Progressive Learning*

Avidit Acharya† and Juan Ortner‡

April 10, 2017

Abstract

We study a dynamic principal-agent relationship with adverse selection and limited commitment. We show that when the relationship is subject to productivity shocks, the principal may be able to improve her value over time by progressively learning the agent's private information. She may even achieve her first best payoff in the long-run. The relationship may also exhibit path dependence, with early shocks determining the principal's long-run value. These findings contrast sharply with the results of the ratchet effect literature, in which the principal persistently obtains low payoffs, giving up substantial informational rents to the agent.

JEL Classification Codes: C73, D86

Key words: principal-agent, dynamic contracting, adverse selection, ratchet effect, inefficiency, learning, path dependence.

---

*For helpful comments, we would like to thank Steve Callander, Stephane Wolton, and seminar audiences at Berkeley, Boston University, Collegio Carlo Alberto, Northwestern Kellogg Strategy, Stanford, and the LSE/NYU political economy conference. The paper also benefited from the thoughtful feedback of a Co-Editor and four anonymous referees.

†Assistant Professor of Political Science, 616 Serra Street, Stanford University, Stanford CA 94305 (email: avidit@stanford.edu).

‡Assistant Professor of Economics, Boston University, 270 Bay State Road, Boston MA 02215 (email: jortner@bu.edu).

1 Introduction

Consider a long-term relationship between an agent who has persistent private information and a principal who cannot commit to long-term contracts. If the parties are sufficiently forward-looking, then the relationship is subject to the ratchet effect: the agent is unwilling to disclose his private information, fearing that the principal will update the terms of his contract. This limits the principal's ability to learn the agent's private information, and reduces her value from the relationship.

The ratchet effect literature has shed light on many economic applications including planning problems (Freixas et al., 1985), labor contracting (Gibbons, 1987; Dewatripont, 1989), regulation (Laffont and Tirole, 1988), optimal taxation (Dillen and Lundholm, 1996), repeated buyer-seller relationships (Hart and Tirole, 1988; Schmidt, 1993), and relational contracting (Halac, 2012; Malcomson, 2015).

A natural feature in virtually all of these applications is that productivity shocks to the economy have the potential to change the incentive environment over time. In this paper, we show that the classic ratchet effect results may not hold when the principal-agent relationship is subject to time-varying productivity shocks. In particular, the principal may gradually learn the agent's private information, which increases the value that she obtains from the relationship over time. The principal may even achieve her first-best payoff in the long run.

We study a stochastic game played between a principal and an agent. At each period, the principal offers the agent a transfer in exchange for taking an action that benefits her. The principal is able to observe the agent's action, but the agent's cost of taking the action is his private information, and constant over time. The principal has short-term, but not long-term, commitment power: she can credibly promise to pay a transfer in the current period if the agent takes the action, but cannot commit to future transfers. The realization of a productivity shock affects the size of the benefit that the principal obtains from having the agent take the action. The realization of the current period shock is publicly observed by both the principal and the agent at the start of the period, and the shock evolves over time as a Markov process.

Hart and Tirole (1988) and Schmidt (1993) study the special case of our model in which productivity is constant over time. The equilibrium of this special case differs qualitatively from the equilibrium of our model in which productivity changes over time. We explain the three main differences as follows.

First, we find that in the presence of productivity shocks the equilibrium may be persistently inefficient. This contrasts with the equilibrium of the model without the shocks, which is efficient.

Second, productivity shocks give the principal the opportunity to progressively learn the agent's private information. As a result, the principal's value from the relationship gradually improves over time. We show that under natural assumptions, the principal is only able to get the agent to disclose some of his private information when productivity is low; that is, learning takes place in "bad times." We also show that productivity shocks can enable the principal to obtain profits that are arbitrarily close to her full commitment profits. Lastly, we derive conditions under which the principal ends up fully learning the agent's private information and attains her first-best payoffs in the long-run.

Third, we show that learning by the principal may be path dependent: the degree to which the principal learns the agent's private information may depend critically on the order in which productivity shocks were realized early on in the relationship. This is true even when the process governing the evolution of productivity is ergodic. As a result, early shocks can have a lasting impact on the principal's value from the relationship.

Our model generates two testable predictions. First, the agent's performance in our model will typically be higher after the firm experiences negative shocks. This is consistent with Lazear et al. (2016), who find evidence that workers' productivity increases following a recession. Second, there will be hysteresis in the agent's compensation: the current wage of the agent is negatively affected by previous negative shocks. This result is consistent with Kahn (2010) and Oreopoulos et al. (2012), who find evidence that recessions have a long lasting impact on workers' compensation.

The key feature of our model that drives these dynamics is that the agent's incentive to conceal his private information changes over time. When current productivity is low and the future looks dim, the informational rents that low cost types expect to earn by mimicking a higher cost type are small. When these rents are small, it is cheap for the principal to get a low cost agent to reveal his private information. These changes in the cost of inducing information disclosure make it possible for the principal to progressively screen the different types of agents, giving rise to our equilibrium dynamics.

Related literature. Our work relates to prior papers that have suggested different ways of mitigating the ratchet effect. Kanemoto and MacLeod (1992) show that competition for second-hand workers may alleviate the ratchet effect. Carmichael and MacLeod (2000) show that the threat of future punishment may deter the principal from updating the terms of the agent's contract, mitigating the ratchet effect. Fiocco and Strausz (2015) show that the principal can incentivize information disclosure by delegating contracting to an independent third party. Our paper differs from these studies in that we do not introduce external sources of contract enforcement, nor do we reintroduce commitment by allowing for non-Markovian strategies.

Instead, we focus on the role that shocks play in ameliorating the principal's commitment problem. This connects our paper with Ortner (2016), who considers a durable goods monopolist who lacks commitment power and who faces time-varying production costs. In contrast to the classic results on the Coase conjecture (Fudenberg et al., 1985; Gul et al., 1986), Ortner (2016) shows that time-varying costs may enable the monopolist to extract rents from high value buyers. A key difference between Ortner (2016) and the current paper is that the interaction between the monopolist and buyers is one-shot in the Coasian environment. As a result, issues of information revelation, which are central to the current paper, are absent in that model.[1]

Blume (1998) generalizes the Hart and Tirole (1988) model to a setting in which the consumer's valuation has both permanent and transient components. Blume (1998) shows that optimal renegotiation-proof contracts in this environment give the buyer the chance to exit in the future in case his valuation falls. Gerardi and Maestri (2015) study a dynamic contracting model with no commitment in which the agent's private information affects his marginal cost of effort. They find that the principal's lack of commitment may push her to offer inefficient pooling contracts.

Our model is strategically equivalent to a setting in which the agent has at each period an outside option, whose value varies over time and is publicly observed. This relates our model to papers studying how outside options affect equilibrium dynamics in the classic Coasian model (Fuchs and Skrzypacz, 2010; Board and Pycia, 2014; Hwang and Li, 2017).[2] The key difference, again, is that we study the effect that time-varying outside options have in settings with repeated interaction.[3]

[1] The current paper also differs from Ortner (2016) in terms of results. Ortner (2016) shows that the monopolist's ability to extract rents diminishes as the support of the value distribution becomes dense. In contrast, the equilibrium dynamics of our model hold independently of how dense the support of the agent's cost distribution is.

[2] See also Compte and Jehiel (2002), who study the effect that outside options have in models of reputational bargaining.

[3] Our model also relates to Kennan (2001), who studies a bilateral bargaining game in which a long-run seller faces a long-run buyer. The buyer is privately informed about her valuation, which evolves over time as a Markov chain. Kennan (2001) shows that time-varying private information gives rise to cycles in which the seller's offer depends on the buyer's past purchasing decisions.

The path-dependence result relates our paper to a series of recent studies in organization economics that attempt to explain the persistent performance differences among seemingly identical firms (Gibbons and Henderson, 2012). Chassang (2010) shows that path-dependence may arise when a principal must learn how to effectively monitor the agent. Li and Matouschek (2013) study relational contracting environments in which the principal has private information, and show that this private information may give rise to cycles. Callander and Matouschek (2014) show that persistent performance differences may arise when managers engage in trial and error experimentation. Halac and Prat (2015) show that path-dependence arises due to the agent's changing beliefs about the principal's monitoring ability. We add to this literature by providing a new explanation for persistent performance differences, with new testable implications.

Finally, our paper relates to a broader literature on dynamic games with private information (Hart, 1985; Sorin, 1999; Wiseman, 2005; Peski, 2008, 2014). In this literature our paper relates closely to work by Watson (1999, 2002), who studies a private information partnership game, and shows that the value of the partnership increases over time as the players gradually increase the stakes of their relationship to screen out bad types.

2 Two Period Example

Consider the following two-period game played between a principal and an agent. At $t = 0$, the agent learns her cost of work $c \in \{c_L, c_H\}$. Let $\mu \in (0, 1)$ be the probability that the agent's cost is c_L. At the start of each period $t = 0, 1$, the principal's benefit $b_t \in \{b_L, b_H\}$ from having the agent work is publicly revealed. After observing b_t, the principal offers the agent a transfer $T_t \geq 0$ for working. The agent then publicly chooses whether or not to work. The payoffs of the principal and an agent of type c are

$$(1 - \delta)(b_0 - T_0)a_0 + \delta(b_1 - T_1)a_1,$$
$$(1 - \delta)(T_0 - c)a_0 + \delta(T_1 - c)a_1,$$

where $a_t \in \{0, 1\}$ denotes whether or not the agent works in period $t = 0, 1$ and $\delta \in (0, 1)$ measures the importance of period $t = 1$ relative to period $t = 0$. We assume

$$0 \leq c_L < b_L < c_H < b_H \qquad \text{and} \qquad \mu < \frac{b_H - c_H}{b_H - c_L} =: \overline{\mu}$$

Lastly, we assume that the benefit b_t is drawn i.i.d. over time, with $\text{prob}(b_t = b_L) = q \in [0, 1]$ for $t = 0, 1$. We consider pure strategy equilibria of this game.

Consider play at $t = 1$. Since we focus on pure strategy equilibria, at the start of $t = 1$ the principal' beliefs are either equal to her prior or are degenerate. If the principal's beliefs are equal to her prior, she finds it optimal to offer a transfer $T_1 = c_H$ that both types accept if $b_1 = b_H$ (since $\mu < \overline{\mu}$), and she finds it optimal to offer transfer $T_1 = c_L$ that only a low cost type accepts if $b_1 = b_L$. If the principal learned that the agent's cost is c, she finds it optimal to offer $T_1 = c$, which the agent accepts if and only if $b_1 > c$.

Consider now play at $t = 0$. Suppose first that $b_0 = b_L$. In this case, the principal must choose between two options: make a low offer that both types reject, or make a higher offer that only the low cost type accepts. Making an offer that both types accept is not profitable since $b_L < c_H$. Suppose the principal makes a separating offer T_0 that only a low cost type accepts. Note that a low cost agent reveals his private information by accepting, so his payoff is $(1 - \delta)(T_0 - c_L) + \delta 0$. Also note that the low cost type can obtain a payoff of $\delta(1 - q)(c_H - c_L)$ by rejecting the offer, so we must have $T_0 \geq c_L + \frac{\delta}{1-\delta}(1 - q)(c_H - c_L)$. Since the high cost type rejects offer T_0 if and only if $T_0 \leq c_H$, we must have $c_H \geq c_L + \frac{\delta}{1-\delta}(1 - q)(c_H - c_L)$, or

$$\frac{\delta}{1 - \delta}(1 - q) \leq 1. \tag{1}$$

When the future is sufficiently valuable (i.e., $\delta > 1/2$), this inequality holds only if the probability $1 - q$ of high productivity tomorrow is low enough; i.e., if the future looks dim. When (1) holds, the principal finds it optimal to make a separating offer, since such an offer gets the low cost type to work at time $t = 0$. In contrast, when (1) does not hold the principal makes a low offer that both types reject.

Suppose next that $b_0 = b_H$. In this case, the assumption that $\mu < \overline{\mu}$ implies that it is optimal for the principal to make a pooling offer $T_0 = c_H$ that both types accept. In particular, if the benefit is large with probability 1 (i.e., $q = 0$), the principal is never able to learn the agent's type.

There are three main takeaways from this example. First, productivity shocks may enable the principal to learn the agent's private information. Second, learning happens when times are bad and the future looks dim. Third, there is path dependence: the value that the principal derives in the second period depends on the first period shock.

In the rest of the paper, we consider an infinite horizon model in which both the agent's type and the principal's benefit can take finitely many values. The three main takeaways of the two period model extend to this environment. But the infinite horizon model gives rise to new results as well. First, the principal may learn the agent's private information gradually over time. Second, even when learning takes place, learning may stop before the principal achieves her first best payoff. And finally, the principal's payoff may display path dependence even in the long run, and even when the process governing the evolution of productivity is ergodic.

3 Model

3.1 Setup

We study a repeated interaction between a principal and an agent. Time is discrete and indexed by $t = 0, 1, 2, ..., \infty$. At the start of each period t, a state b_t is drawn from a finite set of states $\mathcal{B}$, and is publicly revealed. The evolution of b_t is governed by a Markov process with transition matrix $[Q_{b,b'}]_{b,b' \in \mathcal{B}}$. After observing $b_t \in \mathcal{B}$, the principal decides how much transfer $T_t \geq 0$ to offer the agent in exchange for taking a productive action. The agent then decides whether or not to take the action. We denote the agent's choice by $a_t \in \{0, 1\}$, where $a_t = 1$ means that the agent takes the action at period t. The action provides the principal a benefit equal to b_t.

The agent incurs a cost $ac \geq 0$ when choosing action $a \in \{0, 1\}$. The agent's cost c of taking the action is his private information, and it is fixed throughout the game. c may take one of K possible values from the set $\mathcal{C} = \{c_1, ..., c_K\}$. The principal's prior belief about the agent's cost is denoted $\mu_0 \in \Delta(\mathcal{C})$, which we assume has full support. At the end of each period the principal observes the agent's action and updates her beliefs about the agent's cost. The players receive their payoffs and the game moves to the next period.[4] Both players are risk-neutral expected utility maximizers and share a common discount factor $\delta < 1$.[5] The payoffs to the principal and an agent of cost $c = c_k$ at the end of period t are, respectively,

$$u(b_t, T_t, a_t) = (1 - \delta)\,(b_t - T_t)\,a_t,$$
$$v_k(b_t, T_t, a_t) = (1 - \delta)\,(T_t - c_k)\,a_t.$$

[4]As in Hart and Tirole (1988) and Schmidt (1993), the principal can commit to paying the transfer within the current period, but cannot commit to a schedule of transfers in future periods.

[5]The results are qualitatively the same when the players have different discount factors.

We assume, without loss of generality, that the agent's possible costs are ordered so that $0 < c_1 < c_2 < ... < c_K$. To avoid having to deal with knife-edge cases, we further assume that $b \neq c_k$ for all $b \in \mathcal{B}$ and $c_k \in \mathcal{C}$. Then, it is socially optimal for an agent with cost c_k to take action $a = 1$ at state $b \in \mathcal{B}$ if and only if $b - c_k > 0$. Let the set of states at which it is socially optimal for an agent with cost c_k to take the action be

$$E_k := \{b \in \mathcal{B} : b > c_k\}.$$

We refer to E_k as the *efficiency set* for type c_k. Note that by our assumptions on the ordering of types, the efficiency sets are nested, i.e. $E_{k'} \subseteq E_k$ for all $k' \geq k$.

We assume that process $\{b_t\}$ is persistent and that players are moderately patient. To formalize this, first define the following function: for any $b \in \mathcal{B}$ and $B \subseteq \mathcal{B}$, let

$$X(b, B) := (1-\delta)\mathbb{E}\left[\sum_{t=1}^{\infty} \delta^t \mathbf{1}_{\{b_t \in B\}} | b_0 = b\right],$$

where $\mathbb{E}[\cdot|b_0 = b]$ denotes the expectation operator with respect to the Markov process $\{b_t\}$, given that the period 0 state is b. Thus $X(b, B)$ is the expected discounted amount of time that the realized state is in B in the future, given that the current state is b. For any $b \in \mathcal{B}$, let $b^+ := \{b' \in \mathcal{B} : b' \geq b\}$. We maintain the following assumption throughout.

Assumption 1 (discounting/persistence) $X(b, b^+) > 1 - \delta$ *for all* $b \in \mathcal{B}$.

When there are no shocks to productivity (i.e., when the state is fully persistent) this assumption holds when $\delta > 1/2$. In general, for any $\delta > 1/2$, it holds whenever the process $\{b_t\}$ is sufficiently persistent. For any ergodic process $\{b_t\}$, there is a cutoff $\overline{\delta} \in (1/2, 1)$ such that the assumption holds whenever $\delta > \overline{\delta}$.

3.2 Histories, Strategies and Equilibrium Concept

A history $h_t = \langle (b_0, T_0, a_0), ..., (b_{t-1}, T_{t-1}, a_{t-1}) \rangle$ records the states, transfers and agent's action from the beginning of the game until the start of period t. For any two histories $h_{t'}$ and h_t with $t' \geq t$, we write $h_{t'} \succeq h_t$ if the first t period entries of $h_{t'}$ are the same as the t period entries of h_t. Let H_t denote the set of histories of length t and $H = \bigcup_{t \geq 0} H_t$ the set of all histories. A pure strategy for the principal is a function $\tau : H \times \mathcal{B} \to \mathbb{R}_+$, which maps histories and the current state to transfer offers T. A pure strategy for the agent is a collection of mappings $\{\alpha_k\}_{k=1}^K$, $\alpha_k : H \times \mathcal{B} \times \mathbb{R}_+ \to \{0, 1\}$, each of which maps the current history, current state and current transfer offer to the action choice $a \in \{0, 1\}$ for a particular type c_k.

For conciseness, we restrict attention to pure strategy perfect Bayesian equilibrium (PBE) in the body of the paper. We consider mixed strategies in Online Appendix OA2.2; see also Remark 2 below. Pure strategy PBE are denoted by the pair (σ, μ), where $\sigma = (\tau, \{\alpha_k\}_{k=1}^K)$ is a strategy profile and $\mu : H \to \Delta(\mathcal{C})$ gives the principal's beliefs about the agent's type after each history. For any PBE (σ, μ), the continuation payoffs of the principal and an agent with cost c_k after history h_t and shock realization b_t are denoted $U^{(\sigma,\mu)}[h_t, b_t]$ and $V_k^{(\sigma,\mu)}[h_t, b_t]$. For any $\mu_0 \in \Delta(\mathcal{C})$, any PBE (σ, μ) and any state $b \in \mathcal{B}$, we denote by $W^{(\sigma,\mu)}[\mu_0, b]$ the principal's payoff at the start of a game with prior μ_0 under the PBE (σ, μ) when the initial state is b.

We restrict attention to pure strategy PBE that satisfy a sequential optimality condition for the principal, defined as follows. For each integer $n \leq K$, define $S_n := \{\lambda \in \Delta(\mathcal{C}) : |\mathsf{supp}\,\lambda| = n\}$. Let Σ_0 denote the set of pure strategy PBE. For all $k = 1, 2, ..., K$, we define the sets Σ_k recursively as follows:

$$\Sigma_k := \left\{ (\sigma, \mu) \in \Sigma_{k-1} : \begin{array}{c} \sigma \text{ is a pure strategy profile and} \\ \forall (h_t, b_t) \text{ with } \mu[h_t] \in S_k \text{ and } \forall (\sigma', \mu') \in \Sigma_{k-1}, \\ U^{(\sigma,\mu)}[h_t, b_t] \geq W^{(\sigma',\mu')}[\mu[h_t], b_t] \end{array} \right\}.$$

Thus, Σ_1 is the set of pure strategy PBE that deliver the highest possible payoff to the principal at histories at which her beliefs are degenerate. For all $k > 1$, Σ_k is the set of pure strategy PBE in Σ_{k-1} that deliver the highest possible payoff to the principal (among all PBE in Σ_{k-1}) at histories at which the support of her beliefs contains k elements. In what follows, we restrict attention to PBE in Σ_K (recall that $|\mathcal{C}| = K$) and use the word *equilibrium* to refer such a PBE. By forcing continuation play to be efficient for the principal at all histories, this solution concept naturally captures lack of commitment.

Alternatively, our solution concept can be thought of as capturing renegotiation-proofness. At histories at which the principal learned the agent's type, players must play a continuation PBE that is optimal for the principal (who is assumed to have all the bargaining power). Similarly, at histories at which the principal believes that the agent may be one of two possible types, players play a continuation PBE that is optimal for the principal among the set of PBE that are principal-optimal at histories at which beliefs are degenerate. And so on.[6]

[6]This solution concept is similar in spirit to the refinement used by Gerardi and Maestri (2015).

