Progressive Learning*

Avidit Acharya† and Juan Ortner‡

September 22, 2016

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. Moreover, the relationship can exhibit path dependence, with early shock realizations 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 model, adverse selection, ratchet effect, inefficiency, learning, path dependence

*For helpful comments, we would like to thank Steve Callander, Stephane Wolton, and seminar audiences at Boston University, Stanford, Berkeley and the LSE/NYU political economy conference.

†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 then update the terms of their contract. This limits the principal's ability to learn the agent's private information, and reduces her value from their relationship.

This insight from the ratchet effect literature has shed light on many applications including planning problems (Freixas et al., 1985), labor contracting (Gibbons, 1987; Dewatripont, 1989), the economics of 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 fail to hold when the relationship between the principal and the agent 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 is 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 she cannot commit to any 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 a special case of our model in which productivity is constant over time. We find that the equilibrium of our model with productivity shocks differs qualitatively in three ways from the equilibrium of the same model without the shocks.

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 over time. As a result, the principal's value from the relationship gradually improves. We show that under natural assumptions, the principal is only able to induce the agent to reveal his private information at times at which productivity is low. That is, learning takes place in "bad times." We also derive conditions under which the principal ends up fully learning the agent's private information and therefore attains her first-best payoffs in the long-run. These findings also contrast with the key finding of the ratchet effect literature that the principal is unable to induce the agent to reveal his private information.

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 and the equilibrium is stationary. As a result, early shocks can have a lasting impact on the principal's value from the relationship.

The key feature of our model that drives these three results 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 rents that low cost agents 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 allow 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. For example, Kanemoto and MacLeod (1992) show that competition for second-hand workers can alleviate the ratchet effect and allow firms to achieve efficient outcomes. Carmichael and MacLeod (2000) show that, in an infinite repeated game, the threat of future punishment may incentivize the principal not to use the information that the agent discloses to her advantage, thus mitigating the ratchet effect. Fiocco and Strausz (2015) show that the principal can induce information revelation when contracting is strategically delegated to an independent third party. Our paper differs from these studies in that we do not introduce external sources of contract enforce-

ment, nor do we (effectively) reintroduce commitment by allowing for non-Markovian strategies.

Instead, we focus on the role that shocks play in ameliorating the principal's commitment problem.[1] Our approach is to exploit the non-stationarity of the environment created by changes in productivity. In this sense, our work relates to Blume (1998), who also studies the ratchet effect in a non-stationary environment where the agent's private information (his valuation) rather than the level of productivity, changes over time. While our results contrast with the results of the ratchet effect literature, Blume's (1988) results generalize the main findings in the literature to a different setting.

Our model is also related 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.

Our 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). For example, 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 i.i.d. private information about her opportunity cost of paying the agent 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.

Finally, our paper relates to a broader literature on dynamic games with private information (Hart, 1985; Sorin, 1999; Wiseman, 2005; Peski, 2008, 2014). Within that literature our paper relates closely to work by Watson (1999, 2002), who studies a partnership game between privately informed players 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.

[1]The idea that time-varying shocks can ameliorate a player's lack of commitment also appears in Ortner (2016), who studies the problem of a durable goods monopolist with time-varying production costs. He shows that in contrast to the classic Coase conjecture results in Gul et al. (1986) and Fudenberg et al. (1985), a monopolist with time-varying costs may extract rents from consumers.

2 Model

2.1 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 $a \times c > 0$ when choosing action $a \in \{0, 1\}$. The agent's cost c is his private information, and it is fixed throughout the game. Cost 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.[2] The payoffs to the principal and an agent of cost $c = c_k$ at the end of period t are, respectively,

$$u(b_t, T_t, a_t) = a_t \left[b_t - T_t \right],$$
$$v_k(b_t, T_t, a_t) = a_t \left[T_t - c_k \right].$$

Both players are risk-neutral expected utility maximizers and share a common discount factor $\delta < 1$.[3]

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 (efficient) 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

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

[2] 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.

[3] Our results remain qualitatively unchanged if the principal and agent have different discount factors.

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

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

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.

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

When there are no shocks (i.e., the state is fully persistent) the above assumption holds when $\delta > 1/2$. In general, for any ergodic process $\{b_t\}$, the assumption holds whenever δ is above a cutoff $\underline{\delta} > 1/2$. Conversely, for any $\delta > 1/2$, the assumption holds whenever the process $\{b_t\}$ is sufficiently persistent; that is, whenever $Q_{b,b} = \text{Prob}(b_{t+1} = b | b_t = b)$ is sufficiently large for all $b \in \mathcal{B}$.

2.2 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 choices that have been realized 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. As usual, we let H_t denote the set of period t histories 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.

A pure strategy perfect Bayesian equilibrium (PBE) is a strategy profile σ and posterior beliefs μ for the principal such that the strategies form a Bayesian Nash equilibrium in every continuation game given the posterior beliefs, and beliefs are consistent with Bayes rule whenever possible. Thus, pure strategy PBE can be denoted by the pair (σ, μ). For any pure strategy PBE (σ, μ), we denote the principal's belief about the agent's cost after history h_t by $\mu[h_t]$. For any pure strategy 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 $V_k^{(\sigma,\mu)}[h_t, b_t]$, respectively.

Our equilibrium concept is pure strategy PBE that are Markovian with respect to the shock variable and principal's beliefs, and optimal for the principal after every history among such equilibria. More precisely, we use "equilibrium" to refer to pure strategy PBE that satisfy the two conditions below, where we identify $\tau(h_t, \cdot) : \mathcal{B} \to \mathbb{R}_+$ and $\alpha_k(h_t, \cdot, \cdot) : \mathcal{B} \times \mathbb{R}_+ \to \{0, 1\}$ with the continuation strategies of the principal and agent with cost c_k, given history h_t.

