Dynamic Mechanisms without Money

Yingni Guo* Johannes Hörner†

October 2017

Abstract

We analyze the optimal design of dynamic mechanisms in the absence of transfers. The designer uses future allocation decisions to elicit private information. Values evolve according to a two-state Markov chain. We solve for the optimal allocation mechanism. Unlike with transfers, efficiency decreases over time. In the long run, polarization obtains. A simple implementation is provided. The agent is endowed with a "quantified option," corresponding to the number of units he is entitled to claim in a row.

Keywords: Mechanism design. Principal-Agent. Token mechanisms.

JEL numbers: C73, D82.

1 Introduction

This paper is concerned with the dynamic allocation of resources when transfers are not allowed and the resources' optimal use is the private information of a strategic agent.

We seek the social choice mechanism that maximizes efficiency. Here, efficiency and implementability are understood to be Bayesian: the agent and society understand the probabilistic nature of uncertainty and update beliefs based on it. The societal decision not to allow money—for economic, physical, legal or ethical reasons—is assumed. Throughout, we assume that the good to be allocated is perishable.[1] Absent private information, the allocation problem is trivial: the good should be provided if and only if its value exceeds

*Department of Economics, Northwestern University, yingni.guo@northwestern.edu.

†Yale University, 30 Hillhouse Ave., New Haven, CT 06520, USA, johannes.horner@yale.edu.

[1] Many allocation decisions involve goods or services that are perishable, such as how a nurse or a worker should divide time; which patients should receive scarce medical resources (blood or treatments); or which investments and activities should be approved by a firm.

its cost.[2] However, in the presence of private information, and in the absence of transfers, linking future allocation decisions to current decisions is the only instrument available to elicit truthful information. Our goal is to understand this link.

In each round, the agent's value for a good is either high or low and privately known to him. The agent enjoys consumption regardless of the value, but supplying the good entails a fixed cost. Efficiency demands supplying it only when the value is high. The value evolves over time. It is autocorrelated: tomorrow's value is more likely to be high, the higher today's value is. We examine the full spectrum from independent to perfectly persistent value. Persistence not only widens the range of applications but also provides new insights into the optimum.

Our goal is to solve for the efficient policy. We cast this problem as an agency model, in which a principal with commitment power chooses when to supply the good, given the agent's reports. There is no statistical or hard evidence concerning the agent's value, even *ex post*. This setup allows us to focus on the strategic rather than statistical role of information.

The optimal mechanism is a *quantified option*, corresponding to the number of units that the agent is entitled to claim consecutively with no questions asked. This option is updated according to the agent's choice. Whenever he claims the unit, his option decreases by one. Whenever he does not, his option is augmented by a certain amount. The increment is selected such that the low-value agent is indifferent between consuming the unit today and consuming those incremental units, *if he were to cash in all his units as fast as possible, with the incremental units last*.

This mechanism results in simple dynamics. In an initial phase, the agent claims the good if and only if his value is high. In the long run, this option either drops to zero or becomes literally infinite. In the former case, the unit is never supplied again—in essence, the agent is terminated. In the latter case, the unit is always supplied—the agent is "tenured." Hence, an episode of efficient allocations leads to one of these two inefficient outcomes: polarization arises eventually, as the agent receives either his worst or best outcome forever.

The intuition for these dynamics is straightforward, although the proofs are somewhat involved. Whenever the value is high, the interests of the two parties are most aligned. There is no better time to supply the unit than in such a round, as the agent's future expected values cannot be higher. Whenever the value is low, the interests of the two parties are the least aligned. Granting entitlement to future units is a way for the principal to induce

[2] This is because the supply of the perishable good is taken as given. There is a considerable literature on the optimal ordering policy for perishable goods, beginning with Fries (1975).

truth-telling. The appropriate "currency" for the agent's option is the number of units that he can claim in a row. Controlling for impatience, the later the entitlement comes, the better it is for the principal, as the low-value agent's future expected values drift upward.

As mentioned above, the optimal policy is implemented by a calibrated quantified option. The updating of the option when the agent forgoes the unit depends on the number of units to which he is already entitled. For the low-value agent, each incremental unit (attached to the tail of his option) is especially valuable if his "budget" is already large, as his value is more likely to be high by the time he claims the incremental units. Hence, an agent with a small option must be granted more increments than an agent with a large option.