We end this section by noting that our equilibrium refinement facilitates a direct comparison with prior papers on the ratchet effect, e.g. Hart and Tirole (1988) and Schmidt (1993). As we will show below, this refinement selects a unique equilibrium that naturally generalizes the equilibrium studied in these papers. In particular, when there are no productivity shocks (i.e., when $\mathcal{B}$ is a singleton), our equilibrium coincides with the equilibrium in Hart and Tirole (1988) and Schmidt (1993).

4 Equilibrium Analysis

4.1 Incentive Constraints

Fix an equilibrium $(\sigma, \mu) = \big((\tau, \{\alpha_k\}_{k=1}^K), \mu\big)$. Recall that for any $h_t \in H$, $\mu[h_t]$ are the principal's beliefs at history h_t. We use $C[h_t] \subset \mathcal{C}$ to denote the support of $\mu[h_t]$, and $\overline{k}[h_t] := \max\{k : c_k \in C[h_t]\}$ to denote the highest type index in $C[h_t]$. Since $c_1 < ... < c_K$, $c_{\overline{k}[h_t]}$ is the highest cost in the support of $\mu[h_t]$. Finally, we let $a_{t,k}$ be the random variable indicating the action in $\{0, 1\}$ that agent type c_k takes in period t.

For any history h_t, any pair $c_i, c_j \in C[h_t]$, and any productivity level $b \in \mathcal{B}$, let

$$V_{i\to j}^{(\sigma,\mu)}[h_t, b] := (1-\delta)\mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\infty} \delta^{t'-t} a_{t',j}(T_{t'} - c_i)|h_t, b_t = b\right]$$

be the expected discounted payoff that an agent with cost c_i obtains after history h_t when $b_t = b$ from following the equilibrium strategy of an agent with cost c_j. Here, $\mathbb{E}^{(\sigma,\mu)}[\cdot|h_t, b_t]$ denotes the expectation over future play under equilibrium (σ, μ) conditional on history h_t and current shock b_t. Note that for any $c_i \in C[h_t]$, the continuation value of an agent with cost c_i at history h_t and current shock b is simply $V_i^{(\sigma,\mu)}[h_t, b] := V_{i\to i}^{(\sigma,\mu)}[h_t, b]$. Also note that

$$\begin{aligned} V_{i\to j}^{(\sigma,\mu)}[h_t, b] &= (1-\delta)\mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}\left(a_{t',j}(T_{t'} - c_j) + a_{t',j}(c_j - c_i)\right)|h_t, b_t = b\right] \\ &= V_j^{(\sigma,\mu)}[h_t, b] + (c_j - c_i)A_j^{\sigma}[h_t, b] \end{aligned} \quad (2)$$

where $V_j^{(\sigma,\mu)}[h_t, b]$ is type c_j's continuation value at (h_t, b) and

$$A_j^{(\sigma,\mu)}[h_t, b] := (1-\delta)\mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\infty} \delta^{t'-t} a_{t',j}|h_t, b_t = b\right]$$

is the expected discounted number of times that type c_j takes the productive action after (h_t, b) under equilibrium (σ, μ). Equation (2) says that type c_i's payoff from deviating to c_j's strategy can be decomposed into two parts: type c_j's continuation value, and an *informational rent* $(c_j - c_i)A_j^{(\sigma,\mu)}[h_t, b_t]$, which depends on how frequently c_j is expected to take the action in the future. In any equilibrium (μ, σ),

$$V_i^{(\sigma,\mu)}[h_t, b_t] \geq V_{i\to j}^{(\sigma,\mu)}[h_t, b_t] \qquad \forall (h_t, b_t), \forall c_i, c_j \in C[h_t] \tag{3}$$

which represents the set of incentive constraints that must be satisfied in equilibrium. We then have the following lemma, which we prove in the Online Appendix. Part (i) says that, in any equilibrium, the highest cost type in the support of the principal's beliefs obtains a continuation payoff equal to zero. Part (ii) says that "local" incentive constraints bind.

Lemma 0. *Fix an equilibrium (σ, μ) and a history h_t, and if necessary renumber the types so that $C[h_t] = \{c_1, c_2, ..., c_{\overline{k}[h_t]}\}$ with $c_1 < c_2 < ... < c_{\overline{k}[h_t]}$. Then, for all $b \in \mathcal{B}$,*

(i) $V_{\overline{k}[h_t]}^{(\sigma,\mu)}[h_t, b] = 0$.

(ii) If $|C[h_t]| \geq 2$, $V_i^{(\sigma,\mu)}[h_t, b] = V_{i\to i+1}^{(\sigma,\mu)}[h_t, b]$ for all $c_i \in C[h_t] \backslash \{c_{\overline{k}[h_t]}\}$.

4.2 Equilibrium Characterization

We now describe the (essentially) unique equilibrium of the game. Recall that $c_{\overline{k}[h_t]}$ is the highest cost in the support of the principal's beliefs at history h_t, and E_k is the set of productivity levels at which it is socially optimal for type $c_k \in \mathcal{C}$ to take the action.

Theorem 1. *The set of equilibria is non-empty. In any equilibrium (μ, σ), for every history h_t and every $b_t \in \mathcal{B}$, we have:*

(i) If $b_t \in E_{\overline{k}[h_t]}$, the principal offers transfer $T_t = c_{\overline{k}[h_t]}$ and all types in $C[h_t]$ take action $a = 1$.

(ii) If $b_t \notin E_{\overline{k}[h_t]}$ and $X(b_t, E_{\overline{k}[h_t]}) > 1 - \delta$, all types in $C[h_t]$ take action $a = 0$.

(iii) If $b_t \notin E_{\overline{k}[h_t]}$ and $X(b_t, E_{\overline{k}[h_t]}) \leq 1 - \delta$, there is a threshold type $c_{k^*} \in C[h_t]$ such that types in $C^- := \{c_k \in C[h_t] : c_k < c_{k^*}\}$ take action $a = 1$, while types in $C^+ := \{c_k \in C[h_t] : c_k \geq c_{k^*}\}$ take action $a = 0$. If C^- is non-empty, the transfer that the principal offers (and which is accepted by types in C^-) satisfies

$$T_t = c_{j^*} + \frac{1}{1-\delta} V_{j^*\to k^*}^{(\sigma,\mu)}[h_t, b_t], \tag{$*$}$$

where $c_{j^*} = \max C^-$.

Theorem 1 says that at histories (h_t, b_t) satisfying the conditions in parts (i) or (ii), all the agent types in $C[h_t]$ take the same action. Hence, the principal learns nothing about the agent's type at such states. When the history (h_t, b_t) satisfies these conditions, an agent with cost $c_i < c_{\overline{k}[h_t]}$ gets large rents by mimicking an agent with cost $c_{\overline{k}[h_t]}$. Since low cost types anticipate that the principal would leave them with no rents in the future if they were to reveal their private information, the principal is unable to learn.

Equilibrium behavior is, however, quite different at histories satisfying the conditions in parts (i) and (ii). When $b_t \in E_{\overline{k}[h_t]}$, there is an *efficient* ratchet effect. At these productivity levels the agent takes the socially efficient action $a = 1$, and the principal compensates him as if he was the highest cost type. This replicates the main finding of the ratchet effect literature. For example, Hart and Tirole (1988) and Schmidt (1993) consider a special case of our model in which the benefit from taking the action is constant over time and strictly larger than the highest cost (i.e., for all times t, $b_t = b > c_K$). Thus, part (i) of Theorem 1 applies: the principal offers a transfer $T = c_K$ that all agent types accept in every period, and she never learns anything about the agent's type.[7]

Part (ii), in contrast, characterizes histories (h_t, b_t) at which there is an *inefficient* ratchet effect. In these histories, low cost types pool with high cost types and don't take the productive action even if the principal is willing to fully compensate their costs. This contrasts with the results in Hart and Tirole (1988) and Schmidt (1993), where the equilibrium is always socially optimal.

Part (iii) characterizes histories (h_t, b_t) at which learning may take place. Specifically, the principal learns about the agent's type when a subset of the types take the action (i.e., when the set C^- is nonempty). In contrast to states in part (ii), the informational rent that type $c_i < c_{\overline{k}[h_t]}$ gets from mimicking an agent with the highest cost $c_{\overline{k}[h_t]}$ are small when $X(b_t, E_{\overline{k}[h_t]}) \leq 1 - \delta$. As a result, the principal is able to get low cost types to reveal their private information. In Appendix A.1.3 we provide a characterization of the threshold cost c_{k^*} in part (iii) of the theorem as the solution to a finite maximization problem. Building on this, we also characterize the principal's equilibrium payoffs as the fixed point of a contraction mapping.

[7] Hart and Tirole (1988) and Schmidt (1993) consider games with a finite deadline. In such games, the principal is only able to induce information revelation at the very last periods prior to the deadline. As the deadline grows to infinity, there is no learning by the principal along the equilibrium path.

Remark 1. (Markovian equilibrium) Note that the equilibrium characterized in Theorem 1 is Markovian: at each period t, the behavior of principal and agent depends solely on the principal's beliefs $\mu[h_t]$ and the current shock realization b_t.

Remark 2. (mixed strategies) In the Online Appendix OA2.2, we extend our analysis and consider a broad class of mixed strategies. In particular, we look at the class of finitely revealing PBE (Peski, 2008); i.e., PBE in which, along any history, the principal's beliefs are updated only finitely many times.

Let Σ_0^M denote the set of PBE that are finitely revealing. For $k = 1, ..., K$, define the sets Σ_k^M recursively as follows:

$$\Sigma_k^M := \left\{ (\sigma, \mu) \in \Sigma_{k-1}^M : \begin{array}{c} \sigma \text{ is finitely revealing} \\ \forall (h_t, b_t) \text{ with } \mu[h_t] \in S_k \text{ and } \forall (\sigma', \mu') \in \Sigma_{k-1}^M \\ U^{(\sigma,\mu)}[h_t, b_t] \geq W^{(\sigma',\mu')}[\mu[h_t], b_t] \end{array} \right\}.$$

This is the natural generalization to the mixed strategy case of our refinement capturing renegotiation-proofness in the pure strategy case, but with the added restriction that the principal updates her beliefs a bounded number of times in any equilibrium outcome.

Let (σ^P, μ^P) denote the PBE in Theorem 1. We show in the appendix that $(\sigma^P, \mu^P) \in \Sigma_K^M$. This implies that any equilibrium in Σ_K^M must give the principal the same payoff as (σ^P, μ^P) at every history. Moreover, we show along the way that generically any equilibrium in the set Σ_K^M is outcome-equivalent to (σ^P, μ^P).

Remark 3. (full-commitment benchmark) We can compare the principal's equilibrium profits to what she would obtain if she had full commitment. A principal with commitment power will in general want to make a high-cost agent take action $a = 1$ inefficiently few times, to reduce the informational rents of low cost types. Time-varying shocks enable the principal to approximate the full-commitment solution. At histories (h_t, b_t) with $b_t \notin E_{\overline{k}[h_t]}$ and $X(b_t, E_{\overline{k}[h_t]}) \leq 1 - \delta$, the principal can truthfully commit to contract infrequently with the highest cost agent $c_{\overline{k}[h_t]}$ in the future. This reduces the rents for lower cost types, and enables the principal to learn about the agent's type.

In Online Appendix OA2.3, we illustrate this result for the case of two types, $\mathcal{C} = \{c_1, c_2\}$. We show that if $X(b, E_2) = \epsilon \leq 1 - \delta$ for some productivity level $b \in E_1 \backslash E_2$, then the principal's equilibrium payoff at histories (h_t, b_t) such that $C[h_t] = \mathcal{C}$ and $b_t = b$ are within ϵM of her full commitment payoff for some constant M.

4.3 Examples

We end this section with a couple of two-type, two-shock examples that illustrate some of the main equilibrium features of our model. The first highlights the fact that equilibrium outcome in our model can be inefficient. The second illustrates a situation in which the principal learns the agent's type, and the equilibrium outcome is efficient.

Example 1. (inefficient ratchet effect) Suppose that there are two states, $\mathcal{B} = \{b_L, b_H\}$, and two types, $\mathcal{C} = \{c_1, c_2\}$ with $c_1 < b_L < c_2 < b_H$, so that $E_1 = \{b_L, b_H\}$ and $E_2 = \{b_H\}$. *Assume further that* $X(b_L, \{b_H\}) > 1 - \delta$.

Consider a history h_t such that $C[h_t] = \{c_1, c_2\}$. Theorem 1(i) implies that, at such a history, both types take the action if $b_t = b_H$, receiving a transfer equal to c_2. On the other hand, Theorem 1(ii) implies that neither type takes the action if $b_t = b_L$. Indeed, when $X(b_L, \{b_H\}) > 1 - \delta$ the benefit that a c_1-agent obtains by pooling with a c_2-agent is so large that there does not exist an offer that a c_1-agent would accept but a c_2-agent would reject. As a result, the principal never learns the agent's type. Inefficiencies arise in all periods t in which $b_t = b_L$: an agent with cost c_1 never takes the action when the state is b_L, even though it is socially optimal for him to do so. □

Example 2. (efficiency and learning) The environment is the same as in Example 2, with the only difference that $X(b_L, \{b_H\}) < 1 - \delta$. Consider a history h_t such that $C[h_t] = \{c_1, c_2\}$. As in Example 2, both types take the action in period t if $b_t = b_H$. The difference is that, if $b_t = b_L$, the principal offers a transfer T_t that a c_2-agent rejects, but a c_1-agent accepts. The principal's offer T_t exactly compensates type c_1 for revealing his type: $(1-\delta)(T_t - c_1) = X(b_L, \{b_H\})(c_2 - c_1)$.[8] Note that $X(b_L, \{b_H\}) < 1 - \delta$ implies that $T_t < c_2$, so an agent with cost c_2 rejects offer T_t. The principal finds it optimal to make such an offer, since it induces an agent with cost $c_L < b_L$ to take the efficient action.

We note that the principal learns the agent's type at time $\bar{t} = \min\{t : b_t = b_L\}$, and the outcome is efficient from time $\bar{t} + 1$ onwards: type c_i takes the action at time $t' > \bar{t}$ if and only if $b_{t'} \in E_i$. Moreover, Lemma 0(i) guarantees that the principal extracts all of the surplus from time $\bar{t} + 1$ onwards, paying the agent a transfer equal to his cost. □

The inefficiency in Example 1 contrasts with the results of the ratchet effect literature in which the outcome is always efficient. The results of Example 2 contrast sharply

[8]The payoff a low cost agent obtains by accepting offer T is $(1 - \delta)(T - c_1) + \delta 0$, since the principal learns that the agent's cost is c_1. On the other hand, the payoff such an agent obtains from rejecting the offer and mimicking a high cost agent is $X(b_L, \{b_H\})(c_2 - c_1)$.

with this literature as well, in which learning never takes place. The key features of this example are that (i) learning by the principal takes place only if productivity is low, (ii) the principal eventually achieves her first best payoff, and (iii) the equilibrium exhibits a form of path-dependence: equilibrium play at time t depends on the entire history of shocks up to period t.[9] These features motivate the results of the next section.

5 Implications

5.1 The Consequences of Bad Shocks

In Example 2 above, the principal learns the agent's type and learning takes place the first time the relationship hits the low productivity state. In addition, as soon as the low productivity state is reached for the first time, the agent's compensation falls permanently. In this section, we present conditions under which these results generalize.

Consider the following assumption, which is a monotonicity condition on the stochastic process Q that governs the evolution of productivity.

Assumption 2 *For all $c_k \in \mathcal{C}$, $X(b, E_k) \leq X(b', E_k)$ for all $b, b' \in \mathcal{B}$ with $b < b'$.*

The assumption is natural; for example, it holds when $\{Q_{b,\tilde{b}}\}_{\tilde{b}\in\mathcal{B}}$ and $\{Q_{b',\tilde{b}}\}_{\tilde{b}\in\mathcal{B}}$ satisfy the monotone likelihood ratio property.[10]

Now refer to history (h_t, b_t) as a *history of information revelation* if $\mu[h_{t+1}] \neq \mu[h_t]$; i.e., if learning takes place at history (h_t, b_t). The following proposition states that under Assumption 2, learning takes place only in periods of low productivity.

Proposition 1. (learning in bad times) *Suppose that Assumption 2 holds. For every history h_t there exists a productivity level $b[h_t] \in \mathcal{B}$ such that (h_t, b_t) is a history of information revelation only if $b_t < b[h_t]$.*

Proof. By Theorem 1, $\mu[h_{t+1}] \neq \mu[h_t]$ only if b_t is such that $X(b_t, E_{\overline{k}[h_t]}) \leq 1 - \delta$. By Assumption 2, there exists $b[h_t]$ such that $X(b_t, E_{\overline{k}[h_t]}) \leq 1-\delta$ if and only if $b_t < b[h_t]$.[11] □

[9] Before state b_L is reached for the first time, the principal pays a transfer equal to the agent's highest cost c_2 to get both types to take the action. After state b_L is visited, if the principal finds that the agent has low cost, then she pays a lower transfer equal to c_1.

[10] That is, for every $b > b'$, $\frac{Q_{b,\tilde{b}}}{Q_{b',\tilde{b}}} = \frac{\text{prob}(b_{t+1}=\tilde{b}|b_t=b)}{\text{prob}(b_{t+1}=\tilde{b}|b_t=b')}$ is increasing in $\tilde{b}$.

[11] When $b[h_t] = \min \mathcal{B}$, $X(b, E_k) > 1 - \delta$ for all $b \in \mathcal{B}$. In this case, the principal's beliefs remain unchanged after history h_t.

To see why the result holds, note that under Assumption 2 the future expected discounted surplus of the relationship is decreasing in the current shock b_t. This implies that the informational rent that agents with type $c_i < c_{\overline{k}[h_t]}$ get from mimicking an agent with the highest cost $c_{\overline{k}[h_t]}$ is also decreasing in b_t. As a result, the principal is only able to learn about the agent's type in periods where the productivity b_t is low.

Next, recall that according to Theorem 1, if (h_t, b_t) is a history of information revelation, then there exists a type $c_{j^*} \in C[h_t]$ such that only agents with cost at most c_{j^*} take action the action at time t. We refer to type c_{j^*} the *marginal type* in period t. Also, for every history (h_t, b_t) and every type $c_j \in C[h_t]$, define

$$P_j[h_t, b_t] := -(1-\delta)\mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\infty} \delta^{t'-t} |\mathbf{1}_{b_{t'} \in E_j} - a_{t',j}|(b_{t'} - c_j) \middle| h_t, b_t\right]$$

which is a measure of how efficient the equilibrium actions of type c_j are. The following proposition, which follows directly from Theorem 1, states two results: (i) that productivity increases after histories of information revelation, and (ii) that the agent's compensation may fall permanently after such histories.

Proposition 2. (productivity and compensation) *Let (h_t, b_t) be a history of information revelation, and c_{j^*} the marginal type at time t. Then, for all (h_τ, b_τ) with $h_\tau \succeq h_t$,*

(i) $P_{j^*}[h_\tau, b_\tau] = 0$, and

(ii) $V_{j^*}^{(\sigma,\mu)}[h_\tau, b_\tau] = 0$.

Part (i) of this result, combined with Proposition 1, implies that agents' productivity will increase after the relationship goes through bad times. The result is in line with Lazear et al. (2016), who find evidence that workers' productivity increases after a recession. Part (ii) combined with Proposition 1 implies that the agents' compensation may be permanently lowered after the relationship experiences negative shocks. This finding is consistent with Kahn (2010) and Oreopoulos et al. (2012), who provide evidence that recessions have a persistent negative effect on worker compensation.

5.2 Long-Run First-Best Payoffs

Another notable feature of Example 2 is that full learning takes place, and as a result, the principal's value increases permanently to the first best level. Here, we characterize general conditions under which the principal obtains her first-best payoff in the long-run, as well as conditions under which she doesn't. Before stating our results, we introduce some additional notation and make a preliminary observation.

An equilibrium outcome can be written as an infinite sequence $h_\infty = \langle b_t, T_t, a_t\rangle_{t=0}^{\infty}$, or equivalently as an infinite sequence of equilibrium histories $h_\infty = \{h_t\}_{t=0}^{\infty}$ such that $h_{t+1} \succeq h_t$ for all t. For any equilibrium outcome h_∞, there exists a time $t^*[h_\infty]$ such that $\mu[h_t] = \mu[h_{t^*[h_\infty]}]$ for all $h_t \succeq h_{t^*[h_\infty]}$. That is, given an equilibrium outcome, learning always stops after some time $t^*[h_\infty]$. Given an equilibrium outcome h_∞, in every period after $t^*[h_\infty]$ the principal's continuation payoff depends only on the realization of the current period shock. Formally, given any equilibrium outcome $h_\infty = \{h_t\}_{t=0}^{\infty}$, the principal's equilibrium continuation value at time $t \geq t^*[h_\infty]$ can be written as $U_{LR}^{(\sigma,\mu)}(b_t|h_{t^*[h_\infty]})$.