R1. (Markovian) *For all histories h_t and $h_{t'}$, if $\mu[h_t] = \mu[h_{t'}]$ then*

$$\tau(h_t, \cdot) = \tau(h_{t'}, \cdot) \text{ and } \alpha_k(h_t, \cdot, \cdot) = \alpha_k(h_{t'}, \cdot, \cdot) \text{ for all } k = 1, ..., K.$$

R2. (best for principal) *There is no history h_t, shock $b_t \in \mathcal{B}$ and pure strategy PBE (σ', μ') that also satisfies the Markovian condition, R1, for which*

$$U^{(\sigma',\mu')}[h_t, b_t] > U^{(\sigma,\mu)}[h_t, b_t]$$

We impose these restrictions to rule out indirect sources of commitment for the principal. Condition R1 rules out equilibria in which the threat of future punishment enforces high continuation payoffs for the agent. Condition R2 rules out Markovian equilibria in which off path beliefs are constructed in ways that make the principal give the agent extra rents.[4] Without restrictions R1 and R2, the principal could use the promise of a high continuation payoff to induce the agent to reveal his private information. As we show in Lemma 0 below, the main implications of these restrictions is that the highest cost type in the support of the principal's belief has a zero continuation payoff at every history, and the agent's local incentive constraints always hold with equality.

We end this section by noting that our equilibrium refinement facilitates a direct comparison with prior papers on the ratchet effect; for instance, Hart and Tirole (1988) and Schmidt (1993).[5] As we will show below, this refinement selects a unique equilibrium that naturally generalizes the equilibrium studied in these papers. Indeed, 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]Markovian equilibria in which the principal offers high transfers can be constructed by specifying off path beliefs that "punish" an agent who accepts low transfers. Such beliefs incentivize the agent to reject low transfers, and by doing this they also incentivize the principal to offer high transfers.

[5]Hart and Tirole (1988) and Schmidt (1993) study games with long but finite time horizons, while Gerardi and Maestri (2015) assume that the time horizon is infinite. We take the latter approach, and focus on pure strategy PBE that satisfy R1 and R2.

3 Equilibrium Analysis

3.1 Incentive Constraints

Fix an equilibrium $(\sigma, \mu) = ((\tau, \{\alpha_k\}), \mu)$. Recall that for any $h_t \in H$, $\mu[h_t]$ denotes the principal's beliefs after history h_t. We will 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 index of types in the support of the principal's beliefs. 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 an agent of type c_k takes in period t.

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

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

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

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

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

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

is the expected discounted number of times that type c_j takes the action after (h_t, b) under equilibrium (σ, μ). Equation (1) 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),

$$V_i^{(\sigma,\mu)}[h_t, b_t] \geq V_{i\to j}^{(\sigma,\mu)}[h_t, b_t] \qquad \forall (h_t, b_t), \forall c_i, c_j \in C[h_t] \tag{2}$$

We then have the following lemma, which follows from conditions R1 and R2. Part (i) says that the highest cost type in the support of the principal's beliefs obtains a continuation payoff equal to zero, while part (ii) says that "local" incentive constraints bind. All proofs that don't appear in the main text are in the Appendix.

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* $c_i \in C[h_t] \backslash \{c_{\overline{k}[h_t]}\}$.

3.2 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 states 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 $X(b_t, E_{\overline{k}[h_t]}) > 1$, all types in $C[h_t]$ take action $a = 0$.

(iii) If $b_t \notin E_{\overline{k}[h_t]}$ and $X(b_t, E_{\overline{k}[h_t]}) \leq 1$, 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 (which is accepted by types in C^-) satisfies

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

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

Theorem 1 describes the (essentially) unique equilibrium. At states $(\mu[h_t], b_t)$ satisfying the conditions in parts (i) or (ii), all the agent types that the principal believes are possible take the same action. Hence, the principal learns nothing about the agent's type at such states. The proof of Theorem 1 shows that, when the state $(\mu[h_t], b_t)$ satisfies these conditions, an agent with cost $c_i < c_{\overline{k}[h_t]}$ gets large rents by mimicking an agent with cost $c_{\overline{k}[h_t]}$. At the same time, low cost agents anticipate that the principal would leave them with no rents in the future if they were to reveal their private information. As a result, at these states the principal is unable to induce information revelation by low cost agents.

Equilibrium behavior is, however, quite different at states satisfying the conditions in parts (i) and (ii). When $b_t \in E_{\overline{k}[h_t]}$, there is an *efficient* ratchet effect at play. At these states the agent takes the socially efficient action $a = 1$, but the principal compensates him as if he was the highest cost type, replicating the main finding in the 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 at every periods, and she never learns anything about the agent's type.[6]

Part (ii), in contrast, characterizes states $(\mu[h_t], b_t)$ at which an *inefficient* ratchet effect is at play. At such states, 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. To see why, recall that for any $b \in \mathcal{B}$ and any $B \subset \mathcal{B}$, $X(b, B)$ measures how often the process $\{b_t\}$ visits states in B, conditional on the current state being b. When $X(b_t, E_{\overline{k}[h_t]}) > 1$, low cost agents expect to obtain large rents by mimicking an agent with cost $c_{\overline{k}[h_t]}$, making it impossible for the principal to induce them to reveal their type.

Part (iii) characterizes states $(\mu[h_t], b_t)$ 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 rents that agents with type $c_i < c_{\overline{k}[h_t]}$ get from mimicking an agent with the highest cost $c_{\overline{k}[h_t]}$ are small when $X(b_t, E_{\overline{k}[h_t]}) \leq 1$. As a result, the principal is able to induce low cost agents to reveal their private information. In Appendix B.3 we provide a characterization of the threshold cost c_{k^*} in part (iii) as the solution to a finite maximization problem. Building

[6]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.

on this, we also characterize the principal's equilibrium payoffs as the fixed point of a contraction mapping.

