, K^{\#\#}_{\gamma_n}\right) \le h_1\left(\xi^*_1, c^*_{1,1}, K^*_1\right)$ for each n, then imply that

$$h_{\gamma_n}\left(\xi^*_{\gamma_n}, c^*_{\gamma_n,1}, K^*_{\gamma_n}\right) \to h_1\left(\xi^*_1, c^*_{1,1}, K^*_1\right) \tag{S16}$$

as $n \to +\infty$. The result in Property S-(vi) then follows from (S16) by considering the functional

$$\hat{h}_{\hat{\gamma}}\left(\xi, c_1, K; \gamma\right) = \mathbb{E}\left[\begin{array}{c} \tilde{\theta}_1 + \xi_1(\tilde{\theta}_1) + \hat{\gamma}\tilde{\theta}_1 + \tilde{\varepsilon} + \xi_2(\tilde{\theta}_1, \gamma\tilde{\theta}_1 + \tilde{\varepsilon}) - c_1(\tilde{\theta}_1) \\ -w\left(\begin{array}{c} \psi(\xi_1(\tilde{\theta}_1)) + \psi\left(\xi_2(\tilde{\theta}_1, \gamma\tilde{\theta}_1 + \tilde{\varepsilon})\right) + \int_{\underline{\theta}_1}^{\tilde{\theta}_1}\left\{\psi'\left(\xi_1(s)\right) + \hat{\gamma}\mathbb{E}^{\tilde{\varepsilon}}\left[\psi'(\xi_2(s, \gamma s + \tilde{\varepsilon}))\right]\right\} ds \\ + \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} \psi'(\xi_2(\tilde{\theta}_1, \gamma\tilde{\theta}_1 + s)) ds - \mathbb{E}^{\tilde{\varepsilon}}\left[\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} \psi'(\xi_2(\tilde{\theta}_1, \gamma\tilde{\theta}_1 + s)) ds\right] \\ + K - v(c_1(\tilde{\theta}_1)) \end{array}\right) \end{array}\right].$$

In particular, it follows from observing that $\hat{h}_{\hat{\gamma}}\left(\xi, c_1, K; \gamma\right)$ is continuous in $\hat{\gamma}$ uniformly over $\gamma \in [\gamma', 1]$ and over policies (ξ, c_1, K) satisfying the aforementioned bounds, and that $h_{\gamma_n}\left(\xi^*_{\gamma_n}, c^*_{\gamma_n,1}, K^*_{\gamma_n}\right) = \hat{h}_{\gamma_n}\left(\xi^*_{\gamma_n}, c^*_{\gamma_n,1}, K^*_{\gamma_n}; \gamma_n\right)$ while $h_1\left(\xi'_{\gamma_n}, c'_{\gamma_n,1}, K'_{\gamma_n}\right) = \hat{h}_1\left(\xi^*_{\gamma_n}, c^*_{\gamma_n,1}, K^*_{\gamma_n}; \gamma_n\right)$.

Our construction of $\left(\xi^{\#}_{\gamma_n}, c^{\#}_{\gamma_n,1}, K^{\#}_{\gamma_n}\right)$ for each n is as follows. Let $K^{\#}_{\gamma_n} = K^*_1 + 2\left(1 - \gamma_n\right) M \left(\bar{\theta}_1 - \underline{\theta}_1\right)$. Then, let $\xi^{\#}_{\gamma_n,1}\left(\theta_1\right) = \gamma_n \xi^*_{1,1}\left(\theta_1\right)$ and $\xi^{\#}_{\gamma_n,2}\left(\theta_1, \theta_2\right) = \xi^*_{1,2}\left(\theta_1, \theta_1 + \left(\theta_2 - \gamma_n\theta_1\right)\right)$. Finally, let $c^{\#}_{\gamma_n,1} = c^*_{1,1}$. Note that, since $\left(\xi^*_1, c^*_{1,1}, K^*_1\right)$, together with (7), defines an implementable policy when $\gamma = 1$, $\left(\xi^{\#}_{\gamma_n}, c^{\#}_{\gamma_n,1}, K^{\#}_{\gamma_n}\right)$ also defines an implementable policy when $\gamma = \gamma_n$. This is verified with respect to the conditions in Proposition 1. The only condition that is not immediate to check is B(i), or (using that ψ is quadratic) that, for all $\theta_1, \hat{\theta}_1$,