For all $b \in \mathcal{B}$ and all $c_k \in \mathcal{C}$, the principal's first best payoffs conditional on the current shock being b and the agent's type being $c = c_k$ are given by

$$U^*(b|c_k) := (1-\delta)\mathbb{E}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}(b_{t'} - c_k)\mathbf{1}_{\{b_{t'} \in E_k\}} \;\middle|\; b_t = b\right].$$

Under the first best outcome the agent takes the action whenever it is socially optimal and the principal always compensates the agent his exact cost. Say that an equilibrium (σ, μ) is *long run first best* if for all $c_k \in \mathcal{C}$, the set of equilibrium outcomes h_∞ such that

$$U_{LR}^{(\sigma,\mu)}(b_t|h_{t^*[h_\infty]}) = U^*(b_t|c_k) \quad \forall t > t^*[h_\infty] \text{ and } \forall b_t \in \mathcal{B},$$

has probability 1 when the agent's type is $c = c_k$. The next result, which we prove in Appendix A.2, reports a sufficient condition for the equilibrium to be long run first best.

Proposition 3. (long run first best) *Suppose that $\{b_t\}$ is ergodic and that for all $c_k \in \mathcal{C}\backslash\{c_K\}$ there exists a productivity level $b \in E_k\backslash E_{k+1}$ such that $X(b, E_{k+1}) < 1-\delta$. Then, the equilibrium is long run first best.*

The condition in the statement of Proposition 3 guarantees that, for any history h_t such that $|C[h_t]| \geq 2$, there exists at least one state $b \in \mathcal{B}$ at which the principal finds it optimal to make an offer that only a strict subset of types accept. So if the process $\{b_t\}$ is ergodic, then it is certain that the principal will eventually learn the agent's type, and from that point onwards she gets her first best payoffs.

If an equilibrium is long run first best then it is also *long run efficient*, i.e. for all $c_k \in \mathcal{C}$, with probability one an agent with cost c_k takes the action at time $t > t^*[h_\infty]$ if and only if $b_t \in E_k$. However, the converse of this statement is not true. Because of this, there are weaker sufficient conditions under which long run efficiency holds. One such condition is that $\{b_t\}$ is ergodic and for all $c_k \in \mathcal{C}$ such that $E_k \neq E_K$, there exists $b \in E_k \backslash E_{\underline{k}}$ such that $X(b, E_{\underline{k}}) < 1 - \delta$, where $\underline{k} = \min\{j \geq k : E_j \neq E_k\}$. This condition guarantees that the principal's beliefs will eventually place all the mass on the set of types that share the same efficiency set with the agent's true type. After this happens, even if the principal does not achieve her first best payoff by further learning the agent's type, the agent takes the action if and only if it is socially optimal to do so. The argument mirrors that of Proposition 3.

Our next result provides a partial counterpart to Proposition 3. The result is an immediate consequence of Theorem 1.

Proposition 4. (no long run first best; no long run efficiency) *Let h_t be an equilibrium history such that $|C[h_t]| \geq 2$ and suppose that $X(b, E_{\overline{k}[h_t]}) > 1 - \delta$ for all $b \in \mathcal{B}$. Then $\mu[h_{t'}] = \mu[h_t]$ for all histories $h_{t'} \succeq h_t$ (and thus $|C[h_{t'}]| \geq 2$), so the equilibrium is not long run first best. If, in addition, there exists $c_i \in C[h_t]$ such that $E_i \neq E_{\overline{k}[h_t]}$, then the equilibrium is not long run efficient either.*

5.3 Long-Run Path Dependence

The third notable feature of Example 2 was that the equilibrium exhibits a form of path-dependence: equilibrium play at time t depends on the entire history of shocks up to period t. Note, however, that the path dependence in this example is short-lived: after state b_L is visited for the first time, the principal learns the agent's type and the equilibrium outcome from that point onwards is independent of the prior history of shocks. Here we show that this is *not* a general property of our model.

Say that an equilibrium (σ, μ) exhibits *long run path dependence* if for some type of the agent $c = c_k \in \mathcal{C}$ there exists $U_1 : \mathcal{B} \to \mathbb{R}$ and $U_2 : \mathcal{B} \to \mathbb{R}$, $U_1 \neq U_2$, such that conditional on the agent's type being c_k, the set of outcomes h_∞ with $U_{LR}^{(\sigma,\mu)}(\cdot|h_{t^*[h_\infty]}) = U_i(\cdot)$ has positive probability for $i = 1, 2$. That is, the equilibrium exhibits long run path dependence if, with positive probability, the principal's long run payoffs may take more than one value conditional on the agent's type.

The next example shows that equilibrium may exhibit long-run path dependence when process $\{b_t\}$ is not ergodic.

Example 3. (path dependence with non-ergodic shocks) Let $\mathcal{C} = \{c_1, c_2\}$, and $\mathcal{B} = \{b_L, b_M, b_H\}$, with $b_L < b_M < b_H$. Suppose that $E_1 = \{b_L, b_M, b_H\}$ and $E_2 = \{b_M, b_H\}$.

Suppose further that the transition matrix $[Q_{b,b'}]$ satisfies: (i) $X(b_L, E_2) < 1 - \delta$, and (ii) $Q_{b_H,b_H} = 1$ and $Q_{b,b'} \in (0, 1)$ for all $(b, b') \neq (b_H, b_H)$. Thus, state b_H is absorbing. By Theorem 1, if $b_t = b_H$, from time t onwards the principal makes an offer equal to $c_{\overline{k}[h_t]}$ and all agent types in $C[h_t]$ accept.

Consider history h_t with $C[h_t] = \{c_1, c_2\}$. By Theorem 1, if $b_t = b_M$ the principal makes an offer $T_t = c_2$ that both types of agents accept. If $b_t = b_L$, the principal makes an offer $T_t = c_1 + \frac{1}{1-\delta} X(b_L, E_2)(c_2 - c_1) \in (c_1, c_2)$ that type c_1 accepts and type c_2 rejects. Therefore, the principal learns the agent's type.

Now suppose that the agent's true type is $c = c_1$, and consider the following two histories, h_t and $\tilde{h}_t$:

$$
\begin{aligned}
h_t &= \langle (b_{t'} = b_M, T_{t'} = c_2, a_{t'} = 1)_{t'=1}^{t-1} \rangle, \\
\tilde{h}_t &= \langle (b_{t'} = b_M, T_{t'} = c_2, a_{t'} = 1)_{t'=1}^{t-2}, (b_{t-1} = b_L, T_{t-1} = \tilde{T}, a_{t-1} = 1) \rangle.
\end{aligned}
$$

Under history h_t, $b_{t'} = b_M$ for all $t' \leq t - 1$, so the principal's beliefs after h_t is realized are equal to her prior. Under history $\tilde{h}_t$ the principal learns that the agent's type is c_1 at time $t - 1$. Suppose that $b_t = b_H$, so that $b_{t'} = b_H$ for all $t' \geq t$. Under history h_t, the principal doesn't know the agent's type at t, and therefore offers a transfer $T_{t'} = c_2$ for all $t' \geq t$, which both agent types accept. However, under history $\tilde{h}_t$ the principal knows that the agent's type is c_1, and therefore offers transfer $T_{t'} = c_1$ for all $t' \geq t$, and the agent accepts it. Therefore, when the agent's type is c_1, the principal's continuation payoff at history $(h_t, b_t = b_H)$ is $b_H - c_2$, while her payoff at history $(\tilde{h}_t, b_t = b_H)$ is $b_H - c_1$. □

Path-dependence in this example is driven by the non-ergodicity of the productivity shocks. Since $b_H > c_2$ is absorbing, Theorem 1 implies that the principal will stop learning once the shock reaches this state. At the same time, the principal is able to screen the different types when the shock reaches state b_L (since $X(b_L, E_2) < 1 - \delta$), but is unable to screen them at state b_M. Therefore, the principal only learns the agent's type at histories such that shock b_L is realized before shock b_H.

We highlight, however, that the model may give rise to path-dependence even when the evolution of productivity is governed by an ergodic process. The following example, which is fully developed in Online Appendix OA2.4, illustrates this.

Example 4. (path dependence with ergodic shocks) Let $\mathcal{C} = \{c_1, c_2, c_3\}$ and $\mathcal{B} = \{b_L, b_{ML}, b_{MH}, b_H\}$, $b_L < b_{ML} < b_{MH} < b_H$. Suppose that $E_1 = E_2 = \{b_{ML}, b_{MH}, b_H\}$ and

$E_3 = \{b_H\}$. Suppose further that the transition matrix $[Q_{b,b'}]$ satisfies: (a) $Q_{b,b'} > 0$ for all $b, b' \in \mathcal{B}$, and (b) $X(b_{MH}, \{b_H\}) > 1 - \delta$ and $X(b_{ML}, \{b_H\}) < 1 - \delta$.

In Online Appendix OA2.4 we show that the unique equilibrium has the following properties:

(i) For histories h_t such that $C[h_t] = \{c_1, c_2\}$, $\mu[h_{t'}] = \mu[h_t]$ for all $h_{t'} \succeq h_t$ (i.e., there is no more learning by the principal from time t onwards);

(ii) For histories h_t such that $C[h_t] = \{c_2, c_3\}$: if $b_t = b_L$ or $b_t = b_{MH}$, types c_2 and c_3 take action $a = 0$; if $b_t = b_{ML}$, type c_2 takes action $a = 1$ and type c_3 takes action $a = 0$; and if $b_t = b_H$, types c_2 and c_3 take action $a = 1$;

(iii) For histories h_t such that $C[h_t] = \{c_1, c_2, c_3\}$: if $b_t = b_L$, type c_1 takes action $a = 1$ while types c_2 and c_3 take action $a = 0$; if $b_t = b_{ML}$, types c_1 and c_2 take action $a = 1$ and type c_3 takes action $a = 0$; if $b_t = b_{MH}$, all agent types take action $a = 0$; and if $b_t = b_H$, all agent types take action $a = 1$.

An immediate consequence of these facts is that when the agent's type is c_1, the principal learns the agent's type at histories such that state b_L is visited before b_{ML}. In contrast, at histories at which b_{ML} is visited before b_L, the principal only learns that the agent's type is in $\{c_1, c_2\}$. From this point onwards, her beliefs are never again updated. As a result, the principal's long run value when the agent's type is c_1 depends on whether or not shock b_L was realized before shock b_{ML}. □

To understand Example 4, note that the informational rents that type c_1 gets by mimicking type c_2 depend on how often c_2 is expected to take the productive action in the future. In turn, how often c_2 takes the productive action depends on the principal's beliefs. If the principal learns along the path of play that the agent's type is not c_3, from that time onwards type c_2 will take the action whenever the state is in $E_2 = \{b_{ML}, b_{MH}, b_H\}$.

In contrast, at histories at which the principal has not ruled out types c_2 and c_3, type c_2 will not take the productive action at time t if $b_t = b_{MH}$ (since, by assumption, $X(b_{MH}, E_3) > 1 - \delta$). Therefore, type c_2 is expected to take the action significantly less frequently in the future at a history after which the support of the principal's beliefs is $\{c_1, c_2, c_3\}$ than at a history at which it is $\{c_1, c_2\}$.

As a consequence of this, the cost of getting a c_1-agent to reveal his private information depends on the principal's beliefs. In particular, when the current productivity level is b_L, getting a c_1-agent to reveal his private information is cheaper at histories where all three types are in the support of the principal's beliefs than at histories at which only c_1 and c_2 are in the support. This difference makes it optimal for the principal to get a c_1-agent to reveal his type when productivity is b_L and the support of the principal's beliefs is $\{c_1, c_2, c_3\}$, and at the same time it makes it suboptimal to get this agent type to reveal himself when productivity is b_L and the support is $\{c_1, c_2\}$.

6 Final Remarks

Productivity shocks are a natural feature of most economic environments, and the incentives that economic agents face in completely stationary environments can be very different than the incentives they face in environments subject to these shocks. Our results demonstrate the consequences of this fact for the traditional ratchet effect literature. A key takeaway from this literature is that outside institutions that provide contract enforcement can help improve the principal's welfare. However, our results show that even without such institutions, a strategic principal can use productivity shocks to her advantage to gradually learn the agent's private information and improve her own welfare.

Our model has several natural extensions. For example, we have assumed that the benefit b_t that the principal obtains when the agent takes the action is publicly observed. This assumption is natural in settings in which the principal's benefits depends on the cost of some key input (like oil or cement), or when these benefits are linked to the aggregate state of the economy. However, it is also interesting to consider settings in which the benefit b_t is privately observed by the principal.

For concreteness, consider the setting of Examples 1 and 2, in which b_t can take values $\{b_L, b_H\}$ and the agent's cost can take values $\{c_1, c_2\}$, with $b_H > c_2 > b_L > c_1$ (i.e., $E_1 = \{b_L, b_H\}$ and $E_2 = \{b_L\}$). Consider first that $X(b, E_2) > 1 - \delta$ for $b = b_L, b_H$. In this case, the equilibrium outcome in Theorem 1 remains an equilibrium even when b_t is privately observed. Indeed, under this condition, a low cost agent is not willing to disclose his private information, regardless of whether he observes the realization of b_t or not.

On the other hand, when $X(b_L, E_2) < 1 - \delta$, the equilibrium outcome in Theorem 1 fails to be an equilibrium. Indeed, in this case, when the benefit is b_H the principal would prefer to make an offer as if the benefit were b_L, to induce the low cost agent to reveal his private information. In this setting, one can construct PBE under which the principal's transfer offer perfectly reveals her private information at each point (i.e., her transfer T_t reveals the realization of benefit b_t at every period t). Under such equilibria, at histories

(h_t, b_t) with $b_t = b_L$ and $C[h_t] = \{c_1, c_2\}$ the principal makes a low offer $T_t \in (c_1, c_2)$ that leaves low cost agents indifferent between accepting and rejecting. Such an offer is rejected by high cost agents, and accepted with probability $\alpha_t \in [0,1]$ by low cost agents. The probability of acceptance α_t is calibrated to provide incentives to the principal to make a high offer $T_t = c_2$ at histories (h_t, b_t) with $b_t = b_H$.[12]

A Appendix

A.1 Proof of Theorem 1

The proof proceeds in three steps. First we analyze the case where $b_t \in E_{\overline{k}[h_t]}$, establishing part (i) of the theorem. Then we analyze the case where $b_t \notin E_{\overline{k}[h_t]}$, establishing parts (ii) and (iii). Finally, we show that equilibrium exists and has unique payoffs. In doing so, we also characterize the threshold type c_{k^*} defined in part (iii).

A.1.1 Proof of part (i) (the case of $b_t \in E_{\overline{k}[h_t]}$)

We prove part (i) by strong induction on the cardinality of $C[h_t]$. If $C[h_t]$ is a singleton $\{c_k\}$, the result holds: in any PBE in Σ_K, the principal offers the agent a transfer $T_{t'} = c_k$ at all times $t' \geq t$ such that $b_{t'} \in E_k$ and the agent accepts, and she offers some transfer $T_{t'} < c_k$ at all times $t' \geq t$ such that $b_{t'} \notin E_k$, and the agent rejects.

Suppose next that the claim is true for all histories $h_{t'}$ such that $|C[h_{t'}]| \leq n-1$. Let (h_t, b_t) be a history such that $|C[h_t]| = n$ and $b_t \in E_{\overline{k}[h_t]}$. We show that, at such a history (h_t, b_t) the principal makes an offer $T_t = c_{\overline{k}[h_t]}$ that all agent types accept.

Note first that, in a PBE in Σ_K, it cannot be that at (h_t, b_t) the principal makes an offer that no type in $C[h_t]$ accepts. Indeed, suppose that no type in $C[h_t]$ takes the action. Consider an alternative PBE which is identical to the original PBE, except that at history (h_t, b_t) the principal makes an offer $T = c_{\overline{k}[h_t]}$, and all agent types in $C[h_t]$ accept any offer weakly larger than $T = c_{\overline{k}[h_t]}$. The principal's beliefs after this period are equal to $\mu[h_t]$ regardless of the agent's action. Since $T = c_{\overline{k}[h_t]}$, it is optimal for all agent types to accept this offer. Moreover, it is optimal for the principal to make offer T. Finally, since $b_t \in E_{\overline{k}[h_t]}$, the payoff that the principal gets from this PBE is larger than her payoff under the original PBE. But this cannot be, since the original PBE is in Σ_K. Hence, if $b_t \in E_{\overline{k}[h_t]}$, at least a subset of types in $C[h_t]$ take the action at time t if $b_t \in E_{\overline{k}[h_t]}$.

[12]Further details about such equilibria are available from the authors upon request.

We now show that, in a PBE in Σ_K, it cannot be that at (h_t, b_t) the principal makes an offer T_t that only a strict subset $C \subsetneq C[h_t]$ of types accept. Towards a contradiction, suppose that a strict subset $C \subsetneq C[h_t]$ of types accept T_t, and let $c_j = \max C$. There are two possible cases: (a) $c_j = c_{\overline{k}[h_t]}$, and (b) $c_j < c_{\overline{k}[h_t]}$. Consider case (a). By Lemma 0, the continuation payoff of an agent with cost $c_{\overline{k}[h_t]}$ is zero at all histories. This implies that $T_t = c_{\overline{k}[h_t]}$. Let $c_i = \max C[h_t] \backslash C$ (note that $C[h_t] \backslash C$ is non-empty). Since c_i rejects the offer today and becomes the highest cost in the support of the principal's beliefs tomorrow, Lemma 0 implies that $V_i^{(\sigma,\mu)}[h_t, b_t] = 0$. But this cannot be, since this agent can guarantee a payoff of at least $(1-\delta)(T_t - c_i) = (1-\delta)(c_{\overline{k}[h_t]} - c_i) > 0$ by accepting the offer. Hence, if only a strict subset $C \subsetneq C[h_t]$ of types accept, $c_j = \max C < c_{\overline{k}[h_t]}$.

Consider next case (b). By Lemma 0, the payoff of type c_j from taking the productive action at time t is $(1-\delta)(T_t - c_j) + 0$. Indeed, at period $t+1$, c_j will be the highest cost in the support of the principal's beliefs if he takes the action at t. Since an agent with cost c_j can mimic the strategy of type $c_{\overline{k}[h_t]}$, incentive compatibility implies that

$$\begin{aligned}(1-\delta)(T_t - c_j) &\geq V_{\overline{k}[h_t]}^{(\sigma,\mu)}[h_t, b_t] + (c_{\overline{k}[h_t]} - c_j) A_{\overline{k}[h_t]}^{(\sigma,\mu)}[h_t, b_t] \\ &\geq (c_{\overline{k}[h_t]} - c_j) X(b_t, E_{\overline{k}[h_t]}) > (1-\delta)(c_{\overline{k}[h_t]} - c_j) \qquad (4)\end{aligned}$$

The first inequality follows from equation (3) in the main text. The second inequality follows from Lemma 0 and the fact that $A_{\overline{k}[h_t]}^{(\sigma,\mu)}[h_t, b_t] \geq X(b_t, E_{\overline{k}[h_t]})$. To see why this last inequality holds, note that $c_{\overline{k}[h_t]} \notin C$, so at most $n-1$ types accept the principal's offer. Thus, the inductive hypothesis implies that if the agent rejects the offer, then at all periods $t' > t$ the principal will get all the remaining types to take the action whenever $b_t \in E_{\overline{k}[h_t]}$, and so $A_{\overline{k}[h_t]}^{(\sigma,\mu)}[h_t, b_t] \geq X(b_t, E_{\overline{k}[h_t]})$. The last inequality in equation (4) follows from the fact $X(b_t, E_{\overline{k}[h_t]}) \geq X(b_t, b_t^+) > 1 - \delta$ where the first inequality holds because $b_t \in E_{\overline{k}[h_t]}$ and the second follows by Assumption 1.

On the other hand, because Lemma 0 implies that an agent with type $c_{\overline{k}[h_t]}$ has a continuation value of zero, the transfer T_t that the principal offers must be weakly smaller than $c_{\overline{k}[h_t]}$; otherwise, if $T_t > c_{\overline{k}[h_t]}$, an agent with type $c_{\overline{k}[h_t]}$ could guarantee himself a strictly positive payoff by accepting the offer. But this contradicts (4). Hence, it cannot be that only a strict subset of types in $C[h_t]$ accept the principal's offer at (h_t, b_t).