3.3 Examples

We now present two examples that illustrate some of the main equilibrium features of our model. The first example highlights the fact that equilibrium outcome in our model can be inefficient. This contrasts with the results in Hart and Tirole (1988) and Schmidt (1993), where the equilibrium is always socially optimal.

Example 1 (inefficient ratchet effect) Suppose that there are two states, $\mathcal{B} = \{b_L, b_H\}$, with $0 < b_L < b_H$, and two types, $\mathcal{C} = \{c_1, c_2\}$ (recall our assumption that $c_2 > c_1$). Assume further that $E_1 = \{b_L, b_H\}$, $E_2 = \{b_H\}$, and $X(b_L, \{b_H\}) > 1$.

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 $X(b_L, \{b_H\}) > 1$ 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. □

The next example illustrates a situation in which the principal learns the agent's type, and the equilibrium outcome is efficient. This too contrasts with earlier work on the ratchet effect in which there is no learning by the principal.

Example 2 (efficiency and learning) The environment is the same as in Example 1, with the only difference that $X(b_L, \{b_H\}) < 1$. Consider a history h_t such that $C[h_t] = \{c_1, c_2\}$. As in Example 1, 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. To see why, note first that by Theorem 1, an agent of type c_2 does not take the action at time t if $b_t = b_L$. Suppose by contradiction that type c_1 does not take the action when $b_t = b_L$ either. Since the equilibrium is Markovian, this implies that the principal never learns the agent's type, and her payoff at time t when $b_t = b_L$ is $\underline{U} = X(b_L, \{b_H\})[b_H - c_2]$.

Consider an alternative equilibrium in which at such a history the principal makes an offer T that only an agent with cost c_1 accepts. Suppose that the principal's offer exactly

compensates type c_1 for revealing his type; i.e. $T - c_1 = X(b_L, \{b_H\})(c_2 - c_1)$.[7] Note that $X(b_L, \{b_H\}) < 1$ implies that $T < c_2$, so an agent with cost c_2 rejects offer T. Conditional on the agent's cost being c_1, the principal's payoff from making offer T when $b_t = b_L$ is

$$\begin{aligned}\overline{U}[c_1] &= b_L - T + X(b_L, \{b_L\})[b_L - c_1] + X(b_L, \{b_H\})[b_H - c_1] \\ &= [1 + X(b_L, \{b_L\})][b_L - c_1] + X(b_L, \{b_H\})[b_H - c_2],\end{aligned}$$

where the first equality follows since, under the proposed equilibrium, the principal would learn at time t that the agent's type is c_1. On the other hand, conditional on the agent's type being c_2, the agent would reject the offer and the principal's payoff would be $\overline{U}[c_2] = X(b_L, \{b_H\})[b_H - c_2] = \underline{U}$. Under this equilibrium, the principal's expected payoff at history h_t with $C[h_t] = \{c_1, c_2\}$ when $b_t = b_L$ is $\mu_0[c_1]\overline{U}[c_1] + \mu_0[c_2]\overline{U}[c_2] > \underline{U}$, where $\mu_0[c_j]$ is the prior probability that the agent's cost is c_j. The inequality follows since $\overline{U}[c_2] = \underline{U}$ and since $\overline{U}[c_1] > \underline{U}$. Therefore, in any Markovian PBE satisfying R2, type c_1 takes the action in state b_L when $C[h_t] = \{c_1, c_2\}$.

Finally, note that the principal learns the agent's type at time $\bar{t} = \min\{t : b_t = b_L\}$, and the equilibrium 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. □

Example 2 has three notable features. First, despite of her lack of commitment the principal learns the agent's type. Learning takes place the first time the relationship hits the low productivity state. In the next subsection we present conditions under which this result generalizes.

Second, the principal's value increases over time, since the surplus she extracts from the agent increases as she learns the agent's type. In Section 4 we characterize general conditions under which the principal obtains her first-best payoff in the long-run.

Third, the equilibrium exhibits a form of path-dependence: equilibrium play at time t depends on the entire history of shocks up to period t. 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 transfer equal to c_1. Note, however, that the path dependence in

[7] The payoff a low cost agent obtains by accepting offer T is $T - c_1 + \delta \times 0$, 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 $X(b_L, \{b_H\})(c_2 - c_1)$.

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. It turns out, however, that this is *not* a general property of our model. In Section 4 we show by example that the equilibrium may also display long-run path dependence.

3.4 Learning in Bad Times

We now show that, under natural conditions on the process $\{b_t\}$ that governs the evolution of the stochastic shock, the principal learns about the agent's type only in "bad times," i.e., only when the benefit b_t is small. Recall that $\mu[h_t]$ denotes the principal's beliefs about the agent's type after history h_t. Then, we have:

Proposition 1. (learning in bad times) *Suppose that for all $c_k \in \mathcal{C}$, $X(b, E_k) \leq X(b', E_k)$ for all $b, b' \in \mathcal{B}$ with $b < b'$. Then, in any equilibrium and for every history h_t there exists a state $b[h_t] \in \mathcal{B}$ such that $\mu[h_{t+1}] \neq \mu[h_t]$ only if $b_t < b[h_t]$.*

Proof. By Theorem 1, $\mu[h_{t+1}] \neq \mu[h_t]$ only if the shock b_t is such that $X(b_t, E_{\overline{k}[h_t]}) \leq 1$. By the assumption that for all types c_k, $b < b'$ implies $X(b, E_k) \leq X(b', E_k)$, there exists a state $b[h_t] \in \mathcal{B}$ such that $X(b_t, E_{\overline{k}[h_t]}) \leq 1$ if and only if $b_t < b[h_t]$.[8] $\square$

Proposition 1 provides conditions under which the principal only updates her beliefs about the agent's type at states at which the benefits from taking the productive action are sufficiently small. Intuitively, under the conditions in Proposition 1, 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 at times at which the benefit b_t is low.