What is the impact of persistence? As values become more persistent, a low current value has a larger impact on the expected value in future rounds. Hence, the principal has to promise more in the future for the low-value agent to forego the current unit. Hence, persistence jeopardizes the principal's ability to use future allocations to induce truth-telling. Efficiency decreases with persistence.

Our findings contrast with static design with multiple units (*e.g.*, Jackson and Sonnenschein, 2007), as the optimal mechanism is not a (discounted) quota mechanism as commonly studied in the literature: the order of the sequence of reports matters. The total (discounted) number of units that the agent receives is not fixed. Depending on the sequence of reports, the agent might ultimately receive few or many units.[3] By backloading inefficiencies, our mechanism exploits the agent's ignorance regarding future values. Examining the continuous-time limit of our model, we show that our mechanism results in a convergence rate that is higher than under a quota mechanism. (See Lemma 8 in Online Appendix D.)

Backloading inefficiency contrasts with the outcome in dynamic mechanisms with transfers (Battaglini, 2005). Unlike models with transfers, here, efficiency decreases over time. As explained above, the allocation stops being efficient from some random time onward. Eventually, the agent is either granted the unit forever or never again. Hence, polarization is ineluctable, but immiseration is not. This long-run polarization differs from the long-run immiseration in principal-agent models with risk-aversion (Thomas and Worrall, 1990).

Allocation problems without transfers are plentiful. Our results inform best practices concerning how to implement algorithms to improve allocations. For instance, consider nurses who must decide whether to take seriously alerts triggered by patients. The opportunity cost is significant. Patients appreciate quality time with nurses irrespective of whether their

[3]This is not a bankruptcy mechanism (Radner, 1986) either, because the order of the report sequence matters.

condition necessitates it. This discrepancy produces a challenge with which every hospital must contend: ignore alarms and risk that a patient with a serious condition is not attended to, or heed all alarms and overwhelm the nurses. "Alarm fatigue" is a problem that health care must confront (see, *e.g.*, Sendelbach, 2012). We suggest the best approach for trading off the risks of neglecting a patient in need and attending to one who simply cries wolf.

Our results readily extend to the case in which there is moral hazard, in addition to adverse selection. For instance, the agent might be a borrower, who privately observes his credit needs, and the principal is a lender, who can grant the desired short-term loan. If granted a loan, the borrower can default or honor it—an observable choice. Such a standard model of lending (see, *e.g.*, Bulow and Rogoff, 1989) is readily accommodated. All that is necessary is an additional obedience constraint. We discuss this extension in Section 5.

Related Literature. Our work is closely related to the literature on dynamic mechanisms with transfers and on "linking incentive constraints."

Transfers: The obvious benchmark is Battaglini (2005),[4] who considers the same model as ours but allows for transfers. Another important difference is his focus on revenue maximization, a meaningless objective in the absence of transfers. Section 4.5 explains in detail the role of transfers in dynamic mechanism design.

Because of transfers, his results are diametrically opposed to ours. In Battaglini, efficiency improves over time. Here, efficiency decreases over time, with an asymptotic outcome that is at best the outcome of the static game. Krishna, Lopomo and Taylor (2013) provide an analysis of limited liability (although transfers are allowed) in a model closely related to that of Battaglini, and their findings suggest that transfers affect both the optimal contract and the dynamics.[5]

Linking Incentives: This refers to the notion that as the number of identical decision problems increases, linking them improves on each isolated problem. See Fang and Norman (2006) and Jackson and Sonnenschein (2007) for papers specifically devoted to this. (See also Radner, 1981; Rubinstein and Yaari, 1983.) Hortala-Vallve (2010) provides an analysis of the unavoidable inefficiencies that must be incurred away from the limit, and Cohn (2010) demonstrates the suboptimality of the commonly used quota mechanisms.

Unlike the bulk of this literature, we focus on the exactly optimal mechanism for a fixed discount factor. This allows us to identify results (backloading of inefficiency, for instance)

[4]See also Zhang (2012) for an exhaustive analysis of Battaglini's model, as well as Fu and Krishna (2017).

[5]Note that there is an important exception to the quasi-linearity commonly assumed in the dynamic mechanism design literature, namely, Garrett and Pavan (2015).

that need not hold for other asymptotically optimal mechanisms and to clarify the role of the dynamic structure. Discounting is not the issue; the key is the fact that the agent learns the value of the units as they arrive.

Frankel (2015) computes the exactly optimal mechanism in a related environment. While he allows for more than two types and actions, he restricts attention to types that are serially independent. More importantly, he assumes that the preferences of the agent are independent of the agent's type, which allows for a drastic simplification of the problem.