By the arguments above, all agents in $C[h_t]$ take action $a = 1$ at (h_t, b_t) with $b_t \in E_{\overline{k}[h_t]}$. Since an agent with cost $c_{\overline{k}[h_t]}$ obtains a payoff of zero after every history (Lemma 0), the transfer that the principal offers at time t is $T_t = c_{\overline{k}[h_t]}$. □

A.1.2 Proof of parts (ii) & (iii) (the case of $b_t \notin E_{\overline{k}[h_t]}$)

In both parts (ii) and (iii) of the theorem, the highest cost type in the principal's support $c_{\overline{k}[h_t]}$ does not take the productive action when $b_t \notin E_{\overline{k}[h_t]}$. We prove this in Lemma A.1 below, and use the lemma to prove parts (ii) and (iii) separately.

Lemma A.1. *Fix any equilibrium (σ, μ) and history h_t. If $b_t \notin E_{\overline{k}[h_t]}$, then an agent with cost $c_{\overline{k}[h_t]}$ does not take the productive action.*

Proof. Suppose for the sake of contradiction that an agent with type $c_{\overline{k}[h_t]}$ does take the action at time t if $b_t \notin E_{\overline{k}[h_t]}$. Since, by Lemma 0, this type's payoff must equal zero at all histories, it must be that the offer that is accepted is $T_t = c_{\overline{k}[h_t]}$. We now show that if the principal makes such an offer, then all agent types will accept the offer and take the productive action. To see this, suppose some types reject the offer. Let c_j be the highest cost type that rejects the offer. By Lemma 0, type c_j's continuation payoff is zero, because this type becomes the highest cost in the support of the principal's beliefs following a rejection. However, this type can guarantee himself a payoff of at least $(1-\delta)(T_t - c_j) = (1-\delta)(c_{\overline{k}[h_t]} - c_j) > 0$ by accepting the current offer. Hence, it cannot be that some types reject offer $T_t = c_{\overline{k}[h_t]}$ when type $c_{\overline{k}[h_t]}$ accepts it.

It then follows that if type $c_{\overline{k}[h_t]}$ accepts the offer, then the principal will not learn anything about the agent's type. Since $b_t \notin E_{\overline{k}[h_t]}$, her flow payoff from making the offer is $(1-\delta)(b_t - c_{\overline{k}[h_t]}) < 0$. Consider an alternative PBE which is identical to the original PBE, except that at history (h_t, b_t) the principal makes an offer $T = 0$, and all agent types in $C[h_t]$ reject this offer. The principal's beliefs after this period are equal to $\mu[h_t]$ regardless of the agent's action. Note that it is optimal for all types to reject this offer. Moreover, since $b_t \notin E_{\overline{k}[h_t]}$, the payoff that the principal gets from this PBE is larger than her payoff under the original PBE. But this cannot be, since the original PBE is in Σ_K. Hence, if $b_t \notin E_{\overline{k}[h_t]}$, an agent with type $c_{\overline{k}[h_t]}$ does not take the action at time t. □

Proof of part (ii). Fix a history h_t and let $b_t \in \mathcal{B} \backslash E_{\overline{k}[h_t]}$ be such that $X(b_t, E_{\overline{k}[h_t]}) > 1-\delta$. By Lemma A.1, type $c_{\overline{k}[h_t]}$ doesn't take the productive action at time t if $b_t \notin E_{\overline{k}[h_t]}$. Suppose, for the sake of contradiction, that there is a nonempty set of types $C \subsetneq C[h_t]$ that do take the productive action. Let $c_j = \max C$. By Lemma 0 type c_j obtains a continuation payoff of zero starting in period $t+1$. Hence, type c_j receives a payoff $(1-\delta)(T_t - c_j) + \delta 0$ from taking the productive action in period t. Since this payoff must be weakly larger than the payoff the agent would obtain by not taking the action and mimicking the strategy of agent $c_{\overline{k}[h_t]}$ in all future periods, it follows that

$$\begin{aligned}(1-\delta)(T_t - c_j) &\geq V^{(\sigma,\mu)}_{\overline{k}[h_t]}[h_t, b_t] + (c_{\overline{k}[h_t]} - c_j)A^{(\sigma,\mu)}_{\overline{k}[h_t]}[h_t, b_t] \\ &\geq (c_{\overline{k}[h_t]} - c_j)X(b_t, E_{\overline{k}[h_t]}) \qquad (5) \\ &> (1-\delta)(c_{\overline{k}[h_t]} - c_j), \qquad (6)\end{aligned}$$

where the first line follows from incentive compatibility, the second line follows from the fact that $a_{t',\overline{k}[h_t]} = 1$ for all times $t' \geq t$ such that $b_{t'} \in E_{\overline{k}[h_t]}$ (by the result of part (i) proven above), and the third line follows since $X(b_t, E_{\overline{k}[h_t]}) > 1 - \delta$ by assumption. The inequalities in (6) imply that $T_t > c_{\overline{k}[h_t]}$. But then by Lemma 0, it would be strictly optimal for type $c_{\overline{k}[h_t]}$ to deviate by accepting the transfer and taking the productive action, a contradiction. So it must be that all agent types in $C[h_t]$ take action $a_t = 0$. □

Proof of part (iii). Fix a history h_t and let $b_t \in \mathcal{B} \backslash E_{\overline{k}[h_t]}$ be such that $X(b_t, E_{\overline{k}[h_t]}) \leq 1 - \delta$. We start by showing that the set of types that accept the offer has the form $C^- = \{c_k \in C[h_t] : c_k < c_{k^*}\}$ for some $c_{k^*} \in C[h_t]$. The result is clearly true if no agent type takes the action, in which case set $c_{k^*} = \min C[h_t]$; or if only an agent with type $\min C[h_t]$ takes the action, in which case set c_{k^*} equal to the second lowest cost in $C[h_t]$.

Therefore, suppose that an agent with type larger than $\min C[h_t]$ takes the action, and let $c_{j^*} \in C[h_t]$ be the highest cost agent that takes the action. Since $b_t \notin E_{\overline{k}[h_t]}$, by Lemma A.1 it must be that $c_{j^*} < c_{\overline{k}[h_t]}$. By Lemma 0, type c_{j^*}'s payoff is $(1-\delta)(T_t - c_{j^*})$, since from date $t+1$ onwards this type will be the highest cost type in the support of the principal's beliefs if the principal observes that the agent took the action at time t. Let $c_{k^*} = \min\{c_k \in C[h_t] : c_k > c_{j^*}\}$. By incentive compatibility, it must be that

$$(1-\delta)(T_t - c_{j^*}) \geq V^{(\sigma,\mu)}_{k^*}[h_t, b_t] + (c_{k^*} - c_{j^*})A^{(\sigma,\mu)}_{k^*}[h_t, b_t], \qquad (7)$$

since type c_{j^*} can obtain the right-hand side of (7) by mimicking type c_{k^*}. Furthermore, type c_{k^*} can guarantee himself a payoff of $(1-\delta)(T_t - c_{k^*})$ by taking the action at time t and never taking the action again. Therefore, it must be that

$$\begin{aligned}V^{(\sigma,\mu)}_{k^*}[h_t, b_t] \geq (1-\delta)(T_t - c_{k^*}) &\geq (1-\delta)(c_{j^*} - c_{k^*}) + V^{(\sigma,\mu)}_{k^*}[h_t, b_t] + (c_{k^*} - c_{j^*})A^{(\sigma,\mu)}_{k^*}[h_t, b_t] \\ \implies 1 - \delta &\geq A^{(\sigma,\mu)}_{k^*}[h_t, b_t] \qquad (8)\end{aligned}$$

where the second inequality in the first line follows from (7).

We now show that all types $c_i \in C[h_t]$ with $c_i < c_{j^*}$ also take the action at time t. Suppose for the sake of contradiction that this is not true, and let $c_{i^*} \in C[h_t]$ be the highest cost type lower than c_{j^*} that does not take the action. The payoff that this type would get by taking the action at time t and then mimicking type c_{j^*} is

$$\begin{aligned} V_{i^* \to j^*}^{(\sigma,\mu)}[h_t, b_t] &= (1-\delta)(T_t - c_{j^*}) + (c_{j^*} - c_{i^*}) A_{j^*}^{(\sigma,\mu)}[h_t, b_t] \\ &= (1-\delta)(T_t - c_{j^*}) + (c_{j^*} - c_{i^*})(1 - \delta + X(b_t, E_{j^*})) \\ &\geq (c_{j^*} - c_{i^*})(1 - \delta + X(b_t, E_{j^*})) + V_{k^*}^{(\sigma,\mu)}[h_t, b_t] + (c_{k^*} - c_{j^*}) A_{k^*}^{(\sigma,\mu)}[h_t, b_t] \end{aligned} \tag{9}$$

where the first line follows from the fact that type c_{j^*} is the highest type in the support of the principal's beliefs in period $t+1$, so he receives a payoff of 0 from $t+1$ onwards; the second follows from part (i) and Lemma 2, which imply that type c_{j^*} takes the action in periods $t' \geq t+1$ if and only if $b_{t'} \in E_{j^*}$ (note that type c_{j^*} also takes the action at time t); and the third inequality follows from (7).

On the other hand, by Lemma 0(ii), the payoff that type c_{i^*} gets by rejecting the offer at time t is equal to the payoff she would get by mimicking type c_{k^*}, since the principal will believe for sure that the agent's type is not in $\{c_{i^*+1}, ..., c_{j^*}\} \subseteq C[h_t]$ after observing a rejection. That is, type c_{i^*}'s payoff is

$$V_{i^*}^{(\sigma,\mu)}[h_t, b_t] = V_{i^* \to k^*}^{(\sigma,\mu)}[h_t, b_t] = V_{k^*}^{(\sigma,\mu)}[h_t, b_t] + (c_{k^*} - c_{i^*}) A_{k^*}^{(\sigma,\mu)}[h_t, b_t] \tag{10}$$

From equations (9) and (10), it follows that

$$V_{i^*}^{(\sigma,\mu)}[h_t, b_t] - V_{i^* \to j^*}^{(\sigma,\mu)}[h_t, b_t] \leq (c_{j^*} - c_{i^*}) \left[A_{k^*}^{(\sigma,\mu)}[h_t, b_t] - [1 - \delta + X(b_t, E_{j^*})] \right] < 0,$$

where the strict inequality follows after using (8). Hence, type c_{i^*} strictly prefers to mimic type c_{j^*} and take the action at time t than to not take it, a contradiction. Hence, all types $c_i \in C[h_t]$ with $c_i \leq c_{j^*}$ take the action at t, and so the set of types taking the action takes the form $C^- = \{c_j \in C[h_t] : c_j < c_{k^*}\}$.

Finally, it is clear that in equilibrium, the transfer that the principal will pay at time t if all agents with type $c_i \in C^-$ take the action is given by $(*)$. The payoff that an agent with type $c_{j^*} = \max C^-$ gets by accepting the offer is $(1-\delta)(T_t - c_{j^*})$, while her payoff from rejecting the offer and mimicking type $c_{k^*} = \min C[h_t] \backslash C^-$ is $V_{k^*}^{(\sigma,\mu)}[h_t, b_t] +$

$(c_{k^*} - c_{j^*})A_{k^*}^{(\sigma,\mu)}[h_t, b_t]$. Hence, the lowest offer that a c_{j^*}-agent accepts is $(1-\delta)T_t = (1-\delta)c_{j^*} + V_{k^*}^{(\sigma,\mu)}[h_t, b_t] + (c_{k^*} - c_{j^*})A_{k^*}^{(\sigma,\mu)}[h_t, b_t]$. □

A.1.3 Proof of Existence and Uniqueness

For each history h_t and each $c_j \in C[h_t]$, let $C_{j+}[h_t] = \{c_i \in C[h_t] : c_i \geq c_j\}$. For each history h_t and state realization $b_t \in \mathcal{B}$, let

$$A_{j+}^{(\sigma,\mu)}[h_t, b_t] := (1-\delta)\mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t+1}^{\infty} \delta^{t'-t} a_{t',j} | (h_t, b_t) \text{ and } C[h_{t+1}] = C_{j+}[h_t]\right].$$

That is, $A_{j+}^{(\sigma,\mu)}[h_t, b_t]$ is the expected discounted fraction of time that an agent with type c_j takes the action after history (h_t, b_t) if the beliefs of the principal at time $t+1$ have support $C_{j+}[h_t]$. We then have:

Lemma A.2. *Fix any equilibrium (σ, μ) and history (h_t, b_t). Then, there exists an offer $T \geq 0$ such that types $c_i \in C[h_t], c_i < c_j$, accept at time t and types $c_i \in C[h_t], c_i \geq c_j$, reject if and only if $A_{j+}^{(\sigma,\mu)}[h_t, b_t] \leq 1 - \delta$.*

Proof. First, suppose such an offer T exists, and let c_k be the highest type in $C[h_t]$ that accepts T. Let c_j be the lowest type in $C[h_t]$ that rejects the offer, and note that $c_k < c_j$. By Lemma 0, the expected discounted payoff that an agent with type c_k gets from accepting the offer is $(1-\delta)(T - c_k) + \delta 0$. The payoff that type c_k obtains by rejecting the offer and mimicking type c_j from time $t+1$ onwards is $V_j^{(\sigma,\mu)}[h_t, b_t] + (c_j - c_k)A_{j+}^{(\sigma,\mu)}[h_t, b_t]$. Therefore, the offer T that the principal makes must satisfy

$$(1-\delta)(T - c_k) \geq V_j^{(\sigma,\mu)}[h_t, b_t] + (c_j - c_k)A_{j+}^{(\sigma,\mu)}[h_t, b_t]. \tag{11}$$

Note that an agent with type c_j can guarantee herself a payoff of $(1-\delta)(T - c_j)$ by taking the action in period t and then never taking it again; therefore, incentive compatibility implies

$$\begin{aligned} V_j^{(\sigma,\mu)}[h_t, b_t] \geq (1-\delta)(T - c_j) &\geq V_j^{(\sigma,\mu)}[h_t, b_t] + (c_j - c_k)\left[A_{j+}^{(\sigma,\mu)}[h_t, b_t] - (1-\delta)\right] \\ \implies 1 - \delta &\geq A_{j+}^{(\sigma,\mu)}[h_t, b_t] \end{aligned}$$

where the second inequality in the first line follows after substituting T from (11).

Suppose next that $A_{j+}^{(\sigma,\mu)}[h_t, b_t] \leq 1 - \delta$, and suppose the principal makes offer T such that $(1-\delta)(T - c_k) = V_j^{(\sigma,\mu)}[h_t, b_t] + (c_j - c_k)A_{j+}^{(\sigma,\mu)}[h_t, b_t]$, which only agents with type

$c_\ell \in C[h_t], c_\ell \leq c_k$ are supposed to accept. The payoff that an agent with cost c_k obtains by accepting the offer is $(1-\delta)(T-c_k)$, which is exactly what he would obtain by rejecting the offer and mimicking type c_j. Hence, type c_k has an incentive to accept such an offer. Similarly, one can check that all types $c_\ell \in C[h_t], c_\ell < c_k$ also have an incentive to accept the offer. If the agent accepts such an offer and takes the action in period t, the principal will be believe that the agent's type lies in $\{c_\ell \in C[h_t] : c_\ell \leq c_i\}$. Note that, in all periods $t' > t$, the principal will never offer $T_{t'} > c_k$.

Consider the incentives of an agent with type $c_i \geq c_j > c_k$ at time t. The payoff that this agent gets from accepting the offer is $(1-\delta)(T-c_i)$, since from $t+1$ onwards the agent will never accept any equilibrium offer. This is because all subsequent offers will be lower than $c_k < c_j \leq c_i$. On the other hand, the agent's payoff from rejecting the offer is

$$\begin{aligned} V_i^{(\sigma,\mu)}[h_t,b_t] &\geq V_{i\to j}^{(\sigma,\mu)}[h_t,b_t] = V_j^{(\sigma,\mu)}[h_t,b_t] + (c_j - c_i)A_{j+}^{(\sigma,\mu)}[h_t,b_t] \\ &\geq (1-\delta)(T-c_i) = (1-\delta)(c_k - c_i) + V_j^{(\sigma,\mu)}[h_t,b_t] + (c_j - c_k)A_{j+}^{(\sigma,\mu)}[h_t,b_t], \end{aligned}$$

where the second inequality follows since $A_{j+}^{(\sigma,\mu)}[h_t,b_t] \leq 1-\delta$. □

The proof of existence and uniqueness relies on Lemma A.2 and uses strong induction on the cardinality of $C[h_t]$. Clearly, Σ_1 is non-empty, and all PBE in Σ_1 give the same payoff to the principal at histories (h_t, b_t) such that $C[h_t] = \{c_k\}$: in this case, the principal offers the agent a transfer $T_{t'} = c_k$ (which the agent accepts) at times $t' \geq t$ such that $b_{t'} \in E_k$ and offers some transfer $T_{t'} < c_k$ (which the agent rejects) at times $t' \geq t$ such that $b_{t'} \notin E_k$.

Suppose next that Σ_{k-1} is non-empty for all $k \leq n-1$, and that for all $k \leq n-1$, all PBE in Σ_k give the principal the same payoff at histories (h_t, b_t) with $|C[h_t]| = k$. We now show that Σ_n is non-empty, and that all all PBE in Σ_n give the principal the same payoff at histories (h_t, b_t) with $|C[h_t]| = n$.

Consider a history (h_t, b_t) with $|C[h_t]| = n$. If $b_t \in E_{\overline{k}[h_t]}$, then by part (i) it must be that all agent types in $C[h_t]$ take the action in period t and $T_t = c_{\overline{k}[h_t]}$; hence, at such states

$$U^{(\sigma,\mu)}[h_t,b_t] = (1-\delta)(b_t - c_{\overline{k}[h_t]}) + \delta\mathbb{E}\big[U^{(\sigma,\mu)}[h_{t+1},b_{t+1}]|b_t\big]$$

If $b_t \notin E_{\overline{k}[h_t]}$ and $X(b_t, E_{\overline{k}[h_t]}) > 1-\delta$, then by part (ii), all agent types in $C[h_t]$ don't take the action (in this case, the principal makes an offer T small enough that all agents reject); hence, at such states

$$U^{(\sigma,\mu)}[h_t, b_t] = \delta\mathbb{E}\big[U^{(\sigma,\mu)}[h_{t+1}, b_{t+1}]|b_t\big]$$

In either case, the principal doesn't learn anything about the agent's type, since all types of agents in $C[h_t]$ take the same action, so her beliefs don't change.

Finally, consider states $b_t \notin E_{\overline{k}[h_t]}$ with $X(b_t, E_{\overline{k}[h_t]}) \leq 1-\delta$. Two things can happen at such a state: (i) all types of agents in $C[h_t]$ don't take the action, or (ii) a strict subset of types in $C[h_t]$ don't take the action and the rest do.[13] In case (i), the beliefs of the principal at time $t+1$ would be the same as the beliefs of the principal at time t, and her payoffs are

$$U^{(\sigma,\mu)}[h_t, b_t] = \delta\mathbb{E}\big[U^{(\sigma,\mu)}[h_{t+1}, b_{t+1}]|b_t\big]$$

In case (ii), the types of the agent not taking the action has the form $C_{j+}[h_t] = \{c_i \in C[h_t] : c_i \geq c_j\}$ for some $c_j \in C[h_t]$. So in case (ii) the support of the beliefs of the principal at time $t+1$ would be $C_{j+}[h_t]$ if the agent doesn't take the action, and $C[h_t]\backslash C_{j+}[h_t]$ if he does.

By Lemma A.2, there exists an offer that types $C_{j+}[h_t]$ reject and types $C[h_t]\backslash C_{j+}[h_t]$ accept if and only if $A^{(\sigma,\mu)}_{j+}[h_t, b_t] \leq 1-\delta$. Note that, by the induction hypothesis, $A^{(\sigma,\mu)}_{j+}[h_t, b_t]$ is uniquely determined.[14] Let $C^*[h_t, b_t] = \{c_i \in C[h_t] : A^{(\sigma,\mu)}_{i+}[h_t, b_t] \leq 1-\delta\}$. Without loss of generality, renumber the types in $C[h_t]$ so that $C[h_t] = \{c_1, ..., c_{\overline{k}[h_t]}\}$, with $c_1 < ... < c_{\overline{k}[h_t]}$. For each $c_i \in C^*[h_t, b_t]$, let

$$T^*_{t,i-1} = c_{i-1} + \frac{1}{1-\delta}\left(V^{(\sigma,\mu)}_i[h_t, b_t] + A^{(\sigma,\mu)}_{i+}[h_t, b_t](c_i - c_{i-1})\right)$$

be the offer that leaves an agent with type c_{i-1} indifferent between accepting and rejecting when all types in $C_{i+}[h_t]$ reject the offer and all types in $C[h_t]\backslash C_{i+}[h_t]$ accept. Note that $T^*_{t,i-1}$ is the best offer for a principal who wants to get all agents with types in $C[h_t]\backslash C_{i+}[h_t]$ to take the action and all agents with types in types in $C_{i+}[h_t]$ to not take the action.