4 Long Run Properties

In this section, we study the long run properties of the equilibrium. 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

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

$h_{t+1} \succeq h_t$ for all t. Because we focus on pure strategy Markovian equilibria and because the sets of types and states are finite, 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 $t^*[h_\infty]$. Therefore, 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 $U_{LR}^{(\sigma,\mu)}(b_t|h_{t^*[h_\infty]})$. We use this fact in the next two subsections to study properties of the principal's long run value.

4.1 The Principal's Long Run Value

We start by studying the extent to which the principal can learn the agent's type, and how the efficiency of the relationship might improve over time.

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

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

Thus, 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. We then say that an equilibrium (σ, μ) is *long run first best* if for all $c_k \in \mathcal{C}$, under this equilibrium the set of 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 $c = c_k$. This says that no matter what the agent's true type is, once learning has stopped the principal achieves her first best payoff at every subsequent realization of the shock with probability one. The following proposition reports a sufficient condition for the equilibrium to be long run first best.

Proposition 2. (long run first best) *Suppose that $\{b_t\}$ is ergodic and that for all $c_k \in \mathcal{C}\backslash\{c_K\}$ there exists $b \in E_k\backslash E_{k+1}$ such that $X(b, E_{k+1}) < 1$. Then, the equilibrium is long run first best.*

The condition in the statement of Proposition 2 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, it is easy to find 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 $X(b, E_{\underline{k}}) < 1$, 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.[9]

The next and final result of this subsection provides a partial counterpart to Proposition 2 by presenting conditions under which the equilibrium is not long run first best, and conditions under which it is not long run efficient.

Proposition 3. (no long run first best; no long run efficiency) *Let h_t be an equilibrium history such that $|C[h_t]| \geq 2$ and $X(b, E_{\overline{k}[h_t]}) > 1$ 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.*

Proof. Follows from Theorem 1. □

4.2 Path Dependence

In the examples of Section 3.3 the principal always learns the same amount of information about the agent's type in the long run. As a result, even if equilibrium play may exhibit path dependence in the short-run, as in Example 2, the principal's long run value from the relationship, conditional on the agent's type, is independent of the history of play. In this section we show that this is not a general property of the equilibrium of our model. We

[9] The proof of this result is similar to the proof of Proposition 2.

show here that the learning process, and hence the principal's value from the relationship, may be exhibit path dependence even in the long run.

We 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.

In this section, we show by example that equilibrium may exhibit long-run path dependence even when process $\{b_t\}$ is ergodic. A simpler example in Appendix D shows that long run path dependence can arise when process $\{b_t\}$ is not 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 $E_3 = \{b_H\}$. Thus, in the most productive state, it is socially optimal for all types to take the productive action; in the next two most productive states, it is socially optimal for only the two lowest cost types to take the productive action; and in the least productive state it is not socially optimal for any type to take the productive action.

Lemma 1. *Suppose that the transition matrix $[Q_{b,b'}]$ satisfies:*

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

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

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 $a = 1$.

Properties (i)-(iii) in Lemma 1 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)). We then have the following result.

Proposition 4. (long run path dependence) *Under the conditions in Lemma 1, the equilibrium exhibits long-run path dependence.*

Proof. Follows from Lemma 1. □

Proposition 4 shows that, even when the process $\{b_t\}$ is ergodic, the information that the principal learns about the agent's type in the long run might be influenced by the history of productivity shocks early on in the relationship.

The intuition behind this result is as follows. The information rents that a c_1-agent gets by mimicking a c_2-agent depend on how often the c_2-agent is expected to take the productive action in the future (see equation (1)). In turn, how often a c_2-agent 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 a c_2-agent will take the action whenever the state is in $E_2 = \{b_{ML}, b_{MH}, b_H\}$ (see Theorem 1(i)).

In contrast, at histories h_t at which the principal assigns positive probability to types c_2 and c_3 (i.e., histories h_t such that $c_2, c_3 \in C[h_t]$), a c_2-agent will not take the action at time t if $b_t = b_{MH}$. Indeed, since $X(b_{MH}, E_1) > 1$, by Theorem 1(ii) the agent does not take the action time t if $b_t = b_{MH}$ whenever h_t is such that $c_2, c_3 \in C[h_t]$. Moreover, by Lemma 1, at such histories h_t a c_2-agent is expected to take action $a = 1$ at periods $t' > t$ with $b_{t'} = b_{MH}$ only if $t' > t(b_{ML}) = \min\{t : b_t = b_{ML}\}$.[10] Since $Q_{b,b_{ML}} < \epsilon_2$ for all $b \in \mathcal{B}\backslash\{b_{ML}\}$, time $t(b_{ML})$ is large in expectation. Therefore, a c_2-agent is expected to take

[10] At all periods $t' \leq t(b_{ML})$ with $b_{t'} = b_{MH}$, a c_2-agent takes action $a = 0$ (by Lemma 1).

the action significantly less frequently in the future at a history h_t with $C[h_t] = \{c_1, c_2, c_3\}$ than at a history h'_t with $C[h'_t] = \{c_1, c_2\}$.

As a consequence of this, the cost of inducing a c_1-agent to reveal his private information depends on the principal's beliefs. In particular, inducing a c_1-agent to reveal his private information is cheaper at histories $(h_t, b_t = b_L)$ with $C[h_t] = \{c_1, c_2, c_3\}$ than at histories $(h'_t, b_t = b_L)$ with $C[h'_t] = \{c_1, c_2\}$. As the proof of Lemma 1 shows, this difference makes it optimal for the principal to induce a c_1-agent to reveal his type when $C[h_t] = \{c_1, c_2, c_3\}$ and $b_t = b_L$, and at the same time makes it suboptimal to induce this agent to reveal his type at histories h_t such that $C[h_t] = \{c_1, c_2\}$ and $b_t = b_L$.