$$\int_{\hat{\theta}_1}^{\theta_1}\left\{\xi^{\#}_{\gamma_n,1}(\hat{\theta}_1) + \hat{\theta}_1 + \gamma_n\mathbb{E}^{\tilde{\varepsilon}}\left[\xi^{\#}_{\gamma_n,2}(\hat{\theta}_1, \gamma_n s + \tilde{\varepsilon})\right]\right\} ds \le \int_{\hat{\theta}_1}^{\theta_1}\left\{\xi^{\#}_{\gamma_n,1}(s) + s + \gamma_n\mathbb{E}^{\tilde{\varepsilon}}\left[\xi^{\#}_{\gamma_n,2}(s, \gamma_n s + \tilde{\varepsilon})\right]\right\} ds,$$

which, substituting for $\xi^{\#}_{\gamma_n}$, we can rewrite as

$$\begin{aligned} &\gamma_n \int_{\hat{\theta}_1}^{\theta_1}\left\{\xi^*_{1,1}(\hat{\theta}_1) + \hat{\theta}_1 + \mathbb{E}^{\tilde{\varepsilon}}\left[\xi^*_{1,2}(\hat{\theta}_1, \gamma_n s + \tilde{\varepsilon} + \left(1 - \gamma_n\right)\hat{\theta}_1)\right]\right\} ds - \left(1 - \gamma_n\right)\int_{\hat{\theta}_1}^{\theta_1}\left(s - \hat{\theta}_1\right) ds \\ &\le \gamma_n \int_{\hat{\theta}_1}^{\theta_1}\left\{\xi^*_{1,1}(s) + s + \mathbb{E}^{\tilde{\varepsilon}}\left[\xi^*_{1,2}(s, s + \tilde{\varepsilon})\right]\right\} ds. \end{aligned} \tag{S17}$$

To see that (S17) must hold, note that, because ξ_1^* is implementable, $\xi_{1,2}^*(\theta_1,\theta_2)+\theta_2$ is non-decreasing in θ_2, and so the left-hand side is no greater than

$$\begin{aligned}&\gamma_n\int_{\hat\theta_1}^{\theta_1}\left\{\xi_{1,1}^*(\hat\theta_1)+\hat\theta_1+\mathbb{E}^{\tilde\varepsilon}\left[\xi_{1,2}^*(\hat\theta_1,s+\tilde\varepsilon)\right]+(1-\gamma_n)\left(s-\hat\theta_1\right)\right\}ds-(1-\gamma_n)\int_{\hat\theta_1}^{\theta_1}\left(s-\hat\theta_1\right)ds\\&=\gamma_n\int_{\hat\theta_1}^{\theta_1}\left\{\xi_{1,1}^*(\hat\theta_1)+\hat\theta_1+\mathbb{E}^{\tilde\varepsilon}\left[\xi_{1,2}^*(\hat\theta_1,s+\tilde\varepsilon)\right]\right\}ds-(1-\gamma_n)^2\int_{\hat\theta_1}^{\theta_1}\left(s-\hat\theta_1\right)ds\end{aligned}$$

That (S17) holds then follows because condition B(i) of Proposition 1 holds for ξ_1^*, i.e.

$$\int_{\hat\theta_1}^{\theta_1}\left\{\xi_{1,1}^*(\hat\theta_1)+\hat\theta_1+\mathbb{E}^{\tilde\varepsilon}\left[\xi_{1,2}^*(\hat\theta_1,s+\tilde\varepsilon)\right]\right\}\le\int_{\hat\theta_1}^{\theta_1}\left\{\xi_{1,1}^*(s)+s+\mathbb{E}^{\tilde\varepsilon}\left[\xi_{1,2}^*(s,s+\tilde\varepsilon)\right]\right\}ds,$$

since ξ_1^* is an implementable policy.