Dynamic Capital Budgeting: More generally, the notion that virtual budgets can be used as an intertemporal instrument to discipline agents with private information has appeared in several papers in economics, within the context of games and principal-agent models.

Within the context of games, Möbius (2001) is the first to suggest that using the difference in the number of favors granted (between two agents) as a yardstick to grant new favors is a simple but powerful way of sustaining cooperation in long-run relationships.[6] While his token mechanism is suboptimal, it has desirable properties: properly calibrated, it yields an efficient allocation as discounting vanishes. Hauser and Hopenhayn (2008) come the closest to solving for the optimal mechanism (within the class of PPE). They measure the optimality of simple budget rules (according to which each favor is weighted equally, independent of the history), showing that this rule might be too simple (the efficiency cost reaching up to 30% of surplus). Their analysis suggests that the best equilibrium shares many features with the mechanism in our one-player world: the incentive constraint binds, and the efficient policy is followed until it is incompatible with promise keeping. Our model can be viewed as a game with one-sided incomplete information in which the production cost is known. There are some important differences, however. First, our principal has commitment and hence is not tempted to act opportunistically. Second, our agent's information is persistent.[7] Li, Matouschek and Powell (2017) and Lipnowski and Ramos (2017) solve for the PPE in a model similar to our i.i.d. case. Li, Matouschek and Powell (2017) allow for public signals and demonstrate that better monitoring improves performance.

Some principal-agent models find that simple capital budgeting rules are exactly optimal in related models (*e.g.*, Malenko, 2016). Our results suggest how to extend such rules when values are persistent. Indeed, in the i.i.d. case, the policy admits a simple implementation in terms of a two-part tariff. Only through the persistent environment do we learn that it

[6]See also Athey and Bagwell (2001), Abdulkadiroğlu and Bagwell (2012) and Kalla (2010).

[7]This results in a technical difference: our limiting model in continuous time corresponds to the Markovian case in which flow values switch according to a Poisson process. In Hauser and Hopenhayn, the lump-sum value arrives according to a Poisson process, and the process is memoryless.

is more natural to interpret the budget as an entitlement for consecutive units rather than a budget with a "fixed" currency value.

More generally, it has long been recognized that allocation rights to other (or future) units can be used as a "currency" to elicit private information. Hylland and Zeckhauser (1979) explain how this can be viewed as a pseudo-market. Casella (2005) develops a similar idea within the context of voting rights. Miralles (2012) solves a two-unit version, with both values being privately known at the outset. Boleslavsky and Kelly (2014) solves for the optimal regulation in a two-period environment in which the firm privately learns its current compliance cost.

We show that, due to private information, it is necessary to distort the allocation: after some histories, the good is provided independent of the report; after others, the good is never provided again. In this sense, the scarcity of good provision is endogenously determined to elicit information. There is a large body of literature in operations research considering the case in which this scarcity is exogenous—there are only n opportunities to provide the good. The problem is to decide when to exercise these opportunities. Important early contributions include Derman, Lieberman and Ross (1972) and Albright (1977). Their analyses suggest a natural mechanism that can be applied in our environment: the agent owns a number of "tokens" and uses them whenever he pleases. Gershkov and Moldovanu (2010) consider a dynamic allocation problem related to Derman, Lieberman and Ross, in which agents possess private information regarding the value of the good. In their model, agents are myopic and the scarcity of the resource is exogenously assumed. In addition, transfers are allowed. They demonstrate that the optimal policy of Derman, Lieberman and Ross (which is very different from ours) can be implemented via appropriate transfers.

A related literature considers optimal stopping without transfers; see, in particular, Kováč, Krähmer and Tatur (2014). The difference reflects the nature of the good, namely, whether it is perishable or durable. When only one unit is desired, this is a stopping problem. With a perishable good, a decision must be made every round. As a result, incentives and the optimal mechanisms are different. In the stopping problem, the agent might forgo the current unit if the value is low and future prospects are good. This is not the case here—incentives to forgo the unit must be endogenously generated via promises. In the stopping problem, there is only one history that does not terminate the game. Here, policies differ not only in when the good is first provided but also thereafter.

Finally, while the motivations of the papers differ, the techniques for the i.i.d. case that we use borrow numerous ideas from Thomas and Worrall (1990), as we explain in Section 3.

Section 2 introduces the model. Section 3 solves the i.i.d. case, introducing many of the ideas of the paper, while Section 4 solves the general model and develops an implementation for the optimal mechanism. Section 5 discusses extensions.