Finally, by Lemma 1(iii), at histories (h_t, b_t) with $C[h_t] = \{c_1, c_2, c_3\}$ and $b_t = b_{ML}$, the principal finds it optimal to make an offer that only types in $\{c_1, c_2\}$ accept, and type c_3 rejects. The principal stops learning after making this offer, and her long run value is not first best when the agent's type is c_1. At such a history, the principal has the option of waiting until shock b_L is realized, and at that point make an offer that only an agent with type c_1 accepts. By doing so, the principal would obtain first best payoffs in the long run. However, our assumption that Q_{b,b_L} is small for all $b \in \mathcal{B}\backslash\{b_L\}$ makes this deviation unprofitable for the principal: under this assumption, the forgone surplus until state b_L is reached outweighs the long run gains for the principal.

5 Conclusion

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 show that this is true of the traditional ratchet effect literature. A takeaway from this literature is that outside institutions that provide contract enforcement can help improve the principal's welfare. 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. In addition, a relationship that was initially highly inefficient may become efficient over time. Finally, whether or not the relationship ever becomes efficient, and how profitable it becomes for the principal, may be path dependent.

Appendix

A. 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 $|C[h_t]| = 1$.[11] 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 state $(b_{t'}, \mu[h_{t'}])$ the agent is offered a transfer $\tilde{T} \in (c_{\overline{k}[h_t]}, T_{t'})$. Specify the principal's beliefs at state $(b_{t'}, \mu[h_{t'}])$ 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 states, 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 contradicts R2. □

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 $A^{(\sigma,\mu)}_2[h_t, b_t] = A^{(\sigma,\mu)}_1[h_t, b_t]$.

[11] Indeed, if $C[h_t] = \{c_i\}$, then in any PBE satisfying R1 and R2 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, $V_1^{(\sigma,\mu)}[h_t, b_t] = \mathbb{E}^{(\sigma,\mu)}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}(T_{t'} - c_1)a_{t',1}|h_t\right] = V_2^{(\sigma,\mu)}[h_t, b_t] + A_2^{(\sigma,\mu)}[h_t, b_t](c_2 - c_1)$, where we have used the facts that $V_2^{(\sigma,\mu)}[h_t, b_t] = 0$ and $T_{t'} = c_2$ for all t' such that $a_{t',1} = a_{t',2} = 1$ (both of these facts 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 $V_1^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] = T_{\underline{t}} - c_1$. Note further that $T_{\underline{t}} - c_1 \geq V_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_2 - c_1)$, 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 stationary PBE that are optimal for the principal, the transfer that the principal offers the agent at time $\underline{t}$ must be $T_{\underline{t}} = c_1 + V_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_2 - c_1)$, and so

$$V_1^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] = V_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_2 - c_1) = A_2^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_2 - c_1), \tag{3}$$

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,

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

where we have used (3), 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),

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

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 $V_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] = T_{\underline{t}} - c_k + 0 = V_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{k+1} - c_k)$.[12] Note further that $V_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] \geq T_{\underline{t}} - c_{k+1}$, since an agent with cost c_{k+1} can guarantee $T_{\underline{t}} - c_{k+1}$ by taking the action at time $\underline{t}$ and then not taking the action in all future periods. Since $T_{\underline{t}} - c_k = V_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{k+1} - c_k)$, it follows that $A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] \leq 1$.

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

[12] 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 stationary PBE that are optimal for the principal, 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

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

where the inequality follows since $A_{k+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] \leq 1 < A_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}]$.[13] 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

$$V_i^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] = T_{\underline{t}} - c_i + 0 = V_{i+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] + A_{i+1}^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}](c_{i+1} - c_i),$$

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

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

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 $|C[h_t]| = n$. □

B. 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 (ii) and

[13] Recall that, for all (h_t, b_t), $A_k^{(\sigma,\mu)}[h_t, b_t] = \mathbb{E}^{(\mu,\sigma)}[\sum_{t'=t}^{\infty} \delta^{t'-t} a_{t',k} | b_t, h_t]$. 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, $A_k^{(\sigma,\mu)}[h_{\underline{t}}, b_{\underline{t}}] > 1$.

(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).

B.1. Proof of part (i) (the case of $b_t \in E_{\overline{k}[h_t]}$)

We prove part (i) of the theorem by strong induction on the cardinality of $C[h_t]$. If $C[h_t]$ is a singleton $\{c_k\}$, the statement of part (i) holds: by R1-R2, 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 be a history such that $|C[h_t]| = n$, and let $b_t \in E_{\overline{k}[h_t]}$. We show that, at such a history/shock pair (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 that satisfies R1-R2, 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 Markovian PBE which is identical to the original PBE, except that when the principal's beliefs are $\mu[h_t]$ and the shock is 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 satisfies R1-R2. 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 $b_t \in E_{\overline{k}[h_t]}$.