To see that $\left(\xi_{\gamma_n}^{\#},c_{\gamma_n,1}^{\#},K_{\gamma_n}^{\#}\right)$ satisfies (S14), let $\zeta_\gamma(\theta_1,\varepsilon)=\xi_2(\theta_1,\gamma\theta_1+\varepsilon)$ and note that

$$\begin{aligned}h_\gamma(\xi,c_1,K)&=\mathbb{E}^{(\tilde\theta_1,\tilde\varepsilon)}\left[\begin{array}{c}\tilde\theta_1+\xi_1(\tilde\theta_1)+\gamma\tilde\theta_1+\tilde\varepsilon+\zeta_\gamma\left(\tilde\theta_1,\tilde\varepsilon\right)-c_1(\tilde\theta_1)\\-w\left(\begin{array}{c}\psi(\xi_1(\tilde\theta_1))+\psi\left(\zeta_\gamma(\tilde\theta_1,\tilde\varepsilon)\right)\\+\int_{\underline\theta_1}^{\tilde\theta_1}\left\{\psi'\left(\xi_1(s)\right)+\gamma\mathbb{E}^{\tilde\varepsilon}\left[\psi'(\zeta_\gamma(s,\tilde\varepsilon))\right]\right\}ds\\+\int_{\underline\varepsilon}^{\tilde\varepsilon}\psi'(\zeta_\gamma(\tilde\theta_1,s))ds-\mathbb{E}^{\tilde\varepsilon}\left[\int_{\underline\varepsilon}^{\tilde\varepsilon}\psi'(\zeta_\gamma(\tilde\theta_1,s))ds\right]\\+K-v(c_1(\tilde\theta_1))\end{array}\right)\end{array}\right]\\&\equiv d_\gamma\left(\xi_1,\zeta_\gamma,c_1,K\right).\end{aligned}\tag{S18}$$

Let $\zeta_1^*(\theta_1,\varepsilon)=\xi_{1,2}^*(\theta_1,\theta_1+\varepsilon)$ and, for each n, let $\zeta_{\gamma_n}^{\#}(\theta_1,\varepsilon)=\xi_{\gamma_n,2}^{\#}(\theta_1,\gamma_n\theta_1+\varepsilon)$. Now let $\mathcal{E}_1(M)$ be the space of first-period effort policies ξ_1 bounded by M, and endow this space with the sup norm. Let $\mathcal{Z}(M)$ denote the space of functions $\zeta(\theta_1,\varepsilon)$ (essentially) bounded by M, and let $\mathcal{C}_1\left(\bar C\right)$ denote the space of functions $c_1(\theta_1)$ (essentially) bounded by $\bar C$. Then note that $d_\gamma(\xi_1,\zeta,c_1,K)$ is continuous in (γ,ξ_1,K) uniformly over $[\gamma',1]\times\mathcal{E}_1(M)\times\mathcal{Z}(M)\times\mathcal{C}_1\left(\bar C\right)\times\left[0,\bar K+2(1-\gamma')M\left(\bar\theta_1-\underline\theta_1\right)\right]$. Moreover, by construction, for all n, and for all $(\theta_1,\varepsilon)\in\Theta_1\times[\underline\varepsilon,\bar\varepsilon]$, $\zeta_{\gamma_n}^{\#}(\theta_1,\varepsilon)=\zeta_1^*(\theta_1,\varepsilon)$, and $c_{\gamma n},1^{\#}(\theta_1)=c_{1,1}^*(\theta_1)$. These observations, together with the fact that $\left(\gamma_n,\xi_{\gamma_n,1}^{\#},K_{\gamma_n}^{\#}\right)$ converges uniformly to $\left(1,\xi_{1,1}^*,K_1^*\right)$, then imply (S14).

Next, we construct $\left(\xi_{\gamma_n}^{\#\#},c_{\gamma_n,1}^{\#\#},K_{\gamma_n}^{\#\#}\right)$. Let $\xi_{\gamma_n,1}^{\#\#}(\theta_1)=\frac{1}{\gamma_n}\xi_{\gamma_n,1}^*(\theta_1)+\left(2Mb(1-\gamma_n)+\frac{1-\gamma_n}{\gamma_n}\right)\theta_1$ and $\xi_{\gamma_n,2}^{\#\#}(\theta_1,\theta_2)=\xi_{\gamma_n,2}^*(\theta_1,\gamma_n\theta_1+(\theta_2-\theta_1))$. Let

$$K_{\gamma_n}^{\#\#}=K_1^*+\left(\bar\theta_1-\underline\theta_1\right)\left(\frac{1}{\gamma_n}-\gamma_n\right)M.$$

Finally, let $c_{\gamma_n,1}^{\#\#}=c_{1,1}^*$. Note that, since $\left(\xi_{\gamma_n}^*,c_{\gamma_n,1}^*,K_{\gamma_n}^*\right)$, together with (**??**), defines an implementable policy when $\gamma=\gamma_n$, $\left(\xi_{\gamma_n}^{\#\#},c_{\gamma_n,1}^{\#\#},K_{\gamma_n}^{\#\#}\right)$ also defines an implementable policy when $\gamma=1$.