Let $\mathcal{T} = \{T^*_{t,i-1} : c_i \in C^*[h_t, b_t]\}$. At states $b_t \notin E_{\overline{k}[h_t]}$ with $X(b_t, E_{\overline{k}[h_t]}) \leq 1$, the principal must choose optimally whether to make an offer in $\mathcal{T}$ or to make a low offer (for

[13] By Lemma A.1, in equilibrium an agent with cost $c_{\overline{k}[h_t]}$ doesn't take the action.

[14] $A^{(\sigma,\mu)}_{j+}[h_t, b_t]$ is determined in equilibrium when the principal has beliefs with support $C_{j+}[h_t]$, and the induction hypothesis states that the continuation equilibrium is unique when the cardinality of the support of principal's beliefs is less than n.

example, $T_t = 0$) that all agents reject: an offer $T_t = T^*_{t,i-1}$ would be accepted by types in $C[h_t] \backslash C_{i+}[h_t]$ and rejected by types in $C_{i+}[h_t]$, while an offer $T_t = 0$ will be rejected by all types. For each offer $T^*_{t,i-1} \in \mathcal{T}$, let $p(T^*_{t,i-1})$ be the probability that offer $T^*_{t,i-1}$ is accepted; i.e., the probability that the agent has cost weakly smaller than c_{i-1}. Let $U^{(\sigma,\mu)}[h_t, b_t, T^*_{t,i-1}, a_t = 1]$ and $U^{(\sigma,\mu)}[h_t, b_t, T^*_{t,i-1}, a_t = 0]$ denote the principal's expected continuation payoffs if the offer $T^*_{t,i-1} \in \mathcal{T}$ is accepted and rejected, respectively, at history (h_t, b_t). Note that these payoffs are uniquely pinned down by the induction hypothesis: after observing whether the agent accepted or rejected the offer, the cardinality of the support of the principal's beliefs will be weakly lower than $n-1$. For all $b \in \mathcal{B}$, let

$$U^*(h_t, b_t) = \max_{T \in \mathcal{T}} \left\{ p(T)((1-\delta)(b-T) + U^{(\sigma,\mu)}[b, \mu[h_t], T, 1]) + (1-p(T))U^{(\sigma,\mu)}[b, \mu[h_t], T, 0] \right\}$$

and let $T(b)$ be a maximizer of this expression.

Partition the states $\mathcal{B}$ as follows:

$$\begin{aligned} B_1 &= E_{\overline{k}[h_t]} \\ B_2 &= \{b \in \mathcal{B} \backslash B_1 : X(b_t, E_{\overline{k}[h_t]}) > 1 - \delta\} \\ B_3 &= \{b \in \mathcal{B} \backslash B_1 : X(b_t, E_{\overline{k}[h_t]}) \leq 1 - \delta\} \end{aligned}$$

By our arguments above, the principal's payoff $U^{(\sigma,\mu)}[h_t, b_t]$ satisfies:

$$U^{(\sigma,\mu)}[h_t, b_t] = \begin{cases} (1-\delta)(b - c_{\overline{k}[h_t]}) + \delta \mathbb{E}\big[U^{(\sigma,\mu)}[h_{t+1}, b_{t+1}]|b_t\big] & \text{if } b_t \in B_1 \\ \delta \mathbb{E}\big[U^{(\sigma,\mu)}[h_{t+1}, b_{t+1}|b_t\big] & \text{if } b_t \in B_2 \\ \max\{U^*(h_t, b_t), \delta \mathbb{E}\big[U^{(\sigma,\mu)}[h_{t+1}, b_{t+1}]|b_t\big]\} & \text{if } b_t \in B_3 \end{cases} \tag{12}$$

Let $\mathcal{F}$ be the set of functions from $\mathcal{B}$ to $\mathbb{R}$ and let $\Phi : \mathcal{F} \to \mathcal{F}$ be the operator such that, for every $f \in \mathcal{F}$,

$$\Phi(f)(b) = \begin{cases} (1-\delta)(b - c_{\overline{k}[h_t]}) + \delta \mathbb{E}[f[b_{t+1}]|b_t = b] & \text{if } b \in B_1 \\ \delta \mathbb{E}[f[b_{t+1}]|b_t = b] & \text{if } b \in B_2 \\ \max\{U^*(h_t, b), \delta \mathbb{E}[f[b_{t+1}]|b_t = b]\} & \text{if } b \in B_3 \end{cases}$$

One can check that Φ is a contraction of modulus $\delta < 1$, and therefore has a unique fixed point. Moreover, by (12), the principal's equilibrium payoffs $U^{(\sigma,\mu)}[h_t, b_t]$ are a fixed point of Φ. These two observations together imply that the principal's equilibrium payoffs $U^{(\sigma,\mu)}[h_t, b_t]$ are unique. The equilibrium strategies at (h_t, b_t) can be immediately derived from (12). Finally, it can be readily seen that these equilibrium strategies can be taken to be Markovian with respect to the principal's beliefs $\mu[h_t]$ and the shock b_t. $\square$

A.2 Proof of Proposition 3

Fix a history h_t such that $|C[h_t]| \geq 2$ and without loss of generality renumber the types so that $C[h_t] = \{c_1, ..., c_{\overline{k}[h_t]}\}$ with $c_1 < ... < c_{\overline{k}[h_t]}$. We start by showing that for every such history, there exists a shock realization $b \in \mathcal{B}$ with the property that, if $b_s = b$ at time $s \geq t$, then the principal makes an offer that a strict subset of the types in $C[h_t]$ accepts.

Suppose for the sake of contradiction that this is not true. Note that this implies that $\mu[h_{t'}] = \mu[h_t]$ for every $h_{t'} \succeq h_t$. By Theorem 1, this further implies that after history h_t, the agent only takes the action when the shock is in $E_{\overline{k}[h_t]}$, and receives a transfer equal to $c_{\overline{k}[h_t]}$. Therefore, the principal's payoff after history (h_t, b_t) is

$$U^{(\sigma,\mu)}[h_t, b_t] = (1-\delta)\mathbb{E}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}(b_{t'} - c_{\overline{k}[h_t]})\mathbf{1}_{\{b_{t'} \in E_{\overline{k}[h_t]}\}} | b_t = b\right].$$

Let $b \in E_{\overline{k}[h_t]-1}$ be such that $X(b, E_{\overline{k}[h_t]}) < 1 - \delta$. The conditions in the statement of Proposition 3 guarantee that such a shock b exists. Suppose that the shock at time $s \geq t$ is $b_s = b$, and let $\epsilon > 0$ be small enough such that

$$T = c_{\overline{k}[h_t]-1} + \frac{1}{1-\delta} X(b, E_{\overline{k}[h_t]})(c_{\overline{k}[h_t]} - c_{\overline{k}[h_t]-1}) + \epsilon < c_{\overline{k}[h_t]}. \tag{13}$$

Note that at history (h_s, b_s), an offer equal to T is accepted by all types with cost strictly lower than $c_{\overline{k}[h_t]}$, and is rejected by type $c_{\overline{k}[h_t]}$.[15] The principal's payoff from making an offer T conditional on the agent's type being $c_{\overline{k}[h_t]}$ is $U^{(\sigma,\mu)}[h_t, b_t]$. On the other hand, when the agent's type is lower than $c_{\overline{k}[h_t]}$, the principal obtains $(1-\delta)(b - T)$ at period t if she offers transfer T, and learns that the agent's type is not $c_{\overline{k}[h_t]}$. From period $t+1$ onwards, the principal's payoff is bounded below by what she could obtain if at all periods $t' > t$ she offers $T_{t'} = c_{\overline{k}[h_t]-1}$ whenever $b_{t'} \in E_{\overline{k}[h_t]-1}$ (an offer which is accepted by all types), and offers $T_{t'} = 0$ otherwise (which is rejected by all types). The payoff that the

[15]Indeed, by accepting offer T, an agent with cost $c_i < c_{\overline{k}[h_t]}$ obtains a payoff of at least $(1-\delta)(T - c_i) + \delta \times 0$. This payoff is strictly larger than the payoff of $X(b, E_{\overline{k}[h_t]})(c_{\overline{k}[h_t]} - c_i)$ he obtains by rejecting and continuing playing the equilibrium.

principal obtains from following this strategy when the agent's cost is lower than $c_{\overline{k}[h_t]}$ is

$$\begin{aligned}\underline{U} &= (1-\delta)(b-T) + (1-\delta)\mathbb{E}\left[\sum_{t'=s+1}^{\infty} \delta^{t'-s}(b_{t'} - c_{\overline{k}[h_t]-1})\mathbf{1}_{\{b_{t'}\in E_{\overline{k}[h_t]-1}\}}|b_s = b\right] \\ &= (1-\delta)(b - c_{\overline{k}[h_t]-1} - \epsilon) + (1-\delta)\mathbb{E}\left[\sum_{t'=s+1}^{\infty} \delta^{t'-s}(b_{t'} - c_{\overline{k}[h_t]})\mathbf{1}_{\{b_{t'}\in E_{\overline{k}[h_t]}\}}|b_s = b\right] \\ &\qquad + (1-\delta)\mathbb{E}\left[\sum_{t'=s+1}^{\infty} \delta^{t'-s}(b_{t'} - c_{\overline{k}[h_t]-1})\mathbf{1}_{\{b_{t'}\in E_{\overline{k}[h_t]-1}\setminus E_{\overline{k}[h_t]}\}}|b_s = b\right] \\ &= U^{(\sigma,\mu)}[h_t, b] + (1-\delta)(b - c_{\overline{k}[h_t]-1} - \epsilon) \\ &\qquad + (1-\delta)\mathbb{E}\left[\sum_{t'=s+1}^{\infty} \delta^{t'-s}(b_{t'} - c_{\overline{k}[h_t]-1})\mathbf{1}_{\{b_{t'}\in E_{\overline{k}[h_t]-1}\setminus E_{\overline{k}[h_t]}\}}|b_s = b\right],\end{aligned}$$

where the second line follows from substituting (13). Since $b \in E_{\overline{k}[h_t]-1}$, from the third line it follows that if $\epsilon > 0$ is small enough then $\underline{U}$ is strictly larger than $U^{(\sigma,\mu)}[h_t, b]$. But this cannot be, since the proposed strategy profile was an equilibrium. Therefore, for all histories h_t such that $|C[h_t]| \geq 2$, there exists $b \in \mathcal{B}$ with the property that at history (h_s, b_s) with $h_s \succeq h_t$ and $b_s = b$ the principal makes an offer that a strict subset of the types in $C[h_t]$ accept.

We now use this result to establish the proposition. Note first that this result, together with the assumption that process $\{b_t\}$ is ergodic, implies that there is *long run learning* in equilibrium. Indeed, as long as $C[h_t]$ has two or more elements, there will be some shock realization at which the principal makes an offer that only a strict subset of types in $C[h_t]$ accepts. Since there are finitely many types and $\{b_t\}$ is ergodic, with probability 1 the principal will end up learning the agent's type.

Finally, fix a history h_t such that $C[h_t] = \{c_i\}$. Then, from time t onwards the principal's payoff is $U^{(\sigma,\mu)}[h_t, b] = (1-\delta)\mathbb{E}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}(b_{t'} - c_i)\mathbf{1}_{\{b_{t'}\in E_i\}}|b_t = b\right] = U_i^*(b|c = c_i)$, which is the first best payoff. This and the previous arguments imply that the equilibrium is long run first best. □

OA2 Online Appendix

OA2.1 Proof of Lemma 0

Proof of part (i). The proof is by strong induction on the cardinality of the support of the principal's beliefs, $C[h_t]$. Fix an equilibrium (σ, μ), and note that the claim is true for all histories h_t such that $|C[h_t]| = 1$.[16] Suppose next that the claim is true for all histories $\tilde{h}_{\tilde{t}}$ with $|C[\tilde{h}_{\tilde{t}}]| \leq n-1$, and consider a history h_t with $|C[h_t]| = n$.

Suppose by contradiction that $V^{(\sigma,\mu)}_{\overline{k}[h_t]}[h_t, b_t] > 0$. Then, there must exist a state $b_{t'}$ and history $h_{t'} \succeq h_t$ that arises on the path of play with positive probability at which the principal offers a transfer $T_{t'} > c_{\overline{k}[h_t]}$ that type $c_{\overline{k}[h_t]}$ accepts. Note first that, since type $c_{\overline{k}}[h_t]$ accepts offer $T_{t'}$, all types in the support of $C[h_{t'}]$ must also accept it. Indeed, if this were not true, then there would be a highest type $c_k \in C[h_{t'}]$ that rejects the offer. By the induction hypothesis, the equilibrium payoff that this type obtains at history $h_{t'}$ is $V^{(\sigma,\mu)}_k[h_{t'}, b_{t'}] = 0$, since this type would be the highest cost of in the support of the principal's beliefs following a rejection. But this cannot be, since type c_k can get a payoff of at least $T_{t'} - c_k > 0$ by accepting the principal's offer at time t'.

We now construct an alternative strategy profile $\tilde{\sigma}$ that is otherwise identical to σ except that at history $(h_{t'}, b_{t'})$ the agent is offered a transfer $\tilde{T} \in (c_{\overline{k}[h_t]}, T_{t'})$. Specify the principal's beliefs at history $(h_{t'}, b_{t'})$ as follows: regardless of the agent's action, the principal's beliefs at the end of the period are the same as her beliefs at the beginning of the period. At all other histories, the principal's actions and beliefs are the same as in the original equilibrium. Note that, given these beliefs, at history $h_{t'}$ all agent types in $C[h_{t'}]$ find it strictly optimal to accept the principal's offer $\tilde{T}$ and take the action. Thus, the principal's payoff at history $h_{t'}$ is larger than her payoff under the original equilibrium, which cannot be since the original equilibrium was in Σ_K. □

Proof of part (ii). The proof is by induction of the cardinality of $C[h_t]$. Consider first a history h_t such that $|C[h_t]| = 2$. Without loss of generality, let $C[h_t] = \{c_1, c_2\}$, with $c_1 < c_2$. There are two cases to consider: (i) for all histories $h_{t'} \succeq h_t$, $\mu[h_{t'}] = \mu[h_t]$, i.e., there is no more learning; and (ii) there exists a history $h_{t'} \succeq h_t$ such that $\mu[h_{t'}] \neq \mu[h_t]$.

Consider first case (i). Since $\mu[h_{t'}] = \mu[h_t]$ for all $h_{t'} \succeq h_t$, both types of agents take the productive action at the same times. This implies that $A^{(\sigma,\mu)}_2[h_t, b_t] = A^{(\sigma,\mu)}_1[h_t, b_t]$.

[16] Indeed, if $C[h_t] = \{c_i\}$, then in any PBE in Σ_K the agent takes action $a = 1$ at time $t' \geq t$ if and only if $b_{t'} \in E_i$, and the principal pays the agent a transfer equal to c_i every time the agent takes the action.

Moreover, by Lemma 0(i), the transfer that the principal pays when the productive action is taken is equal to c_2. Hence, $V_1^{(\sigma,\mu)}[h_t, b_t] = (1-\delta)\mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}(T_{t'} - c_1)a_{t',1}|h_t\right] = V_2^{(\sigma,\mu)}[h_t, b_t] + A_2^{(\sigma,\mu)}[h_t, b_t](c_2 - c_1)$, where we have used $V_2^{(\sigma,\mu)}[h_t, b_t] = 0$ and $T_{t'} = c_2$ for all t' such that $a_{t',1} = a_{t',2} = 1$ (both of these follow from part (i) of the Lemma).

Consider next case (ii), and let $\underline{t} = \min\{t' \geq t : a_{t',1} \neq a_{t',2}\}$. Hence, at time $\underline{t}$ only one type of agent in $\{c_1, c_2\}$ takes the action. Note that an agent of type c_1 must take the action at time $\underline{t}$. To see why, suppose that it is only the agent of type c_2 that takes the action. By part (i) of the Lemma, the transfer $T_{\underline{t}}$ that the principal pays the agent must be equal to c_2. The payoff that an agent with type c_1 gets by accepting the offer $T_{\underline{t}}$ is bounded below by $c_2 - c_1 > 0$. In contrast, by part (i) of the Lemma, an agent of type c_1 would obtain a continuation payoff of zero by rejecting this offer. Hence, it must be that only an agent with type c_1 takes the action at time $\underline{t}$.

Note that, by part (i) of the Lemma, the total payoff that an agent with type c_1 gets from time $\underline{t}$ onwards is equal to $V_1^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] = (1-\delta)(T_{\underline{t}} - c_1)$. Note further that $(1-\delta)(T_{\underline{t}} - c_1) \geq V_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_2 - c_1)$, since an agent of type c_1 can get a payoff equal to the right-hand side by mimicking an agent with type c_2. Since we focus on PBE in Σ_K, the transfer that the principal offers the agent at time $\underline{t}$ must be $(1-\delta)(T_{\underline{t}} - c_1) = V_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_2 - c_1)$, and so

$$V_1^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] = V_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_2 - c_1) = A_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_2 - c_1), \tag{OA14}$$

where the last equality follows from part (i) of the Lemma.

Note next that, for all $t' \in \{t, ..., \underline{t} - 1\}$, $a_{t',1} = a_{t',2}$, i.e., both types of agents take the same action. Moreover, by part (i) of the Lemma, $T_{t'} = c_2$ whenever $a_{t',1} = a_{t',2} = 1$, i.e., the principal pays a transfer equal to c_2 whenever the high cost agent takes the action. Therefore,

$$\begin{aligned}
V_1^{(\sigma,\mu)}[h_t, b_t] &= \mathbb{E}^{(\mu,\sigma)}\left[\sum_{t'=t}^{\underline{t}-1} \delta^{t'-t}(1-\delta)(T_{t'} - c_1)a_{t',1} + \delta^{\underline{t}-t} V_1^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] \mid h_t, b_t\right] \\
&= \mathbb{E}^{(\mu,\sigma)}\left[\sum_{t'=t}^{\underline{t}-1} \delta^{t'-t}(1-\delta)(c_2 - c_1)a_{t',2} + \delta^{\underline{t}-t} A_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_2 - c_1) \mid h_t, b_t\right] \\
&= A_2^{(\sigma,\mu)}[h_t, b_t](c_2 - c_1) = V_2^{(\sigma,\mu)}[h_t, b_t] + A_2^{(\sigma,\mu)}[h_t, b_t](c_2 - c_1),
\end{aligned}$$

where we have used (OA14), and the fact that $V_2^{(\sigma,\mu)}[h_t, b_t] = 0$. Therefore, the lemma holds for all h_t such that $|C[h_t]| = 2$.

Suppose next that the result holds for all $\tilde{h}_{\tilde{t}}$ such that $|C[\tilde{h}_{\tilde{t}}]| \leq n-1$, and consider a history h_t such that $|C[h_t]| = n$. Consider two "adjacent" types $c_i, c_{i+1} \in C[h_t]$. We have two possible cases: (i) with probability 1, types c_i and c_{i+1} take the same action at all histories $h_{t'} \succeq h_t$; (ii) there exists a history $h_{t'} \succeq h_t$ at which types c_i and c_{i+1} take different actions. Under case (i),

$$\begin{aligned} V_i^{(\sigma,\mu)}[h_t, b_t] &= \mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)(T_{t'} - c_i)a_{t',i}|h_t, b_t\right] \\ &= \mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)(T_{t'} - c_{i+1})a_{t',i+1}|h_t, b_t\right] \\ &\quad + \mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}(1-\delta)(c_{i+1} - c_i)a_{t',i+1}|h_t, b_t\right] \\ &= V_{i+1}^{(\sigma,\mu)}[h_t, b_t] + A_{i+1}^{(\sigma,\mu)}[h_t, b_t](c_{i+1} - c_i). \end{aligned}$$

For case (ii), let $\underline{t} = \min\{t' \geq t : a_{t',i+1} \neq a_{t',i}\}$ be the first time after t at which types c_i and c_{i+1} take different actions. Let $c_k \in C[h_{\underline{t}}]$ be the highest cost type that takes the action at time $\underline{t}$. The transfer $T_{\underline{t}}$ that the principal offers at time $\underline{t}$ must satisfy $V_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] = (1-\delta)(T_{\underline{t}} - c_k) = V_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{k+1} - c_k)$.[17] Note further that $V_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] \geq (1-\delta)(T_{\underline{t}} - c_{k+1})$, since an agent with cost c_{k+1} can guarantee $(1-\delta)(T_{\underline{t}} - c_{k+1})$ by taking the action at time $\underline{t}$ and then not taking the action in all future periods. Since $(1-\delta)(T_{\underline{t}} - c_k) = V_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{k+1} - c_k)$, it follows that $A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] \leq 1-\delta$.