We now show that, in a PBE that satisfies R1-R2, 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: (i) $c_j = c_{\overline{k}[h_t]}$, and (ii) $c_j < c_{\overline{k}[h_t]}$. Consider case (i). 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 $T_t - c_i = c_{\overline{k}[h_t]} - c_i > 0$ 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 (ii). By Lemma 0, the payoff of type c_j from taking the productive action at time t is $T_t - c_j + 0$. 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

$$\begin{aligned} T_t - c_j &\geq V^{(\sigma,\mu)}_{\overline{k}[h_t]}[h_t, b_t] + (c_{\overline{k}[h_t]} - c_j) A^{(\sigma,\mu)}_{\overline{k}[h_t]}[h_t, b_t] \\ &\geq (c_{\overline{k}[h_t]} - c_j) X(b_t, E_{\overline{k}[h_t]}) > c_{\overline{k}[h_t]} - c_j \end{aligned} \tag{4}$$

The first inequality follows from equation (2) in the main text. The second inequality follows from Lemma 0 and the fact that $A^{(\sigma,\mu)}_{\overline{k}[h_t]}[h_t, b_t] \geq X(b_t, E_{\overline{k}[h_t]})$. 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 $A^{(\sigma,\mu)}_{\overline{k}[h_t]}[h_t, b_t] \geq X(b_t, E_{\overline{k}[h_t]})$. The last inequality in equation (4) follows from the fact $X(b_t, E_{\overline{k}[h_t]}) \geq X(b_t, \{b_t\}) > 1$ 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 $T_t = c_{\overline{k}[h_t]}$. □

B.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 2 below, and use the lemma to prove parts *(ii)* and *(iii)* separately.

Lemma 2. *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 $T_t - c_j = c_{\overline{k}[h_t]} - c_j > 0$ by accepting the current offer. Hence, it cannot be that some agents reject the offer $T_t = c_{\overline{k}[h_t]}$ when an agent with type $c_{\overline{k}[h_t]}$ accepts the offer.

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 $b_t - c_{\overline{k}[h_t]} < 0$. Consider an alternative Markovian PBE which is identical to the original PBE, except that when the principal's beliefs are $\mu[h_t]$ and the shock is $b_t \notin E_{\overline{k}[h_t]}$, 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 satisfies R1-R2. Hence, if $b_t \notin E_{\overline{k}[h_t]}$, an agent with type $c_{\overline{k}[h_t]}$ does 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 $X(b_t, E_{\overline{k}[h_t]}) > 1$. By Lemma 2, 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 $T_t - c_j + 0$ 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

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

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 $X(b_t, E_{\overline{k}[h_t]}) > 1$ by assumption. The inequalities in (5) 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 $a_t = 0$. $\square$

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 $X(b_t, E_{\overline{k}[h_t]}) \leq 1$. 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 2 it must be that $c_{j^*} < c_{\overline{k}[h_t]}$. By Lemma 0, type c_{j^*}'s payoff is $T_t - c_{j^*}$, since from date $t+1$ onwards this type will be the largest cost 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

$$T_t - c_{j^*} \geq V_{k^*}^{(\sigma,\mu)}[h_t, b_t] + (c_{k^*} - c_{j^*}) A_{k^*}^{(\sigma,\mu)}[h_t, b_t], \tag{6}$$

since type c_{j^*} can obtain the right-hand side of (6) by mimicking type c_{k^*}. Furthermore, type c_{k^*} can guarantee himself a payoff of $T_t - c_{k^*}$ by taking the action at time t and never taking the action again. Therefore, it must be that

$$\begin{aligned} V_{k^*}^{(\sigma,\mu)}[h_t, b_t] \geq T_t - c_{k^*} \geq c_{j^*} - c_{k^*} + V_{k^*}^{(\sigma,\mu)}[h_t, b_t] + (c_{k^*} - c_{j^*}) A_{k^*}^{(\sigma,\mu)}[h_t, b_t] \\ \implies 1 \geq A_{k^*}^{(\sigma,\mu)}[h_t, b_t] \end{aligned} \tag{7}$$

where the second inequality in the first line follows from (6).

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

$$\begin{aligned} V_{i^* \to j^*}^{(\sigma,\mu)}[h_t, b_t] &= T_t - c_{j^*} + (c_{j^*} - c_{i^*}) A_{j^*}^{(\sigma,\mu)}[h_t, b_t] \\ &= T_t - c_{j^*} + (c_{j^*} - c_{i^*})(1 + X(b_t, E_{j^*})) \\ &\geq (c_{j^*} - c_{i^*})[1 + X(b_t, E_{j^*})] + V_{k^*}^{(\sigma,\mu)}[h_t, b_t] + (c_{k^*} - c_{j^*}) A_{k^*}^{(\sigma,\mu)}[h_t, b_t] \end{aligned} \tag{8}$$

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 (6).

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 does not have type in $\{c_{i^*+1}, ..., c_{j^*}\} \subseteq C[h_t]$ after observing a rejection. That is, type c_{i^*}'s payoff is

$$V_{i^*}^{(\sigma,\mu)}[h_t, b_t] = V_{i^* \to k^*}^{(\sigma,\mu)}[h_t, b_t] = V_{k^*}^{(\sigma,\mu)}[h_t, b_t] + (c_{k^*} - c_{i^*}) A_{k^*}^{(\sigma,\mu)}[h_t, b_t] \tag{9}$$

From equations (8) and (9), it follows that

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

where the strict inequality follows after using (7). 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 $T_t - c_{j^*}$, 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 $T_t = c_{j^*} + V_{k^*}^{(\sigma,\mu)}[h_t, b_t] + (c_{k^*} - c_{j^*}) A_{k^*}^{(\sigma,\mu)}[h_t, b_t]$. □

B.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 $b \in \mathcal{B}$, let (h_t, b_t) denote the concatenation of history $h_t = \langle b_{t'}, T_{t'}, a_{t'} \rangle_{t'=0}^{t-1}$ together with state realization b_t. Let

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

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 3. *Fix any equilibrium (σ, μ) and history-state pair (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 $A_{j+}^{(\sigma,\mu)}[h_t, b_t] \leq 1$.*

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 $T - c_k + 0$. 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

$$T - c_k \geq V_j^{(\sigma,\mu)}[h_t, b_t] + (c_j - c_k)A_{j+}^{(\sigma,\mu)}[h_t, b_t]. \tag{10}$$

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

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

where the second inequality in the first line follows after substituting T from (10).