Again, this is verified by considering Proposition 1. The only condition which is not immediate to check is B(i), or (using that ψ is quadratic) that, for all $\theta_1, \hat{\theta}_1$,

$$\int_{\hat{\theta}_1}^{\theta_1} \left\{ \xi^{\#\#}_{\gamma_n,1}(\hat{\theta}_1) + \hat{\theta}_1 + \mathbb{E}^{\tilde{\varepsilon}} \left[\xi^{\#\#}_{\gamma_n,2}(\hat{\theta}_1, s+\tilde{\varepsilon}) \right] \right\} ds \leq \int_{\hat{\theta}_1}^{\theta_1} \left\{ \xi^{\#\#}_{\gamma_n,1}(s) + s + \mathbb{E}^{\tilde{\varepsilon}} \left[\xi^{\#\#}_{\gamma_n,2}(s, s+\tilde{\varepsilon}) \right] \right\} ds, \tag{S19}$$

which, substituting for $\xi^{\#\#}_{\gamma_n}$, we can rewrite as

$$\begin{aligned} &\int_{\hat{\theta}_1}^{\theta_1} \left\{ \frac{1}{\gamma_n} \left(\xi^*_{\gamma_n,1}(\hat{\theta}_1) + \hat{\theta}_1 \right) + 2\hat{\theta}_1 Mb\left(1-\gamma_n\right) + \mathbb{E}^{\tilde{\varepsilon}} \left[\xi^*_{\gamma_n,2} \left(\hat{\theta}_1, \gamma_n s + \tilde{\varepsilon} + (1-\gamma_n) \left(s - \hat{\theta}_1 \right) \right) \right] \right\} ds \\ &\leq \int_{\hat{\theta}_1}^{\theta_1} \left\{ \frac{1}{\gamma_n} \left(\xi^*_{\gamma_n,1}(s) + s \right) + 2sMb\left(1-\gamma_n\right) + \mathbb{E}^{\tilde{\varepsilon}} \left[\xi^*_{\gamma_n,2}(s, \gamma_n s + \tilde{\varepsilon}) \right] \right\} ds. \end{aligned}$$

Then note that, for any $\hat{\theta}_1$, s,

$$\begin{aligned} &\left| \mathbb{E}^{\tilde{\varepsilon}} \left[\xi^*_{\gamma_n,2} \left(\hat{\theta}_1, \gamma_n s + \tilde{\varepsilon} + (1-\gamma_n) \left(s - \hat{\theta}_1 \right) \right) \right] - \mathbb{E}^{\tilde{\varepsilon}} \left[\xi^*_{\gamma_n,2}(\hat{\theta}_1, \gamma_n s + \tilde{\varepsilon}) \right] \right| \\ &= \left| \int_{\min\left\{ \underline{\varepsilon} + (1-\gamma_n)\left(s-\hat{\theta}_1\right), \underline{\varepsilon} \right\}}^{\max\left\{ \bar{\varepsilon} + (1-\gamma_n)\left(s-\hat{\theta}_1\right), \bar{\varepsilon} \right\}} \xi^*_{\gamma_n,2}(\hat{\theta}_1, \gamma_n s + \varepsilon) d \left[G\left(\varepsilon - (1-\gamma_n) \left(s - \hat{\theta}_1 \right) \right) - G\left(\varepsilon\right) \right] \right| \\ &\leq 2Mb\left(1-\gamma_n\right) \left(s - \hat{\theta}_1 \right), \end{aligned} \tag{S20}$$

where the inequality follows because the density of $\tilde{\varepsilon}$ is bounded by b (which is equivalent to our requirement that the density $f_2\left(\theta_2|\theta_1\right)$ is bounded). The inequality (S20), together with

$$\begin{aligned} &\frac{1}{\gamma_n} \int_{\hat{\theta}_1}^{\theta_1} \left\{ \xi^*_{\gamma_n,1}(\hat{\theta}_1) + \hat{\theta}_1 + \gamma_n \mathbb{E}^{\tilde{\varepsilon}} \left[\xi^*_{\gamma_n,2}(\hat{\theta}_1, \gamma_n s + \tilde{\varepsilon}) \right] \right\} ds \\ &\leq \frac{1}{\gamma_n} \int_{\hat{\theta}_1}^{\theta_1} \left\{ \xi^*_{\gamma_n,1}(s) + s + \gamma_n \mathbb{E}^{\tilde{\varepsilon}} \left[\xi^*_{\gamma_n,2}(s, \gamma_n s + \tilde{\varepsilon}) \right] \right\} ds \end{aligned}$$

(which holds, since ξ^*_1 is an implementable effort policy), implies (S19).