We now show that all types below c_k also take the action at time $\underline{t}$. That is, we show that all agents in the support of $C[h_{\underline{t}}]$ with cost weakly lower than c_k take the action at $\underline{t}$, and all agents with cost weakly greater than c_{k+1} do not take the action. Note that this implies that $c_i = c_k$ (since types c_i and c_{i+1} take different actions at time $\underline{t}$). Suppose for the sake of contradiction that this is not true, and let c_j be the highest cost type below c_k that takes does not take the action. The payoff that this agent gets from not taking the action is $V_{j\to k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] = V_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{k+1} - c_j)$, which follows since at time $\underline{t}$ types c_j and c_{k+1} do not take the action and since, by the induction hypothesis, from time $\underline{t}+1$ onwards the payoff that an agent with cost c_j gets is equal to what this agent would get by mimicking an agent with cost c_{k+1}. On the other hand, the payoff

[17]The first equality follows since, after time $\underline{t}$, type c_k is the highest type in the support of the principal's beliefs if the agent takes action $a = 1$ at time $\underline{t}$. The second equality follows since we focus on PBE in Σ_K, so the transfer $T_{\underline{t}}$ leaves a c_k-agent indifferent between accepting and rejecting.

that agent c_j obtains by taking the action and mimicking type c_k is

$$\begin{aligned}
V_{j\to k}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] &= V_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_k - c_j) \\
&= (1-\delta)(T_{\underline{t}} - c_j) + A_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_k - c_j) \\
&= V_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{k+1} - c_k) + A_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_k - c_j) \\
&> V_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{k+1} - c_j),
\end{aligned}$$

where the inequality follows since $A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] \leq 1-\delta < A_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}]$.[18] Hence, type j strictly prefers to take the action, a contradiction. Therefore, all types below c_k take the action at time $\underline{t}$, and so $c_i = c_k$.

By the arguments above, the payoff that type $c_i = c_k$ obtains at time $\underline{t}$ is

$$V_i^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] = (1-\delta)(T_{\underline{t}} - c_i) = V_{i+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{i+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{i+1} - c_i),$$

since transfer that the principal offers at time $\underline{t}$ satisfies $(1-\delta)(T_{\underline{t}} - c_i) = V_{i+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{i+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{i+1} - c_i)$. Moreover,

$$\begin{aligned}
V_i^{(\sigma,\mu)}[h_t, b_t] &= \mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\underline{t}-1} \delta^{t'-t}(1-\delta)(T_{t'} - c_i)a_{t',i} + \delta^{\underline{t}-t} V_i^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] | h_t, b_t\right] \\
&= \mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\underline{t}-1} \delta^{t'-t}\left((1-\delta)(T_{t'} - c_{i+1})a_{t',i+1} + (1-\delta)(c_{i+1} - c_i)a_{t',i+1}\right) | h_t, b_t\right] \\
&\quad + \mathbb{E}^{(\sigma,\mu)}\left[\delta^{\underline{t}-t}\left(V_{i+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{i+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{i+1} - c_i)\right) | h_t, b_t\right] \\
&= V_{i+1}^{(\sigma,\mu)}[h_t, b_t] + A_{i+1}^{(\sigma,\mu)}[h_t, b_t](c_{i+1} - c_i),
\end{aligned}$$

where the second equality follows since $a_{t',i} = a_{t',i+1}$ for all $t' \in \{t, ..., \underline{t}-1\}$. Hence, the result also holds for histories h_t with $|C[h_t]| = n$. □

OA2.2 Mixed strategies

This appendix extends the results in the main text to allow for mixed strategies. In particular, we show that the equilibrium we characterize in Theorem 1 remains the unique

[18] Recall that, for all (h_t, b_t), $A_k^{(\sigma,\mu)}[h_t, b_t] = (1-\delta)\mathbb{E}^{(\mu,\sigma)}[\sum_{t'=t}^{\infty} \delta^{t'-t} a_{t',k} | b_t, h_t]$. By assumption, an agent with type c_k takes action $a = 1$ at time $\underline{t}$, so $a_{\underline{t},k} = 1$. Moreover, it is easy to show that an agent with cost c_k will take action $a = 1$ with positive probability at some date $t > \underline{t}$. Therefore, $A_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] > 1-\delta$.

PBE that is sequentially optimal for the principal among all finitely revealing PBE; i.e., among all PBE in which, along any history, the principal updates her beliefs a finite number of periods.

Fix a PBE (σ, μ), with $\sigma = (\tau, \{\alpha_k\}_{k=1}^K)$. For any history (h_t, b_t), we say that period t is a period of revelation if (a) $\mu[h_t] \notin S_1$ (i.e., if the principal is uncertain about the agent's type) and (b) there exists $c_i, c_j \in C[h_t]$ such that $\alpha_i(h_t, b_t) \neq \alpha_j(h_t, b_t)$ (i.e., there exists at least two types in the support of the principal's beliefs that play different (possibly mixed) actions at history (h_t, b_t)). We say that an equilibrium (σ, μ) is T-revealing if, for any t and along any history h_t, the number of periods of revelation $t' < t$ is not greater than T.[19]

Three things are worth noting about T-revealing PBE. First, a T-revealing strategy does not put any bound on the occurrence of the last period of revelation. Hence, information may be revealed at any point during the game. Second, a T-revealing strategy does not require the agent to reveal her information fully. Third, since the set of possible types of the agent is finite, any pure strategy PBE is T-revealing for some T.

Let Σ_0^M denote the set of PBE that are finitely revealing (i.e., the set of PBE that are T-revealing for some finite T). For all $k = 1, ..., K$, we define the sets Σ_k^M recursively as follows:

$$\Sigma_k^M := \left\{ (\sigma,\mu) \in \Sigma_{k-1}^M : \begin{array}{c} \sigma \text{ is finitely revealing} \\ \forall (h_t, b_t) \text{ with } \mu[h_t] \in S_k \text{ and } \forall (\sigma', \mu') \in \Sigma_{k-1}^M \\ U^{(\sigma,\mu)}[h_t, b_t] \geq W^{(\sigma',\mu')}[\mu[h_t], b_t] \end{array} \right\}.$$

Let (σ^P, μ^P) denote the PBE characterized in Theorem 1, and note that $(\sigma^P, \mu^P) \in \Sigma_0^M$. The following theorem shows that (σ^P, μ^P) belongs to the set Σ_K^M. Note that this implies that any PBE in Σ_K^M gives the principal the same payoff as (σ^P, μ^P) at every history. Moreover, as the proof the theorem clarifies, any equilibrium $(\sigma, \mu) \in \Sigma_K^M$ induces the same outcome as (σ^P, μ^P).

Theorem OA1. $(\sigma^P, \mu^P) \in \Sigma_K^M$.

Proof. Fix a finitely revealing equilibrium $(\sigma, \mu) \in \Sigma_K^M$, and let T be the upper bound on the periods of revelation under (σ, μ). We start by showing that, at histories at which there have already been T periods of information revelation, players' behavior under (σ, μ) must coincide with their behavior under (σ^P, μ^P).

[19]This definition is borrowed from Peski (2008).

Consider a history (h_t, b_t) at which there have already been T periods of information revelation. Hence, $\mu[h_t] = \mu[h_{t+s}]$ for all $s \geq 0$ and all histories h_{t+s} that follow history h_t. This implies that

$$U^{(\sigma,\mu)}[h_t, b_t] \leq (1-\delta)\mathbb{E}\left[\sum_{s=0}^{\infty} \delta^s \mathbf{1}_{\{b_{t+s} \in E_{\overline{k}[h_t]}\}}(b_{t+s} - c_{\overline{k}[h_t]})|b_t\right], \qquad \text{(OA15)}$$

where $U^{(\sigma,\mu)}[h_t, b_t]$ is the principal's continuation payoff at history (h_t, b_t). To see why the inequality holds, note that all agent types in the support of $\mu[h_t]$ use the same strategy at all periods after time t. Moreover, since an agent of type $c_{\overline{k}[h_t]}$ gets a continuation payoff of 0 at all histories, she only takes the action at time $\tau \geq t$ if $T_\tau = c_{\overline{k}[h_t]}$.[20] These two observations together imply the bound in equation (OA15). Since the principal's continuation payoff at history (h_t, b_t) under equilibrium (σ^P, μ^P) is weakly larger than the right-hand side (OA15), it follows that players' behavior under (σ, μ) must coincide with their behavior under (σ^P, μ^P) at all histories after information revelation has stopped.

Next, consider a history h_t with the property that, for all histories h_{t+s} with $s \geq 1$ that follow history (h_t, b_t), players' behavior under $(\sigma, \mu) \in \Sigma_K^M$ coincides with their behavior under (σ^P, μ^P). We now show that, at such a history (h_t, b_t), the players' behavior under $(\sigma, \mu) \in \Sigma_K^M$ coincides with their behavior under (σ^P, μ^P). Before presenting its proof, we note that this result and the result above together establish Theorem OA1.

To see why the result is true, we consider two separate cases: (i) b_t such that $X(b_t, E_{\overline{k}[h_t]}) > 1-\delta$, and (ii) b_t such that $X(b_t, E_{\overline{k}[h_t]}) \leq 1-\delta$.

Case (i). Let T_t be the principal's offer at history (h_t, b_t) and note that $T_t \leq c_{\overline{k}[h_t]}$ (see footnote 20). We start by showing that if $T_t \leq c_{\overline{k}[h_t]}$ is such that an agent with type $c_{\overline{k}[h_t]}$ rejects the offer with probability 1, then all agents types also reject the offer with probability 1. Suppose by contradiction that the set of types that accept offer T_t with positive probability is non-empty. Let $c_i < c_{\overline{k}[h_t]}$ be the highest cost of a type that accepts T_t with positive probability. The payoff that type c_i obtains by accepting the offer is $(1-\delta)(T_t - c_i) + \delta \times 0 \leq (1-\delta)(c_{\overline{k}[h_t]} - c_i)$, since from $t+1$ onwards type c_i would be the highest type in the support of the principal's beliefs following an acceptance, and since equilibrium (σ, μ) coincides with (σ^P, μ^P) at all histories that follow history (h_t, b_t). In

[20] In any PBE in Σ_K^M, the principal never makes an offer T_t that is larger than the highest cost in the support of her beliefs. Indeed, if $T_t > c_{\overline{k}[h_t]}$ for some history (h_t, b_t), we can construct an alternative finitely-revealing equilibrium in Σ_{k-1}^M (where $k = |C[h_t]|$) that gives the principal strictly more profits than (σ, μ), which would contradict $(\sigma, \mu) \in \Sigma_k^M$.

contrast, the payoff that type c_i gets by rejecting the offer and mimicking type $c_{\overline{k}[h_t]}$ at all times $\tau > t$ is $X(b_t, E_{\overline{k}[h_t]})(c_{\overline{k}[h_t]} - c_i) > (1-\delta)(c_{\overline{k}[h_t]} - c_i)$, a contradiction. Hence, if $T_t \leq c_{\overline{k}[h_t]}$ is such that an agent with type $c_{\overline{k}[h_t]}$ rejects the offer with probability 1, then all agents types also reject the offer with probability 1.

There are two subcases to consider: (ia) $b_t \in E_{\overline{k}[h_t]}$, and (ib) $b_t \notin E_{\overline{k}[h_t]}$. Consider case (ia). We show that, in this case, the principal makes offer $T_t = c_{\overline{k}[h_]}$, and that this offer is accepted by all types with probability 1 (so behavior under equilibrium (σ, μ) coincides with behavior under (σ^P, μ^P)). As a first step, we show that the principal makes offer $T_t = c_{\overline{k}[h_]}$, and that this offer is accepted by an agent of type $c_{\overline{k}[h_t]}$ with positive probability. Indeed, if this was not the case, then by the arguments above no agent type would accept offer T_t, so $\mu[h_{t+1}] = \mu[h_t]$. But then we would be able to construct an alternative finitely revealing equilibrium in Σ^M_{k-1} (where $k = |C[h_t]|$) that gives the principal strictly more profits than (σ, μ), which would contradict $(\sigma, \mu) \in \Sigma^M_k$. To see how, consider an equilibrium in which players' behavior is identical to their behavior under (σ, μ) at every history except for history (h_t, b_t). At history (h_t, b_t), the principal makes offer $T_t = c_{\overline{k}[h_]}$ and every type accepts this offer with probability 1. The principal's beliefs at $t+1$ are identical to $\mu[h_t]$ regardless of whether the agent accepts or not the offer. One can check that this modified strategy profile is a PBE in finitely revealing strategies that lies in Σ^M_{k-1}. Moreover, it delivers the principal a strictly larger payoff at history (h_t, b_t) than (σ, μ), which contradicts $(\sigma, \mu) \in \Sigma^M_K$.

Next, we show that offer $T_t = c_{\overline{k}[h_]}$ is accepted with probability 1 by all agent types $c_i < c_{\overline{k}[h_t]}$. Towards a contradiction, let c_i be the highest cost type below $c_{\overline{k}[h_t]}$ that rejects the offer. The payoff that this type obtains by rejecting is at most $X(b_t, E_{\overline{k}[h_t]})(c_{\overline{k}[h_t]} - c_i)$, since either type c_i will be the second highest cost in the support of $\mu[h_{t+1}]$ (and type $c_{\overline{k}[h_t]}$ will be the highest), or type c_i will be the highest cost in the support of $\mu[h_{t+1}]$. In contrast, by accepting the offer and then mimicking type $c_{\overline{k}[h_t]}$, she obtains $(1-\delta+X(b_t, E_{\overline{k}[h_t]}))(c_{\overline{k}[h_t]} - c_i)$, which cannot be. Hence, offer $T_t = c_{\overline{k}[h_]}$ is accepted with probability 1 by all agent types $c_i < c_{\overline{k}[h_t]}$.

Finally, we show that $T_t = c_{\overline{k}[h_]}$ is accepted by an agent with cost $c_{\overline{k}[h_]}$ with probability 1. Suppose by contradiction this is not true, and consider an alternative finitely revealing equilibrium such that players' behavior coincides with their behavior under (σ, μ) at all histories except (h_t, b_t). At such a history, the principal makes offer $T_t = c_{\overline{k}[h_]}$, and this offer is accepted by all types of the agent with probability 1 (the principal's beliefs after remain equal to $\mu[h_t]$ regardless of the agent's action). One can check that this is a PBE in Σ_{k-1}, and that this PBE gives the principal a strictly larger profit than the original equilibrium (σ, μ), a contradiction. Hence, offer $T_t = c_{\overline{k}[h_]}$ is accepted by all agent types with probability 1.

Consider next case (ib). We show that, in this case, the principal makes an offer $T_t < c_{\overline{k}[h_t]}$ that all agent types reject. From our arguments above, if $T_t \leq c_{\overline{k}[h_t]}$ is rejected by an agent of type $c_{\overline{k}[h_t]}$ with probability 1, then the offer is rejected by all agent types $c_i < c_{\overline{k}[h_t]}$ with probability 1. This implies that any offer $T_t < c_{\overline{k}[h_t]}$ is rejected by every agent type with probability 1. Note that in an equilibrium $(\sigma, \mu) \in \Sigma_K^M$, at such a history the principal would never make an offer $T_t = c_{\overline{k}[h_t]}$ that is accepted by an agent of type $c_{\overline{k}[h_t]}$ with positive probability. If this were the case, and by the same arguments used in case (1a), such an offer would be accepted by all types $c_i < c_{\overline{k}[h_t]}$ with probability 1. Since $b_t < c_{\overline{k}[h_t]}$, the principal would be strictly better off by making an offer $T_t < c_{\overline{k}[h_t]}$ that is rejected by all types with probability 1.[21]

Case (ii). Consider next histories (h_t, b_t) with $X(b_t, E_{\overline{k}[h_t]}) \leq 1$. We show that, in this case, there exists a threshold $c_{k^*} \in C[h_t]$ such that types in $C^- = \{c \in C[h_t] : c < c_{k^*}\}$ accept with probability 1, and that types in $C^+ = \{c \in C[h_t] : c \geq c_{k^*}\}$ reject with probability 1. When C^- is non-empty, the principal offers transfer T_t in equation $(*)$ in the main text.

We start by showing that, at such a history (h_t, b_t), type $c_{\overline{k}[h_t]}$ takes the action with probability 0. Suppose to the contrary that type $c_{\overline{k}[h_t]}$ takes the action with positive probability, so that $T_t = c_{\overline{k}[h_t]}$. If this is so, then all types $c_i < c_{\overline{k}[h_t]}$ must take the action with probability 1. To see why, suppose this is not true, and let c_i be the highest type below $c_{\overline{k}[h_t]}$ that does not take the action with probability 1. Since equilibrium behavior under (σ, μ) coincides with equilibrium behavior under (σ^P, μ^P) at all times $\tau \geq t+1$, the payoff that type c_i obtains by rejecting the offer is at most $X(b_t, E_{\overline{k}[h_t]})(c_{\overline{k}[h_t]} - c_i)$. However, type c_i can guarantee herself a payoff of $(1 - \delta + X(b_t, E_{\overline{k}[h_t]}))(c_{\overline{k}[h_t]} - c_i)$ by accepting the offer today and then mimicking type $c_{\overline{k}[h_t]}$ at all times $\tau \geq t + 1$, a contradiction. Since $c_{\overline{k}[h_t]} < b_t$, then the principal would be strictly better off under an equilibrium in Σ_{k-1}^M that is identical to (σ, μ), except that at history (h_t, b_t) the principal makes offer $T_t = c_{\overline{k}[h_t]}$

[21] Indeed, starting from $t + 1$ equilibrium behavior under (σ, μ) coincides with equilibrium behavior under (σ^P, μ^P). As a result, the profits that the principal obtains from each type of agent $c_i < c_{\overline{k}[h_t]}$ from $t + 1$ onwards do not depend on the relative likelihood that she assigns to type $c_{\overline{k}[h_t]}$. Moreover, the profits that she extracts from type $c_{\overline{k}[h_t]}$ from $t + 1$ onwards are the same regardless of whether this type accepts or not. These two observations imply that, at time t, the principal is better off making an offer that every type of agent rejects.

which is rejected by type $c_{\overline{k}[h_t]}$ and accepted by all types $c_i < c_{\overline{k}[h_t]}$. This contradicts $(\sigma, \mu) \in \Sigma_K^M$. Hence, at history (h_t, b_t) type $c_{\overline{k}[h_t]}$ takes the action with probability 0.