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

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

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

The proof of existence and uniqueness relies on Lemma 3 and uses strong induction on the cardinality of $C[h_t]$. Clearly, equilibrium exists and equilibrium payoffs are unique at histories h_t such that $C[h_t]$ is a singleton $\{c_k\}$: 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 equilibrium exists and equilibrium payoffs are unique for all histories $\tilde{h}_{\tilde{t}}$ such that $|C[\tilde{h}_{\tilde{t}}]| \leq n-1$, and let h_t be a history such that $|C[h_t]| = n$. Fix a candidate for equilibrium (σ, μ), and let $U^{(\sigma,\mu)}[b_t, \mu[h_t]]$ denote the principal's equilibrium payoffs when her beliefs are $\mu[h_t]$ and the shock is b_t. We now show that, when the principal's beliefs are $\mu[h_t]$, equilibrium payoffs are also unique.

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

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

If $b_t \notin E_{\overline{k}[h_t]}$ and $X(b_t, E_{\overline{k}[h_t]}) > 1$, 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

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

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 $X(b_t, E_{\overline{k}[h_t]}) \leq 1$. 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.[14] 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

[14] By Lemma 2, in equilibrium an agent with cost $c_{\overline{k}[h_t]}$ doesn't take the action.

payoffs are

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

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 3, 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 $A_{j+}^{(\sigma,\mu)}[h_t, b_t] \leq 1$. Note that, by the induction hypothesis, $A_{j+}^{(\sigma,\mu)}[h_t, b_t]$ is uniquely determined.[15] Let $C^*[h_t, b_t] = \{c_i \in C[h_t] : A_{i+}^{(\sigma,\mu)}[h_t, b_t] \leq 1\}$. 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

$$T_{t,i-1}^* = c_{i-1} + V_i^{(\sigma,\mu)}[h_t, b_t] + A_{i+}^{(\sigma,\mu)}[h_t, b_t](c_i - c_{i-1})$$

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 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 everyone. 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 $U^{(\sigma,\mu)}[b_t, \mu[h_t], T_{t,i-1}^*, a_t = 1]$ and $U^{(\sigma,\mu)}[b_t, \mu[h_t], T_{t,i-1}^*, a_t = 0]$ denote the principal's expected continuation payoffs if the offer $T_{t,i-1}^* \in \mathcal{T}$ is accepted and rejected, respectively, at state $(\mu[h_t], b_t)$. 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

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

[15] $A_{j+}^{(\sigma,\mu)}[h_t, b_t]$ 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.

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

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

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

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

$$U^{(\sigma,\mu)}[b, \mu[h_t]] = \begin{cases} b - c_{\overline{k}[h_t]} + \delta \mathbb{E}\big[U^{(\sigma,\mu)}[b_{t+1}, \mu[h_t]] | b_t = b\big] & \text{if } b \in B_1 \\ \delta \mathbb{E}\big[U^{(\sigma,\mu)}[b_{t+1}, \mu[h_t]] | b_t = b\big] & \text{if } b \in B_2 \\ \max\{U^*(b, \mu[h_t]), \delta \mathbb{E}\big[U^{(\sigma,\mu)}[b_{t+1}, \mu[h_t]] | b_t = b\big]\} & \text{if } b \in B_3 \end{cases} \tag{11}$$

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 $f \in \mathcal{F}$,

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

One can check that Φ is a contraction of modulus $\delta < 1$, and therefore has a unique fixed point. Moreover, by (11), the principal's equilibrium payoffs $U^{(\sigma,\mu)}[b, \mu[h_t]]$ are a fixed point of Φ. These two observations together imply that the principal's equilibrium payoffs $U^{(\sigma,\mu)}[b, \mu[h_t]]$ are unique. Finally, the equilibrium strategies at (h_t, b_t) can be immediately derived from (11). □

C. Proof of Proposition 2

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 at state $(\mu[h_t], b)$ 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) is

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

Let $b \in E_{\overline{k}[h_t]-1}$ be such that $X(b, E_{\overline{k}[h_t]}) < 1$. The conditions in the statement of Proposition 2 guarantee that such a shock b exists. Suppose that the shock at time t after history h_t is b, and let $\epsilon > 0$ be small enough such that

$$T = c_{\overline{k}[h_t]-1} + X(b, E_{\overline{k}[h_t]})(c_{\overline{k}[h_t]} - c_{\overline{k}[h_t]-1}) + \epsilon < c_{\overline{k}[h_t]}. \tag{12}$$

Note that at state $(\mu[h_t], b)$, 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 $c_{\overline{k}[h_t]}$.[16] The principal's payoff from making an offer T conditional on the agent's type being $c_{\overline{k}[h_t]}$ is $U^{(\sigma,\mu)}[h_t, b]$. On the other hand, when the agent's type is lower than $c_{\overline{k}[h_t]}$, the principal obtains $b - T$ 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 principal obtains from following this strategy when the agent's cost is lower than $c_{\overline{k}[h_t]}$ is

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

[16] Indeed, by accepting offer T, an agent with cost $c_i < c_{\overline{k}[h_t]}$ obtains a payoff of at least $T - c_i + \delta \times 0$. 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.

where the second line follows from substituting (12). 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 state $(\mu[h_t], b)$ 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 $U^{(\sigma,\mu)}[h_t, b] = \mathbb{E}\left[\sum_{t'=t}^{\infty} \delta^{t'-t}(b_{t'} - c_i)\mathbf{1}_{\{b_{t'} \in E_i\}} | b_t = b\right] = U_i^*(b|c = c_i)$, which is the first best payoff. This and the previous arguments imply that the equilibrium is long run first best. □

D. Path Dependence when Shocks are not Ergodic

Here we give an example of how path dependence may arise when the process governing the evolution of shocks is not ergodic.