Finally, (S-15) follows by arguments analogous to those for (S14). ■

We are now ready to establish that Claim S-B is false. Let $\xi^{1/2}_{\gamma_n,1}\left(\theta_1\right) = \frac{1}{2}\xi'_{\gamma_n,1}\left(\theta_1\right) + \frac{1}{2}\xi^*_{1,1}\left(\theta_1\right)$, $\xi^{1/2}_{\gamma_n,2}\left(\theta\right) = \frac{1}{2}\xi'_{\gamma_n,2}\left(\theta\right) + \frac{1}{2}\xi^*_{1,2}\left(\theta\right)$, $c^{1/2}_{1,\gamma_n}\left(\theta_1\right) = w\left(\frac{1}{2}v\left(c'_{\gamma_n,1}\left(\theta_1\right)\right) + \frac{1}{2}v\left(c^*_{1,1}\left(\theta_1\right)\right)\right)$ and $K^{1/2}_{\gamma_n} = \frac{1}{2}K'_{\gamma_n} + \frac{1}{2}K^*_1$, where $\left(\xi'_{\gamma_n}, c'_{\gamma_n,1}, K'_{\gamma_n}\right)$ is defined in Property S-(vi). Recall the construction of $\left(\xi^{\#\#}_{\gamma_n}, c^{\#\#}_{\gamma_n,1}, K^{\#\#}_{\gamma_n}\right)$ in the proof of Property S-(iii), and recall that this policy is implementable when $\gamma = 1$. Then let $\hat{\xi}^{1/2}_{\gamma_n} = \frac{1}{2}\xi^{\#\#}_{\gamma_n} + \frac{1}{2}\xi^*_1$, $\hat{c}^{1/2}_{1,\gamma_n} = w\left(\frac{1}{2}v\left(c^{\#\#}_{\gamma_n,1}\right) + \frac{1}{2}v\left(c^*_{1,1}\right)\right)$ and $\hat{K}^{1/2}_{\gamma_n} = \frac{1}{2}K^{\#\#}_{\gamma_n} + \frac{1}{2}K^*_1$. Using that $\left(\hat{\xi}^{1/2}_{\gamma_n}, \hat{c}^{1/2}_{1,\gamma_n}, \hat{K}^{1/2}_{\gamma_n}\right)$ and $\left(\xi^{1/2}_{\gamma_n}, c^{1/2}_{1,\gamma_n}, K^{1/2}_{\gamma_n}\right)$ are uniformly (essentially) bounded (and hence the continuity of $h_1\left(\cdot,\cdot,\cdot\right)$ over the bounded policies), we have

$$h_1\left(\hat{\xi}^{1/2}_{\gamma_n}, \hat{c}^{1/2}_{1,\gamma_n}, \hat{K}^{1/2}_{\gamma_n}\right) - h_1\left(\xi^{1/2}_{\gamma_n}, c^{1/2}_{1,\gamma_n}, K^{1/2}_{\gamma_n}\right) \to 0 \tag{S21}$$

as $n \to \infty$.

Now note that, if Claim S-B were true, by virtue of Properties S-(iv) and S-(v), we would have that

$$h_1\left(\xi_{\gamma_n}^{1/2}, c_{1,\gamma_n}^{1/2}, K_{\gamma_n}^{1/2}\right) \geq \frac{1}{2} h_1\left(\xi'_{\gamma_n}, c'_{\gamma_n,1}, K'_{\gamma_n}\right) + \frac{1}{2} h_1\left(\xi_1^*, c_{1,1}^*, K_1^*\right) + \kappa(\epsilon) \tag{S22}$$

for all $n \geq N$. By the inequality (S22), the fact that $\kappa(\epsilon) > 0$, Property S-6, and (S21), we conclude that, for all large enough n,

$$h_1\left(\hat{\xi}_{\gamma_n}^{1/2}, \hat{c}_{1,\gamma_n}^{1/2}, \hat{K}_{\gamma_n}^{1/2}\right) > h_1\left(\xi_1^*, c_{1,1}^*, K_1^*\right).$$