Next, we show that at history (h_t, b_t), there exists a threshold $c_{k^*} \in C[h_t]$ such that types in $C^- = \{c \in C[h_t] : c < c_{k^*}\}$ accept with probability 1, and that types in $C^+ = \{c \in C[h_t] : c \geq c_{k^*}\}$ reject with probability 1. The statement is true if all types reject the offer with probability 1. Suppose the set of types in $C[h_t]$ that accept the offer with positive probability is non-empty, and let $c_{j^*} < c_{\overline{k}[h_t]}$ be the highest type in this set. Since equilibrium behavior at times $\tau \geq t+1$ coincides with (σ^P, μ^P), type c_{j^*} obtains a payoff of $(1-\delta)(T_t - c_{j^*}) + \delta \times 0$. Let c_{k^*} be the lowest type in $\{c \in C[h_t] : c > c_{j^*}\}$. Note that the offer that the principal makes must satisfy $(*)$ in the main text:

$$(1-\delta)(T_t - c_{j^*}) = V_{k^*}^{(\sigma^P,\mu^P)}[h_t, b_t] + A_{k^*}^{(\sigma^P,\mu^P)}[h_t, b_t](c_{k^*} - c_{j^*}).$$

Indeed, this transfer leaves type c_{j^*} indifferent between accepting the offer and rejecting it. Since type c_{k^*} rejects the offer with probability 1, it must be that

$$V_{k^*}^{(\sigma^P,\mu^P)}[h_t, b_t] \geq (1-\delta)(T_t - c_{k^*}) \Longleftrightarrow 1-\delta \geq A_{k^*}^{(\sigma^P,\mu^P)}[h_t, b_t]. \tag{OA16}$$

We now show that type c_{j^*} accepts with probability 1. Indeed, the payoff that the principal obtains from type c_{j^*} from $t+1$ onwards if this type accepts the offer is $(1-\delta)\mathbb{E}[\sum_{s=1}^{\infty} \delta^s \mathbf{1}_{b_{t+s} \in E_{j^*}} (b_{t+s} - c_{j^*})|b_t]$, which is the efficient payoff and is clearly higher than what she would obtain from this type if the type rejects the offer.[22]

Next, we show that all types in $c_i \in C[h_t]$ with $c_i < c_{j^*}$ accept offer T_t with probability 1. Towards a contradiction, let c_i be the highest type below c_i that rejects T_t with positive probability. Since equilibrium behavior from $t+1$ onwards under (σ, μ) coincides with equilibrium behavior under (σ^P, μ^P), type c_i obtains payoff $V_{k^*}^{(\sigma^P,\mu^P)}[h_t, b_t] + A_{k^*}^{(\sigma^P,\mu^P)}[h_t, b_t](c_{k^*} - c_i)$ from rejecting offer T_t. In contrast, the payoff that type c_i would obtain from accepting offer T_t and mimicking type c_{j^*} from time $t+1$ onwards is $(1-\delta)(T_t - c_i) + X(b_t, E_{j^*})(c_{j^*} - c_i)$. Note that

$$\begin{aligned}
&(1-\delta)(T_t - c_i) + X(b_t, E_{j^*})(c_{j^*} - c_i) - V_{k^*}^{(\sigma^P,\mu^P)}[h_t, b_t] - A_{k^*}^{(\sigma^P,\mu^P)}[h_t, b_t](c_{k^*} - c_i) \\
=&(c_{j^*} - c_i)\left(1 - \delta + X(b_t, E_{j^*}) - A_{k^*}^{(\sigma^P,\mu^P)}[h_t, b_t]\right) > 0,
\end{aligned}$$

[22]Moreover, if some types $c_i < c_{j^*}$ were to reject the offer, the continuation payoff that the principal would get from them would be weakly higher if type c_{j^*} were to accept offer T_t with probability 1, than if type c_{j^*} were to reject the offer with positive probability. Indeed, if type c_{j^*} is not in the support of the principal's beliefs at time $t+1$, then types $c_i < c_{j^*}$ get smaller informational rents.

where we used equation (OA16).

The arguments above show that, at histories (h_t, b_t) with $X(b_t, E_{\overline{k}[h_t]}) \leq 1$, there exists a threshold $c_{k^*} \in C[h_t]$ such that types in $C^- = \{c \in C[h_t] : c < c_{k^*}\}$ accept with probability 1, and that types in $C^+ = \{c \in C[h_t] : c \geq c_{k^*}\}$ reject with probability 1. Since the threshold c_{k^*} is chosen optimally under equilibrium (σ^P, μ^P), under equilibrium (σ, μ) the principal would choose the same cutoff. Hence, at history (h_t, b_t), players' behavior under $(\sigma, \mu) \in \Sigma_K^M$ coincides with their behavior under (σ^P, μ^P). □

OA2.3 Full Commitment

This appendix studies the problem of a principal who has full commitment power. For conciseness, we focus on the case in which there are two types of agents: $\mathcal{C} = \{c_1, c_2\}$, with $c_1 < c_2$. Let $\mu \in (0, 1)$ be the probability that the agent's cost is c_2.

The principal's problem is to choose processes $\{a_{i,t}, T_{i,t}\}$ for $i = 1, 2$, with $a_{i,t} \in \{0, 1\}$ and $T_{i,t} \in \mathbb{R}$, to solve

$$U^{FC}(b) = \max_{\{a_{i,t}\, T_{i,t}\}_{i=1,2}} (1-\delta)\mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \left((1-\mu)(a_{1,t}b_t - T_{1,t}) + \mu(a_{2,t}b_t - T_{2,t})\right) | b_0 = b\right] \tag{OA17}$$

$$\text{subject to} \quad \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t (T_{i,t} - a_{i,t}c_i) | b_0 = b\right] \geq 0 \text{ for } i = 1, 2$$

$$\text{and} \quad \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t (T_{i,t} - a_{i,t}c_i) | b_0 = b\right] \geq \mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t (T_{j,t} - a_{j,t}c_i) | b_0 = b\right] \text{ for } i, j = 1, 2.$$

By familiar arguments, the participation constraint of type c_1 and the incentive compatibility constraint of type c_2 do not bind. The participation constraint of type c_2 and the incentive compatibility constraint of type c_1 hold with equality at the solution to (OA17). Using these two constraints to solve for the expected discounted transfers $(1-\delta)\mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t T_{i,t} | b_0 = b\right]$ for $i = 1, 2$ and replacing them into the objective yields

$$U^{FC}(b) = \max_{\{a_{i,t}\}_{i=1,2}} (1-\delta)\mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \left((1-\mu)a_{1,t}(b_t - c_1) + \mu a_{2,t}\left(b_t - c_2 - \frac{(1-\mu)}{\mu}(c_2 - c_1)\right)\right) | b_0 = b\right]. \tag{OA18}$$

The solution to problem (OA18) has: $a_{1,t} = 1$ if and only if $b_t \geq c_1$ (i.e., iff $b_t \in E_1$), and $a_{2,t} = 1$ if and only if $b_t \geq c_2 + \frac{(1-\mu)}{\mu}(c_2 - c_1) > c_2$.

The following result shows that, in the presence of stochastic shocks, the principal's equilibrium payoffs can be close to her full commitment payoffs

Proposition OA1. *Let $\mathcal{C} = \{c_1, c_2\}$, and assume there exists $b \in E_2 \backslash E_1$ with $X(b, E_2) = \epsilon < 1 - \delta$. Then, at histories (h_t, b_t) with $C[h_t] = \{c_1, c_2\}$ and $b_t = b$,*

$$U^{FC}(b_t) - U^{\sigma,\mu}[h_t, b_t] \leq (1 - \mu)(c_2 - c_1)\epsilon.$$

Proof. Note that, at such a history, the principal can make a separating offer T with $(1 - \delta)(T - c_1) = X(b, E_1)(c_2 - c_1)$ that only low types accept. Conditional on the agent being a low type, the principal's profits are

$$(1 - \delta)\mathbb{E}\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} \mathbf{1}_{b_\tau \in E_1}(b_\tau - c_1)|b_t = b\right] - X(b, E_1)(c_2 - c_1).$$

Conditional on the agent's type being a high type, the principal's profits are

$$(1 - \delta)\mathbb{E}\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} \mathbf{1}_{b_\tau \in E_2}(b_\tau - c_2)|b_t = b\right].$$

The principal's expected payoff at history (h_t, b_t) from making offer T is then

$$U^{\sigma,\mu}[h_t, b_t] = (1-\delta)\mathbb{E}\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t}\left((1 - \mu)\mathbf{1}_{b_\tau \in E_1}(b_\tau - c_1) + \mu\mathbf{1}_{b_\tau \in E_2}\left(b_t - c_2 - \frac{(1 - \mu)}{\mu}(c_2 - c_1)\right)\right)|b_t\right] \tag{OA19}$$

The principal's full commitment payoffs are

$$U^{FC}(b_t) = (1-\delta)\mathbb{E}\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t}\left((1 - \mu)\mathbf{1}_{b_\tau \in E_1}(b_\tau - c_1) + \mu\mathbf{1}_{b_\tau \in \hat{E}_2}\left(b_t - c_2 - \frac{(1 - \mu)}{\mu}(c_2 - c_1)\right)\right)|b_t\right], \tag{OA20}$$

where $\hat{E}_2 = \{b \in \mathcal{B} : b_t \geq c_2 + (1-\mu)(c_2 - c_1)/\mu\} \subset E_2$. Using (OA19) and (OA20),

$$
\begin{aligned}
U^{FC}(b_t) - U^{\sigma,\mu}[h_t, b_t] &= -(1-\delta)\mu\mathbb{E}\left[\sum_{\tau=t}^{\infty}\delta^{\tau-t}\mathbf{1}_{b_\tau \in E_2\backslash\hat{E}_2}\left(b_t - c_2 - \frac{(1-\mu)}{\mu}(c_2 - c_1)\right)|b_t\right] \\
&\leq (1-\delta)\mu\mathbb{E}\left[\sum_{\tau=t}^{\infty}\delta^{\tau-t}\mathbf{1}_{b_\tau \in E_2\backslash\hat{E}_2}\left(\frac{(1-\mu)}{\mu}(c_2 - c_1)\right)|b_t\right] \\
&= (1-\mu)(c_2 - c_1)X(b_t, E_2\backslash\hat{E}_2) \\
&\leq (1-\mu)(c_2 - c_1)\epsilon,
\end{aligned}
$$

where the first inequality follows since $b_\tau \geq c_2$ for all $b_\tau \in E_2$, and the second inequality follows since $X(b_t, E_2\backslash\hat{E}_2) \leq X(b_t, E_2) = \epsilon$. □

OA2.4 Path Dependence when Shocks are Ergodic

In this appendix, we show by example that the equilibrium may exhibit long-run path dependence when the shock process is ergodic. Let $\mathcal{B} = \{b_L, b_{ML}, b_{MH}, b_H\}$, with $b_L < b_{ML} < b_{MH} < b_H$ and $\mathcal{C} = \{c_1, c_2, c_3\}$. Assume that the efficiency sets are $E_1 = E_2 = \{b_{ML}, b_{MH}, b_H\}$ and $E_3 = \{b_H\}$.

Proposition OA2. *Suppose that the transition matrix $[Q_{b,b'}]$ satisfies:*

(a) $Q_{b,b'} > 0$ for all $b, b' \in \mathcal{B}$;

(b) $X(b_{MH}, \{b_H\}) > 1-\delta$, $X(b_{ML}, \{b_H\}) < 1-\delta$ and $X(b_{ML}, \{b_{ML}\}) > 1-\delta$

Then, there exists $\epsilon_1 > 0, \epsilon_2 > 0, \Delta_1 > 0$ and $\Delta_2 > 0$ such that, if $Q_{b,b_L} < \epsilon_1$ for all $b \in \mathcal{B}\backslash\{b_L\}$ and $Q_{b,b_{ML}} < \epsilon_2$ for all $b \in \mathcal{B}\backslash\{b_{ML}\}$, and if $|b_L - c_1| < \Delta_1$ and $|b_L - c_2| > \Delta_2$, the unique equilibrium satisfies:

(i) For histories h_t such that $C[h_t] = \{c_1, c_2\}$, $\mu[h_{t'}] = \mu[h_t]$ for all $h_{t'} \succeq h_t$ (i.e., there is no more learning by the principal from time t onwards);

(ii) For histories h_t such that $C[h_t] = \{c_2, c_3\}$: if $b_t = b_L$ or $b_t = b_{MH}$, types c_2 and c_3 take action $a = 0$; if $b_t = b_{ML}$, type c_2 takes action $a = 1$ and type c_3 takes action $a = 0$; and if $b_t = b_H$, types c_2 and c_3 take action $a = 1$;

(iii) For histories h_t such that $C[h_t] = \{c_1, c_2, c_3\}$: if $b_t = b_L$, type c_1 takes action $a = 1$ while types c_2 and c_3 take action $a = 0$; if $b_t = b_{ML}$, types c_1 and c_2 take action

$a = 1$ and type c_3 takes action $a = 0$; if $b_t = b_{MH}$, all agent types take action $a = 0$; and if $b_t = b_H$, all agent types take action $a = 1$.

We prove the three properties in Proposition OA2 separately.

Proof of Property (i). Note first that, by Theorem 1, after such a history the principal makes a pooling offer $T = c_2$ that both types accept if $b_t \in E_2 = \{b_{ML}, b_{MH}, b_H\}$. To establish the result, we show that if $b_t = b_L$, types c_1 and c_2 take action $a = 0$ after history h_t. If the principal makes a separating offer that only a c_1 agent accepts, she pays a transfer $T_t = c_1 + \frac{1}{1-\delta}X(b_L, E_2)(c_2 - c_1)$ that compensates the low cost agent for revealing his type. The principal's payoff from making such an offer, conditional on the agent being type c_1, is

$$\begin{aligned}\tilde{U}^{sc}[c_1] &= (1-\delta)(b_L - T_t) + \mathbb{E}\left[\sum_{t'>t} \delta^{t'-t}(1-\delta)\mathbf{1}_{b_t \in E_1}(b_{t'} - c_1)|b_t = b_L\right] \\ &= (1-\delta)(b_L - c_1) + \sum_{b\in\{b_{ML},b_{MH},b_H\}} X(b_L, \{b\})[b - c_2].\end{aligned}$$

Her payoff from making that offer conditional on the agent's type being c_2 is $\tilde{U}^{sc}[c_2] = \sum_{b\in\{b_{ML},b_{MH},b_H\}} X(b_L, \{b\})[b - c_2]$. If she doesn't make a separating offer when $b_t = b_L$, she never learns the agent's type and gets a payoff $\tilde{U}^{nsc} = \sum_{b\in\{b_{ML},b_{MH},b_H\}} X(b_L, \{b\})[b - c_2]$. Since $b_L - c_1 < 0$ by assumption, $\tilde{U}^{nsc} > \mu[h_t][c_1]\tilde{U}^{sc}[c_1] + \mu[h_t][c_2]\tilde{U}^{sc}[c_2]$, and therefore the principal does not to make a separating offer. □

Proof of Property (ii). Theorem 1 implies that, after such a history, the principal makes a pooling offer $T = c_3$ that both types accept if $b_t \in E_3 = \{b_H\}$. Theorem 1 also implies that, if $b_t = b_{MH}$, then after such a history the principal makes an offer that both types reject (since $X(b_{MH}, \{b_H\}) > 1 - \delta$ by assumption). So it remains to show that, after history h_t, the principal makes an offer that a c_2 agent accepts and a c_3 agent rejects if $b_t = b_{ML}$, and that the principal makes an offer that both types reject if $b_t = b_L$.

Suppose $b_t = b_{ML}$. Let $U[c_i]$ be the principal's value at history $(h_t, b_t = b_{ML})$ conditional on the agent's type being $c_i \in \{c_2, c_3\}$, and let V_i be the value of an agent of type c_i at history $(h_t, b_t = b_{ML})$. Note that $U[c_2] + V_2 \leq (1-\delta)(b_{ML} - c_2) + \sum_{b\in\{b_{ML},b_{MH},b_H\}} X(b_{ML}, \{b\})[b - c_2]$, since the right-hand side of this equation corresponds to the efficient total payoff when the agent is of type c_2 (i.e., the agent taking the action if and only if the state is in E_2.) Note also that incentive compatibility implies $V_2 \geq X(b_{ML}, \{b_H\})(c_2 - c_3)$, since a c_2-agent can mimic a c_3-agent forever and obtain

$X(b_{ML},\{b_H\})(c_2-c_3)$. It thus follows that $U[c_2] \leq (1-\delta)(b_{ML}-c_2)+X(b_{ML},\{b_H\})[b_H - c_3] + \sum_{s\in\{b_{ML},b_{MH}\}} X(b_{ML},\{b\})[b-c_2]$.

If when $b_t = b_{ML}$ the principal makes an offer that only a c_2 agent accepts, the offer must satisfy $T_t = c_2 + \frac{1}{1-\delta}X(b_{ML},\{b_H\})(c_3-c_2) < c_3$. The principal's payoff from making such an offer when the agent's type is c_2 is

$$\begin{aligned}\overline{U}[c_2] &= (1-\delta)(b_{ML}-T_t) + \sum_{b\in\{b_{ML},b_{MH},b_H\}} X(b_{ML},\{b\})[b-c_2] \\ &= (1-\delta)(b_{ML}-c_2) + X(b_{ML},\{b_H\})[b_H-c_3] + \sum_{b\in\{b_{ML},b_{MH}\}} X(b_{ML},\{b\})[b-c_2],\end{aligned}$$

which, from the arguments in the previous paragraph, is the highest payoff that the principal can ever get from a c_2 agent after history $(h_t, b_t = b_{ML})$. Hence, it is optimal for the principal to make such a separating offer.[23]

Suppose next that $b_t = b_L$. If the principal makes an offer that a c_2-agent accepts and a c_3-agent rejects, she pays a transfer $T_t = c_2 + \frac{1}{1-\delta}X(b_L, E_3)(c_3-c_2)$. Thus, the principal's payoff from making such an offer, conditional on the agent being type c_2, is

$$\begin{aligned}\tilde{U}^{sc}[c_2] &= (1-\delta)(b_L-T_t) + \sum_{b\in\{b_{ML},b_{MH},b_H\}} X(b_L,\{b\})[b-c_2] \\ &= (1-\delta)(b_L-c_2) + X(b_L,\{b_H\})[b_H-c_3] + \sum_{b\in\{b_{ML},b_{MH}\}} X(b_L,\{b\})[b-c_2].\end{aligned}$$

If the principal makes an offer that both types reject when $b_t = b_L$, then by the arguments above she learns the agent's type the first time at which shock b_{ML} is reached. Let $\check{t}$ be the random variable that indicates the next date at which shock b_{ML} is realized. Then, conditional on the agent's type being c_2, the principal's payoff from making an offer that both types reject when $b_t = b_L$ is

$$\begin{aligned}\tilde{U}^{nsc}[c_2] &= (1-\delta)\mathbb{E}\left[\sum_{t'=t+1}^{\check{t}-1} \delta^{t'-t}\mathbf{1}_{b_{t'}=b_H}(b_H-c_3)|b_t=b_L\right] \\ &\quad + \mathbb{E}\left[\delta^{\check{t}-t}\left((1-\delta)(b_{ML}-T_{\check{t}}) + \sum_{b\in\{b_{ML},b_{MH},b_H\}} X(b_{ML},\{b\})[b-c_2]\right)|b_t=b_L\right].\end{aligned}$$

[23]Indeed, the principal's payoff from making an offer equal to T_t when the agent's type is c_3 is $X(2,\{4\})[b(4)-c_3]$, which is also the most that she can extract from an agent of type c_3.

The offer $T_{\check{t}}$ that the principal makes at time $\check{t}$ satisfies $T_{\check{t}} = c_2 + \frac{1}{1-\delta} X(b_{ML}, \{b_H\})(c_3 - c_2)$. Using this in the equation above,

$$\tilde{U}^{nsc}[c_2] = X(b_L, \{b_H\})[b_H - c_3] + X(b_L, \{b_{ML}\})[b_{ML} - c_2] + \mathbb{E}\left[\delta^{\check{t}-t} | b_t = b_L\right] X(b_{ML}, \{b_{MH}\})[b_{MH} - c_2].$$

Then, we have

$$\tilde{U}^{nsc}[c_2] - \tilde{U}^{sc}[c_2] = -(1-\delta)[b_L - c_2] - \left[X(b_L, \{b_{MH}\}) - \mathbb{E}\left[\delta^{\check{t}-t} | b_t = b_L\right] X(b_{ML}, \{b_{MH}\})\right][b_{MH} - c_2].$$

Since $b_L < c_2$ by assumption, there exists $\Delta_2^1 > 0$ such that, if $(1-\delta)(c_2 - b_L) > \Delta_2^1$, the expression above is positive. Since the principal's payoff conditional on the agent's type being c_3 is the same regardless of whether she makes a separating offer or not when $b_t = b_L$ (i.e., in either case the principal earns $X(b_L, \{b_H\})(b_H - c3)$), when this condition holds the principal chooses not to make an offer that c_2 accepts and c_3 rejects when $b_t = b_L$. □

Proof of Property (iii). Suppose $C[h_t] = \{c_1, c_2, c_3\}$. Theorem 1 implies that all agent types take action $a = 1$ if $b_t = b_H$, and all agent types take action $a = 0$ if $b_t = b_{MH}$ (this last claim follows since $X(b_{MH}, \{b_H\}) > 1 - \delta$).