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 process $[Q_{b,b'}]$ satisfies: (i) $X(b_L, E_2) < 1$, and (ii) $Q_{b_H,b_H} = 1$ and $Q_{b,b'} \in (0,1)$ for all $(b, b') \neq (b_H, b_H)$ (recall that $Q_{b,b'}$ denotes the probability of transitioning to state b' from state b.) Thus, state b_H is absorbing. By Theorem 1, if $b_t = b_H$, then from period t onwards the principal makes an offer equal to $c_{\overline{k}[h_t]}$ and all agent types in $C[h_t]$ accept.

Consider a 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 $b_t = b_L$, by arguments similar to those in Example 2, the principal finds it optimal to make an offer $T_t = c_1 + X(b_L, E_2)(c_2 - c_1) \in (c_1, c_2)$ that an agent with cost c_1 accepts and that an agent with cost c_2 rejects. Therefore, the principal learns the agent's type.

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

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

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. Instead, 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 $\frac{1}{1-\delta}(b_H - c_2)$, while her payoff at history $(\tilde{h}_t, b_t = b_H)$ is $\frac{1}{1-\delta}(b_H - c_1)$. □

E. Proof of Lemma 1

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 $T_t = c_1 + X(b_L, E_2)(c_2 - c_1)$ 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

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

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 $X(b_{MH}, \{b_H\}) > 1$ 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 $U[c_2] + V_2 \leq b_{ML} - c_2 + \sum_{b \in \{b_{ML}, b_{MH}, b_H\}} X(b_{ML}, \{b\})[b - c_2]$, 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 $U[c_2] \leq b_{ML} - c_2 + X(b_{ML}, \{b_H\})[b_H - c_3] + \sum_{s \in \{b_{ML}, b_{MH}\}} X(b_{ML}, \{b\})[b - c_2]$.

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

$$\begin{aligned}\overline{U}[c_2] &= b_{ML} - T_t + \sum_{b \in \{b_{ML}, b_{MH}, b_H\}} X(b_{ML}, \{b\})[b - c_2] \\ &= b_{ML} - c_2 + X(b_{ML}, \{b_H\})[b_H - c_3] + \sum_{b \in \{b_{ML}, b_{MH}\}} X(b_{ML}, \{b\})[b - c_2], \qquad (13)\end{aligned}$$

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.[17]

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 $T_t = c_2 + X(b_L, E_3)(c_3 - c_2)$. Thus, the principal's payoff from making such an offer, conditional on the agent being type c_2, is

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

[17] 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.

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

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

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

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

Then, we have

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

Since $b_L < c_2$ by assumption, there exists $\Delta_2^1 > 0$ such that, if $c_2 - b_L > \Delta_2^1$, 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 $X(b_{MH},\{b_H\}) > 1$).

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

$T_t - c_1 = V_2^{(\sigma,\mu)}[h_t, b_t] + A_2^{(\sigma,\mu)}[h_t, b_t](c_2 - c_1)$. 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 $b_{t'} = b_{ML}$.[18] Therefore, $A_2^{(\sigma,\mu)}[h_t, b_t] \geq X(b_{ML}, \{b_{ML}\}) > 1$, where the last inequality follows from Assumption 1. The payoff that an agent of type c_2 obtains by accepting offer T_t at time t is bounded below by $T_t - c_2 = c_1 - c_2 + V_2^{(\sigma,\mu)}[h_t, b_t] + A_2^{(\sigma,\mu)}[h_t, b_t](c_2 - c_1) > V_2^{(\sigma,\mu)}[h_t, b_t]$, where the inequality follows since $A_2^{(\sigma,\mu)}[h_t, b_t] > 1$. 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 the principal makes an offer that only types in $\{c_1, c_2\}$ accept, or she makes an offer that all types reject.[19]

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 $T_t = c_2 + X(b_{ML}, \{b_H\})(c_3 - c_2)$. 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

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

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

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

[18] 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 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 $E_2 = \{b_{ML}, b_{MH}, b_H\}$.

[19] Note that, by Lemma 2, an agent with cost c_3 takes action $a = 0$ when $C[h_t] = \{c_1, c_2, c_3\}$ and $b_t = b_{ML} \notin E_3$.

where $\hat{t}$ denotes the next period that state b_L is realized.[20] 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 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 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

$$\begin{aligned} T_t - c_1 &= V_2^{(\sigma,\mu)}[h_t, b_L] + A_2^{(\sigma,\mu)}[h_t, b_L](c_2 - c_1) \\ &= X(b_L, \{b_H\})(c_3 - c_1) + \left[X(b_L, \{b_{ML}\}) + \mathbb{E}[\delta^{\check{t}-t} | b_t = b_L] X(b_{ML}, \{b_{MH}\})\right](c_2 - c_1) \end{aligned} \tag{14}$$

where the second equality follows since $V_2^{(\sigma,\mu)}[h_t, b_L] = A_2^{(\sigma,\mu)}[h_t, b_L](c_3 - c_2) = X(b_L, \{b_H\})(c_3 - c_2)$ 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 $\check{t}$.[21] 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

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

[20] 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 $\sum_{b \in E_i} X(b_L, \{b\})(b - c_i)$.

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

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

where the second line follows from substituting the transfer in (14). 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

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

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 $T = c_2 + X(b_{ML}, \{b_H\})(c_3 - c_2)$).

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

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

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

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

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

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 $T - c_2 = X(b_{ML}, \{b_H\})(c_3 - c_2)$. 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

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

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,

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

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 $b_t = b_L$. □

References

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

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.

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.

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.

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.

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*.

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.