However, note that $\left(\hat{\xi}_{\gamma_n}^{1/2}, \hat{c}_{1,\gamma_n}^{1/2}, \hat{K}_{\gamma_n}^{1/2}\right)$ defines (together with (7)) an implementable policy for $\gamma = 1$ (this follows because both $\left(\xi_{\gamma_n}^{\#\#}, c_{\gamma_n,1}^{\#\#}, K_{\gamma_n}^{\#\#}\right)$ and $\left(\xi_1^*, c_{1,1}^*, K_1^*\right)$ are implementable for $\gamma = 1$, and by the conditions in Proposition 1; in particular, because ψ is quadratic, the convex combination of any two effort policies satisfying condition B(i) in Proposition 1 continues to satisfy this condition). This contradicts the optimality of $\left(\xi_1^*, c_{1,1}^*, K_1^*\right)$. That Claim S-B is false then implies the result in Part (b) in the proposition is true, which concludes the proof. Q.E.D.

Numerical Analysis of Section 4. As explained in the main text, in the numerical exercises at the end of Section 4, we assume that, for all $c \geq 0$, $v(c) = \left(c^{1-\eta} - 1\right)/(1-\eta)$, with $\eta \in [0, 1/2]$.[2] Hence, for any $v \geq -1/(1-\eta)$, $w(v) = (v(1-\eta)+1)^{1/(1-\eta)}$. We also assume that θ_1 is drawn from a uniform distribution with support $[0, 1/2]$, while ε is drawn from a uniform distribution with support $[-.5, .5]$. We then characterize the optimal contract in terms of three policies

$$\begin{aligned} \xi_1 &: [0, 1/2] \to \mathbb{R} \\ \hat{\xi}_2 &: [0, 1/2] \times [-.5, .5] \to \mathbb{R} \\ c_1 &: [0, 1/2] \to \mathbb{R} \end{aligned}$$

for different values of the persistence parameter, γ, and coefficient of relative risk aversion, η.

Note that, contrary to the main text, here we find it convenient to express period-2 effort as a function $\hat{\xi}_2(\theta_1, \varepsilon)$ of period-1 productivity, θ_1, and the period-2 shock, ε. Obviously, this alternative representation is inconsequential for the results (for each (θ_1, θ_2) with $\theta_2 \in \Theta_1(\theta_1)$, simply let $\xi_2(\theta_1, \theta_2) = \hat{\xi}_2(\theta_1, \theta_2 - \gamma\theta_1)$.).

[2]Note that, contrary to what assumed in the model setup in the main text, this felicity function is not surjective and Lipschitz continuous over the entire real line. However, the numerical results do not hinge on the lack of these properties. In fact, under the optimal policies identified in the numerical analysis, consumption is bounded away from zero from below. One can then construct extensions $\hat{v}$ of the assumed felicity function v such that (a) $\hat{v}(c) = v(c)$ for all $c > c_0 > 0$, (b) the numerical solutions under $\hat{v}$ coincide with those under v, and (c) $\hat{v}$ satisfies all the conditions in the model setup.

Then let $\hat{\xi} = \left\langle \xi_1, \hat{\xi}_2 \right\rangle$, and, for all $(\theta_1, \varepsilon) \in [0, 1/2] \times [-.5, .5]$ and $\gamma \in [0, 1]$, let

$$\begin{aligned}
\hat{W}\left((\theta_1, \varepsilon); \hat{\xi}\right) &\equiv \frac{\xi_1(\theta_1)^2}{2} + \frac{\hat{\xi}_2(\theta_1, \varepsilon)^2}{2} + \int_0^{\theta_1} \left\{\xi_1(s) + \gamma \mathbb{E}^{\tilde{\varepsilon}}\left[\hat{\xi}_2(s, \tilde{\varepsilon}))\right]\right\} ds \\
&+ \int_{-\frac{1}{2}}^{\varepsilon} \hat{\xi}_2(\theta_1, y) dy - \mathbb{E}^{\tilde{\varepsilon}}\left[\int_{-\frac{1}{2}}^{\tilde{\varepsilon}} \hat{\xi}_2(\theta_1, y) dy\right] \\
&= \frac{\xi_1(\theta_1)^2}{2} + \frac{\hat{\xi}_2(\theta_1, \varepsilon)^2}{2} + \int_0^{\theta_1} \left\{\xi_1(s) + \gamma \int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{\xi}_2(s, y) dy\right\} ds \\
&+ \int_{-\frac{1}{2}}^{\varepsilon} \hat{\xi}_2(\theta_1, y) dy - \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\varepsilon} \hat{\xi}_2(\theta_1, y) dy d\varepsilon.
\end{aligned}$$