Suppose next that $C[h_t] = \{c_1, c_2, c_3\}$ and $b_t = b_{ML}$. Note that, by Lemma 2, an agent with type c_3 takes action $a = 0$ if $b_t = b_{ML} \notin E_3 = \{b_H\}$. We first claim that if the principal makes an offer that only a subset of types accept at state b_{ML}, then this offer must be such that types in $\{c_1, c_2\}$ take action $a = 1$ and type c_3 takes action $a = 0$. To see this, suppose that she instead makes an offer that only an agent with type c_1 accepts, and that agents with types in $\{c_2, c_3\}$ reject. The offer that she makes in this case satisfies $(1-\delta)(T_t - c_1) = V_2^{(\sigma,\mu)}[h_t, b_t] + A_2^{(\sigma,\mu)}[h_t, b_t](c_2 - c_1)$. By property (ii) above, under this proposed equilibrium a c_2-agent will from period $t+1$ onwards take the action at all times $t' > t$ such that $b_{t'} = b_{ML}$.[24] Therefore, $A_2^{(\sigma,\mu)}[h_t, b_t] \geq X(b_{ML}, \{b_{ML}\}) > 1 - \delta$, where the last inequality follows by assumption. The payoff that an agent of type c_2 obtains by accepting offer T_t at time t is bounded below by $(1-\delta)(T_t - c_2) = (1-\delta)(c_1 - c_2) + V_2^{(\sigma,\mu)}[h_t, b_t] + A_2^{(\sigma,\mu)}[h_t, b_t](c_2 - c_1) > V_2^{(\sigma,\mu)}[h_t, b_t]$, where the inequality follows since $A_2^{(\sigma,\mu)}[h_t, b_t] > 1 - \delta$. Thus, type c_2 strictly prefers to accept the offer, a contradiction. Therefore, when $C[h_t] = \{c_1, c_2, c_3\}$ and $b_t = b_{ML}$, either

[24] Under the proposed equilibrium, if the offer is rejected the principal learns that the agent's type is in $\{c_2, c_3\}$. By property (ii), if the agent's type is c_2, the principal will learn the agent's type the next time the shock is b_{ML} (because at that time type c_2 takes the action, while type c_3 doesn't), and from that point onwards the agent will take the action when the shock is in $E_2 = \{b_{ML}, b_{MH}, b_H\}$.

the principal makes an offer that only types in $\{c_1, c_2\}$ accept, or she makes an offer that all types reject.

We now show that, under the conditions in the Lemma, the principal makes an offer that types in $\{c_1, c_2\}$ accept and type c_3 rejects when $b_t = b_{ML}$ and $C[h_t] = \{c_1, c_2, c_3\}$. If she makes an offer that agents with cost in $\{c_1, c_2\}$ accept and a c_3-agent rejects, then she pays a transfer $T_t = c_2 + \frac{1}{1-\delta} X(b_{ML}, \{b_H\})(c_3 - c_2)$. Note then that, by property (i) above, when the agent's cost is in $\{c_1, c_2\}$, the principal stops learning: for all times $t' > t$ the principal makes an offer $T_{t'} = c_2$ that both types accept when $b_{t'} \in E_2$, and she makes a low offer $T_{t'} = 0$ that both types reject when $b_{t'} \notin E_2$. Therefore, conditional on the agent's type being either c_1 or c_2, the principal's payoff from making at time t an offer T_t that agents with cost in $\{c_1, c_2\}$ accept and a c_3-agent rejects is

$$\begin{aligned}\hat{U}^{sc}[\{c_1, c_2\}] &= (1-\delta)(b_{ML} - T_t) + \sum_{b \in \{b_{ML}, b_{MH}, b_H\}} X(b_{ML}, \{b\})[b - c_2] \\ &= (1-\delta)(b_{ML} - c_2) + X(b_{ML}, \{b_H\})[b_H - c_3] + \sum_{b \in \{b_{ML}, b_{MH}\}} X(b_{ML}, \{b\})[b - c_2]\end{aligned}$$

On the other hand, if she does not make an offer that a subset of types accept when $b_t = b_{ML}$, then the principal's payoffs conditional on the agent being of type $c_i \in \{c_1, c_2\}$ is bounded above by

$$\hat{U}^{nsc}[c_i] = \mathbb{E}\left[\sum_{t'=t}^{\hat{t}-1} \delta^{t'-t}(1-\delta)\mathbf{1}_{b_{t'}=b_H}(b_H - c_3) + \delta^{\hat{t}-t} \sum_{b \in E_i} X(b_L, \{b\})(b - c_i) | b_t = b_{ML}\right]$$

where $\hat{t}$ denotes the next period that state b_L is realized.[25] Note that there exists $\epsilon_1 > 0$ small such that, if $Q_{b,b_L} < \epsilon_1$ for all $b \neq b_L$, then $\hat{U}^{sc}[\{c_1, c_2\}] > \hat{U}^{nsc}[c_i]$ for $i = 1, 2$. Finally, note that the payoff that the principal obtains from an agent of type c_3 at history h_t when $b_t = b_{ML}$ is $X(b_{ML}, \{b_H\})(b_H - c_3)$, regardless of the principal's offer. Therefore, if $Q_{b,b_L} < \epsilon_1$ for all $b \neq b_L$, when $C[h_t] = \{c_1, c_2, c_3\}$ and $b_t = b_{ML}$ the principal makes an offer T_t that only types in $\{c_1, c_2\}$ accept.

Finally, we show that when $C[h_t] = \{c_1, c_2, c_3\}$ and $b_t = b_L$, the principal makes an offer that only type c_1 accepts. Let $\check{t}$ be the random variable that indicates the next date

[25] To see why, note that if no type of agent takes the productive action when $C[h_t] = \{c_1, c_2, c_3\}$ and $b_t = b_{ML}$, then the principal can only learn the agent's type when state b_L is realized (i.e., at time $\hat{t}$). At times before $\hat{t}$, all agent types take the action if the shock is b_H (and the principal pays transfer $T = c_3$), and no agent type takes the action at states b_{ML} or b_{MH}. After time $\hat{t}$, the payoff that the principal gets from type c_i is bounded above by her first-best payoff $\sum_{b \in E_i} X(b_L, \{b\})(b - c_i)$.

at which state b_{ML} is realized. If the principal makes an offer T_t that only a c_1-agent accepts, this offer satisfies

$$\begin{aligned}(1-\delta)(T_t-c_1) &= V_2^{(\sigma,\mu)}[h_t,b_L]+A_2^{(\sigma,\mu)}[h_t,b_L](c_2-c_1)\\ &= X(b_L,\{b_H\})(c_3-c_1)+\left[X(b_L,\{b_{ML}\})+\mathbb{E}[\delta^{\check{t}-t}|b_t=b_L]X(b_{ML},\{b_{MH}\})\right](c_2-c_1)\end{aligned} \tag{OA21}$$

where the second equality follows since $V_2^{(\sigma,\mu)}[h_t,b_L]=A_3^{(\sigma,\mu)}[h_t,b_L](c_3-c_2)=X(b_L,\{b_H\})(c_3-c_2)$ and since, by property (ii), when the support of the principal's beliefs is $\{c_2,c_3\}$ and the agent's type is c_2, the principal learns the agent's type at time $\check{t}$.[26] Therefore, conditional on the agent's type being c_1, the principal's equilibrium payoff from making an offer that only an agent with cost c_1 accepts at state b_L is

$$\begin{aligned}\check{U}^{sc}[c_1] &= (1-\delta)(b_L-T_t)+\sum_{b\in\{b_{ML},b_{MH},b_H\}}X(b_L,\{b\})[b-c_1]\\ &= (1-\delta)(b_L-c_1)+X(b_L,\{b_H\})[b_H-c_3]+X(b_L,\{b_{MH}\})[b_{MH}-c_1]\\ &\quad +X(b_L,\{b_{ML}\})[b_{ML}-c_2]-\mathbb{E}[\delta^{\check{t}-t}|b_t=b_L]X(b_{ML},\{b_{MH}\})(c_2-c_1)\end{aligned}$$

where the second line follows from substituting the transfer in (OA21). On the other hand, the principal's payoff from making such an offer at state b_L, conditional on the agent's type being c_2, is

$$\begin{aligned}\check{U}^{sc}[c_2] &= (1-\delta)\mathbb{E}\left[\sum_{t'=t}^{\check{t}-1}\delta^{t'-t}\mathbf{1}_{b_{t'}=b_H}(b_H-c_3)|b_t=b_L\right]\\ &\quad +(1-\delta)\mathbb{E}\left[\delta^{\check{t}-t}\left((b_{ML}-c_2)-\frac{X(b_{ML},\{b_H\})(c_3-c_2)}{1-\delta}\right)+\sum_{t'=\check{t}+1}^{\infty}\delta^{t'-t}\mathbf{1}_{b_{t'}\in E_2}(b_{t'}-c_2)|b_t=b_L\right]\\ &= X(b_L,\{b_H\})(b_H-c_3)+X(b_L,\{b_{ML}\})(b_{ML}-c_2)+\mathbb{E}\left[\delta^{\check{t}-t}X(b_{ML},\{b_{MH}\})|b_t=b_L\right](b_{MH}-c_2),\end{aligned}$$

[26] The fact that the principal learns the agent's type at time $\check{t}$ implies that

$$\begin{aligned}A_2^{(\sigma,\mu)}[h_t,b_L] &= (1-\delta)\mathbb{E}\left[\sum_{t'=t}^{\check{t}-1}\delta^{t'-t}\mathbf{1}_{b_{t'}=b_H}+\delta^{\check{t}-t}\sum_{t'=\check{t}}^{\infty}\delta^{t'-\check{t}}\mathbf{1}_{b_{t'}\in E_2}|b_t=b_L\right]\\ &= X(b_L,\{b_H\})+X(b_L,\{b_{ML}\})+\mathbb{E}\left[\delta^{\check{t}-t}X(b_{ML},\{b_{MH}\})|b_t=b_L\right].\end{aligned}$$

where we used the fact that, when the support of her beliefs is $\{c_2, c_3\}$, the principal makes an offer that only a c_2-agent accepts when the state is b_{ML} (the offer that she makes at that point is $T = c_2 + \frac{1}{1-\delta}X(b_{ML}, \{b_H\})(c_3 - c_2)$).

Alternatively, suppose the principal makes an offer that both c_1 and c_2 accept but c_3 rejects. Then she pays a transfer $T_t = c_2 + \frac{1}{1-\delta}X(b_L, \{b_H\})(c_3 - c_2)$; thus, her payoff from learning that the agent's type is in $\{c_1, c_2\}$ in state b_L is

$$\begin{aligned}\bar{U}^{sc}[\{c_1, c_2\}] &= (1-\delta)(b_L - T_t) + \sum_{b \in \{b_{ML}, b_{MH}, b_H\}} X(b_L, \{b\})(b - c_2) \\ &= (1-\delta)(b_L - c_2) + X(b_L, \{b_H\})[b_H - c_3] \\ &\quad + X(b_L, \{b_{ML}\})[b_{ML} - c_2] + X(b_L, \{b_{MH}\})[b_{MH} - c_2],\end{aligned}$$

where we used the fact that the principal never learns anything more about the agent's type when the support of her beliefs is $\{c_1, c_2\}$ (see property (i) above). Note that there exists $\epsilon_2 > 0$ and $\Delta_2^2 > 0$ such that, if $Q_{b, b_{ML}} < \epsilon_2$ for all $b \neq b_{ML}$ and if $c_2 - b_L > \Delta_2 = \max\{\Delta_2^1, \Delta_2^2\}$, then the following two inequalities hold:

$$\begin{aligned}\check{U}^{sc}[c_1] - \bar{U}^{sc}[\{c_1, c_2\}] &= \left[1 - \delta + X(b_L, \{b_{MH}\}) - \mathbb{E}[\delta^{\check{t}-t} | b_t = b_L] X(b_{ML}, \{b_{MH}\})\right] (c_2 - c_1) > 0 \\ \check{U}^{sc}[c_2] - \bar{U}^{sc}[\{c_1, c_2\}] &= \left[E\left[\delta^{\check{t}-t} X(b_{ML}, \{b_{MH}\}) | b_t = b_L\right] - X(b_L, \{b_{MH}\})\right] (b_{MH} - c_2) \\ &\quad - (1-\delta)(b_L - c_2) > 0.\end{aligned}$$

Therefore, under these conditions, at state b_L the principal strictly prefers to make an offer that a c_1-agent accepts and agents with cost $c \in \{c_2, c_3\}$ reject than to make an offer that agents with cost in $\{c_1, c_2\}$ accept and a c_3-agent rejects.

However, the principal may choose to make an offer that all agent types reject when $b_t = b_L$ and $C[h_t] = \{c_1, c_2, c_3\}$. In this case, by the arguments above, the next time the state is equal to b_{ML} the principal will make an offer that only types in $\{c_1, c_2\}$ accept. The offer that she makes in this case is such that $(1-\delta)(T - c_2) = X(b_{ML}, \{b_H\})(c_3 - c_2)$. Then, from that point onwards, she will never learn more (by property (i) above). In this case, the principal's payoff conditional on the agent's type being $\{c_1, c_2\}$ is

$$\begin{aligned}\bar{U}^{nsc} =&(1-\delta)\mathbb{E}\left[\sum_{\tau=t}^{\check{t}-1} \mathbf{1}_{b_\tau=b_H}(b_\tau - c_3)|b_t = b_L\right] \\ &+ \mathbb{E}\left[\delta^{\check{t}-t}(1-\delta)(b_{ML}-T) + \sum_{b\in E_2} X(b_{ML},\{b\})(b-c_2)|b_t = b_L\right] \\ &= X(b_L,\{b_H\})[b_H - c_3] + X(b_L,\{b_{ML}\})[b_{ML}-c_2] + \mathbb{E}[\delta^{\check{t}-t}|b_t = b_L]X(b_{ML},\{b_{MH}\})[b_{MH}-c_2],\end{aligned}$$

where $\check{t}$ be the random variable that indicates the next date at which state b_{ML} is realized. Note that there exists $\epsilon_2 > 0$ and $\Delta_1 > 0$ such that, if $Q_{b,b_{ML}} < \epsilon_2$ for all $b \neq b_{ML}$, and if $b_L - c_1 > -\Delta_1$, then the following hold:

$$\begin{aligned}&\check{U}^{sc}[c_1] - \bar{U}^{nsc} = (1-\delta)(b_L - c_1) + \left[X(b_L,\{b_{MH}\}) - \mathbb{E}[\delta^{\check{t}-t}|b_t = b_L]X(b_{ML},\{b_{MH}\})\right][b_{MH}-c_1] > 0 \\ &\check{U}^{sc}[c_2] - \bar{U}^{nsc} = 0.\end{aligned}$$

Therefore, under these conditions, the principal makes an offer that type c_1 accepts and types in $\{c_2, c_3\}$ reject when $C[h_t] = \{c_1, c_2, c_3\}$ and $b_t = b_L$. $\square$

Properties (i)-(iii) in Proposition OA2 imply that the equilibrium exhibits long-run path dependence. Suppose that the agent's type is c_1. Then, properties (i)-(iii) imply that the principal eventually learns the agent's type if and only if $t(b_L) := \min\{t \geq 0 : b_t = b_L\} < t(b_{ML}) := \min\{t \geq 0 : b_t = b_{ML}\}$ (i.e., if state b_L is visited before state b_{ML}). Indeed, if b_L is visited before b_{ML}, at time $t(b_L)$ the principal will learn that the agent's type is c_1 (see property (iii)). From that point onwards, the agent will take the productive action at all periods $t > t(b_L)$ such that $b_t \in E_1$ at cost c_1 for the principal.

In contrast, if b_{ML} is visited before b_L, at time $t(b_{ML})$ the principal will learn that the agent's type is in $\{c_1, c_2\}$ (see property (iii)). From that point onwards there will be no more learning (property (i)). As a consequence, the agent will take the productive action at all periods $t > t(b_{ML})$ such that $b_t \in E_2 = E_1$ at cost c_2 for the principal (this follows from Theorem 1(i)).

References

Blume, A. (1998): "Contract Renegotiation with Time-Varying Valuations," *Journal of Economics & Management Strategy*, 7, 397–433.

Board, S. and M. Pycia (2014): "Outside options and the failure of the Coase conjecture," *The American Economic Review*, 104, 656–671.

Callander, S. and N. Matouschek (2014): "Managing on Rugged Landscapes," Tech. rep.

Carmichael, H. L. and W. B. MacLeod (2000): "Worker cooperation and the ratchet effect," *Journal of Labor Economics*, 18, 1–19.

Chassang, S. (2010): "Building routines: Learning, cooperation, and the dynamics of incomplete relational contracts," *The American Economic Review*, 100, 448–465.

Compte, O. and P. Jehiel (2002): "On the role of outside options in bargaining with obstinate parties," *Econometrica*, 70, 1477–1517.

Dewatripont, M. (1989): "Renegotiation and information revelation over time: The case of optimal labor contracts," *The Quarterly Journal of Economics*, 589–619.

Dillen, M. and M. Lundholm (1996): "Dynamic income taxation, redistribution, and the ratchet effect," *Journal of Public Economics*, 59, 69–93.

Fiocco, R. and R. Strausz (2015): "Consumer standards as a strategic device to mitigate ratchet effects in dynamic regulation," *Journal of Economics & Management Strategy*, 24, 550–569.

Freixas, X., R. Guesnerie, and J. Tirole (1985): "Planning under incomplete information and the ratchet effect," *The review of economic studies*, 52, 173–191.

Fuchs, W. and A. Skrzypacz (2010): "Bargaining with arrival of new traders," *The American Economic Review*, 100, 802–836.

Fudenberg, D., D. K. Levine, and J. Tirole (1985): "Infinite-horizon models of bargaining with one-sided incomplete information," in *Bargaining with incomplete information*, ed. by A. Roth, Cambridge Univ Press, 73–98.

Gerardi, D. and L. Maestri (2015): "Dynamic Contracting with Limited Commitment and the Ratchet Effect," Tech. rep., Collegio Carlo Alberto.

Gibbons, R. (1987): "Piece-rate incentive schemes," *Journal of Labor Economics*, 413–429.

Gibbons, R. and R. Henderson (2012): *What Do Managers Do?: Exploring Persistent Performance Differences Among Seemingly Similar Enterprises*, Harvard Business School.

Gul, F., H. Sonnenschein, and R. Wilson (1986): "Foundations of dynamic monopoly and the Coase conjecture," *Journal of Economic Theory*, 39, 155–190.

Halac, M. (2012): "Relational contracts and the value of relationships," *The American Economic Review*, 102, 750–779.

Halac, M. and A. Prat (2015): "Managerial Attention and Worker Performance," .

Hart, O. D. and J. Tirole (1988): "Contract renegotiation and Coasian dynamics," *The Review of Economic Studies*, 55, 509–540.

Hart, S. (1985): "Nonzero-sum two-person repeated games with incomplete information," *Mathematics of Operations Research*, 10, 117–153.

Hwang, I. and F. Li (2017): "Transparency of outside options in bargaining," *Journal of Economic Theory*, 167, 116–147.

Kahn, L. B. (2010): "The long-term labor market consequences of graduating from college in a bad economy," *Labour Economics*, 17, 303–316.

Kanemoto, Y. and W. B. MacLeod (1992): "The ratchet effect and the market for secondhand workers," *Journal of Labor Economics*, 85–98.

Kennan, J. (2001): "Repeated bargaining with persistent private information," *The Review of Economic Studies*, 68, 719–755.

Laffont, J.-J. and J. Tirole (1988): "The dynamics of incentive contracts," *Econometrica: Journal of the Econometric Society*, 1153–1175.

Lazear, E. P., K. L. Shaw, and C. Stanton (2016): "Making do with less: working harder during recessions," *Journal of Labor Economics*, 34, S333–S360.

Li, J. and N. Matouschek (2013): "Managing conflicts in relational contracts," *The American Economic Review*, 103, 2328–2351.

Malcomson, J. M. (2015): "Relational incentive contracts with persistent private information," *Econometrica, forthcoming.*

Oreopoulos, P., T. Von Wachter, and A. Heisz (2012): "The short-and long-term career effects of graduating in a recession," *American Economic Journal: Applied Economics*, 4, 1–29.

Ortner, J. (2016): "Durable goods monopoly with stochastic costs," *Theoretical Economics, forthcoming.*

Peski, M. (2008): "Repeated games with incomplete information on one side," *Theoretical Economics*, 3, 29–84.

——— (2014): "Repeated games with incomplete information and discounting," *Theoretical Economics*, 9, 651–694.

Schmidt, K. M. (1993): "Commitment through incomplete information in a simple repeated bargaining game," *Journal of Economic Theory*, 60, 114–139.

Sorin, S. (1999): "Merging, reputation, and repeated games with incomplete information," *Games and Economic Behavior*, 29, 274–308.

Watson, J. (1999): "Starting small and renegotiation," *Journal of economic Theory*, 85, 52–90.

——— (2002): "Starting small and commitment," *Games and Economic Behavior*, 38, 176–199.

Wiseman, T. (2005): "A partial folk theorem for games with unknown payoff distributions," *Econometrica*, 73, 629–645.