Note that the function $\hat{W}$ is the analog of the function W in the main text, but with arguments (θ_1, ε) as opposed to (θ_1, θ_2). Again, the two functions $\hat{W}$ and W are related by the condition

$$\hat{W}\left((\theta_1, \varepsilon); \left\langle \xi_1, \hat{\xi}_2 \right\rangle\right) = W\left(\theta_1, \gamma\theta_1 + \varepsilon; \langle \xi_1, \xi_2 \rangle\right) \text{ all } (\theta_1, \varepsilon).$$

The objective of the numerical analysis is to maximize by means of the functions $\left(\xi_1, \hat{\xi}_2, c_1\right)$ the firm's expected profits, which, by virtue of Lemma 1 in the main text, can be expressed as follows:

$$\begin{aligned}
&\mathbb{E}\left[\tilde{\theta}_1 + \xi_1(\tilde{\theta}_1) + \gamma\tilde{\theta}_1 + \tilde{\varepsilon} + \hat{\xi}_2(\tilde{\theta}_1, \tilde{\varepsilon}) - c_1(\tilde{\theta}_1) - w\left(\hat{W}\left(\left(\tilde{\theta}_1, \tilde{\varepsilon}\right); \hat{\xi}\right) - v(c_1(\tilde{\theta}_1))\right)\right] \qquad (0.1) \\
&= \mathbb{E}\left[(1+\gamma)\tilde{\theta}_1 + \tilde{\varepsilon}\right] \\
&+ \int_0^{1/2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[\xi_1(\theta_1) + \xi_2(\theta_1, \varepsilon) - c_1(\theta_1) - w\left(\hat{W}\left((\theta_1, \varepsilon); \hat{\xi}\right) - v(c_1(\theta_1))\right)\right] 2 d\varepsilon d\theta_1.
\end{aligned}$$

The Euler conditions for this problem are the analogs of those in Proposition 8 in the main text, expressed in terms of (θ_1, ε) as opposed to (θ_1, θ_2):

$$\xi_1(\theta_1) w'(v(c_1(\theta_1))) = 1 - \int_{\theta_1}^{1/2} w'(v(c_1(r))) dr,$$

$$\begin{aligned}
\hat{\xi}_2(\theta_1, \varepsilon) w'\left(\hat{W}((\theta_1, \varepsilon); \hat{\xi}) - v(c_1(\theta_1))\right) &= 1 - \gamma \int_{\theta_1}^{1/2} w_1'(v(c_1(r))) dr \\
- \int_{\varepsilon}^{1/2} \left\{w'\left(\hat{W}((\theta_1, r); \hat{\xi}) - v(c_1(\theta_1))\right) - w'(v(c_1(\theta_1)))\right\} dr,
\end{aligned}$$

and

$$w'(v(c_1(\theta_1))) = \int_{-\frac{1}{2}}^{\frac{1}{2}} w'\left(\hat{W}\left((\theta_1, \varepsilon); \hat{\xi}\right) - v(c_1(\theta_1))\right) d\varepsilon.$$

We arrive at the policies $\left(\xi_1^R, \hat{\xi}_2^R, c_1^R\right)$ by maximizing (0.1) over the set of policies that solve the above Euler conditions. We then verify that the policies $\left(\xi_1^R, \hat{\xi}_2^R, c_1^R\right)$ can indeed be sustained in a mechanism that is IR and IC for the managers. For this purpose, we verify that the policies identified in the numerical analysis, along with the period-2 consumption policy defined by $c_2^R(\theta_1, \theta_2) = w(\hat{W}(\theta_1, \theta_2 - \gamma\theta_1); \hat{\xi}^R) - v(c_1(\theta_1)))$ satisfy conditions (a)-(e) below:

(a) $\xi_1^R(\theta_1) + \gamma\mathbb{E}\left[\hat{\xi}_2^R(\theta_1, \tilde{\varepsilon})\right]$ nondecreasing in θ_1;

(b) $\hat{\xi}_2^R(\theta_1, \varepsilon)$ nonincreasing in ε, all $\theta_1 \in [0, 1/2]$;

(c) $\xi_2^R(\theta_1, \varepsilon) + \varepsilon$ nondecreasing in ε, all $\theta_1 \in [0, 1/2]$;

(d) $\theta_1 + \xi_1^R(\theta_1) + \gamma\mathbb{E}\left[\gamma\theta_1 + \tilde{\varepsilon} + \hat{\xi}_2^R(\theta_1, \tilde{\varepsilon})\right]$ nondecreasing in θ_1;

(e) $c_1^R(\theta_1), c_2^R(\theta_1, \varepsilon), \xi_1^R(\theta_1), \hat{\xi}_2^R(\theta_1, \varepsilon) \geq 0$, all $(\theta_1, \varepsilon) \in [0, 1/2] \times [-1/2, +1/2]$.

One can easily verify that the above conditions imply that the policies $\left(\xi_1^R, \xi_2^R, c_1^R\right)$, with ξ_2^R defined by $\xi_2^R(\theta_1, \theta_2) = \hat{\xi}_2^R(\theta_1, \theta_2 - \gamma\theta_1)$ all (θ_1, θ_2) with $\theta_2 \in \Theta_2(\theta_1)$, along with the period-2 compensation $c_2^R(\theta_1, \theta_2) = w(\hat{W}(\theta_1, \theta_2 - \gamma\theta_1); \hat{\xi}^R) - v(c_1(\theta_1)))$ satisfy all the conditions in Proposition 1 in the main text (with $K = 0$), and hence are implementable.

Finally, observe that Figures 1, 2, and 4 in the main text depict the distortions

$$D_1(\theta_1) = 1 - \xi_1^R(\theta_1)\, w'\left(v\left(c_1^R(\theta_1)\right)\right)$$

$$\mathbb{E}^{\tilde{\theta}|\theta_1}[D_2(\tilde{\theta})] = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[1 - \hat{\xi}_2^R(\theta_1, \varepsilon)\, w'\left(\hat{W}((\theta_1, \varepsilon)\,; \hat{\xi}^R) - v(c_1^R(\theta_1))\right)\right] d\varepsilon$$

for different values of θ_1, whereas Figure 3 depicts the effort functions

$$\xi_1^R(\theta_1) \text{ and } \mathbb{E}[\hat{\xi}_2^R(\theta_1, \tilde{\varepsilon})] = \int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{\xi}_2^R(\theta_1, \varepsilon)\, d\varepsilon.$$

Finally, Figures 5 and 6 in the main text depict the unconditional expected difference between period-2 and pariod-1 distortions

$$\begin{aligned}
Diff &= \mathbb{E}\left[1 - \hat{\xi}_2^R\left(\tilde{\theta}_1, \tilde{\varepsilon}\right) w'\left(\hat{W}(\left(\tilde{\theta}_1, \tilde{\varepsilon}\right); \hat{\xi}^R) - v(c_1^R(\tilde{\theta}_1))\right)\right] - \mathbb{E}\left[1 - \xi_1^R\left(\tilde{\theta}_1\right)\ w'\left(v\left(c_1^R\left(\tilde{\theta}_1\right)\right)\right)\right] \\
&= \mathbb{E}\left[\xi_1^R\left(\tilde{\theta}_1\right) w'\left(v\left(c_1^R\left(\tilde{\theta}_1\right)\right)\right) - \hat{\xi}_2^R\left(\tilde{\theta}_1, \tilde{\varepsilon}\right) w'\left(\hat{W}(\left(\tilde{\theta}_1, \tilde{\varepsilon}\right); \hat{\xi}^R) - v(c_1^R(\tilde{\theta}_1))\right)\right] \\
&= 2\int_0^{1/2}\left\{\xi_1^R(\theta_1)\, w'\left(v\left(c_1^R(\theta_1)\right)\right) - \int_{-\frac{1}{2}}^{\frac{1}{2}} \hat{\xi}_2^R(\theta_1, \varepsilon)\, w'\left(\hat{W}((\theta_1, \varepsilon)\,; \hat{\xi}^R) - v(c_1^R(\theta_1))\right) d\varepsilon\right\} d\theta_1.
\end{aligned}$$