2 The Model

Time $n = 0, 1, \ldots,$ is discrete, and the horizon is infinite. There are two parties, a principal (she) and an agent (he). In each round, the principal may supply a unit of a good at a cost $c > 0$. The agent's value (or *type*) during round n, v_n, is a random variable that takes value l or h. We assume that $0 < l < c < h$: supplying the good is efficient if and only if the value is high, but the agent's value from consumption is always positive.

The value follows a Markov chain. Namely, for all n,

$$\mathbf{P}[v_{n+1} = h \mid v_n = h] = 1 - \rho_h, \quad \mathbf{P}[v_{n+1} = l \mid v_n = l] = 1 - \rho_l.$$

Here $\rho_h, \rho_l \in [0, 1]$ are the probabilities that the value switches conditional on the current value. The invariant probability of h is $q := \rho_l/(\rho_h + \rho_l)$. For simplicity, we assume that the initial value is drawn according to the invariant distribution $\mathbf{P}[v_0 = h] = q$. The (unconditional) expected value of the good is denoted $\mu := \mathbf{E}[v] = qh + (1 - q)l$. We make no assumption regarding how μ compares to c.

We assume that values are positively correlated. That is, tomorrow's value is more likely to be h conditional on today's value being h than conditional on today's value being l.[8] Let $\kappa := 1 - \rho_h - \rho_l$ be a measure of the persistence of the Markov chain. We assume that $\kappa \geqslant 0$. Two cases of special interest occur when $\kappa = 1$ and $\kappa = 0$. The former corresponds to perfect persistence and the latter to independence, as analyzed in Section 3.

At the beginning of each round, the value is drawn, and only the agent is informed of it. The two parties are impatient and share a common discount factor $\delta < 1$.[9] To exclude trivialities, assume that $\delta > l/\mu$ and $\delta > 1/2$.

The supply decision during round n is denoted by $x_n \in \{0, 1\}$: $x_n = 1$ means that the good is supplied during round n. We assume that the principal internalizes both the cost of supplying the good and the value of providing it to the agent. Thus, given an infinite history

[8] This assumption is commonly adopted in the literature. We comment on the case in which values are negatively correlated at the end of Section 4.3.

[9] The common discount factor is important. However, because we view our principal as a social planner trading off the agent's utility with the social cost of providing the good as opposed to an actual player, it is natural to assume that her discount factor is equal to the agent's.

$\{v_n, x_n\}_{n=0}^{\infty}$, the principal's realized *payoff* is defined as follows:

$$(1-\delta)\sum_{n=0}^{\infty}\delta^n x_n(v_n - c).$$

The agent's realized *utility* is defined as follows:

$$(1-\delta)\sum_{n=0}^{\infty}\delta^n x_n v_n.$$

Throughout, the terms payoff and utility refer to the expectations of these values.

We comment on several important assumptions maintained throughout.

- There are no transfers. This is our point of departure from most of dynamic mechanism design. Furthermore, our objective is efficiency, not revenue maximization. With transfers, there is a trivial mechanism that achieves efficiency: supply the good if and only if the agent pays a fixed price in the range (l, h).

- There is no *ex post* signal regarding the realized value of the agent—the principal does not observe realized payoffs. For some applications, it might be more plausible to assume that a signal of the value obtains at the end of a round, independent of the supply decision. In some other applications, one might wish to assume that this signal occurs only if the good is supplied (*e.g.*, a firm discovers the productivity of a worker only if he is hired). Conversely, statistical evidence might occur only from not supplying the good if supplying it averts a risk (a patient calling for care or police calling for backup). See Li, Matouschek and Powell (2017) for an analysis (with public signals) in a related context. Presumably, the optimal mechanism differs according to the monitoring structure. Understanding what happens without *any* signal is a natural first step.

- The principal commits *ex ante* to a (possibly randomized) mechanism. This assumption brings our analysis closer to the literature on dynamic mechanism design and distinguishes it from the literature on chip mechanisms (as well as Li, Matouschek and Powell, 2017), which assumes no commitment on either side and solves for the (perfect public) equilibria of the game. In Section 5, we discuss how renegotiation-proofness affects our results.

- The good is perishable. Hence, previous supply choices affect neither feasible nor desirable future opportunities. If the good were durable and only one unit were demanded, the problem would be a stopping problem, as in Kováč, Krähmer and Tatur (2014).

Due to commitment, we focus on mechanisms in which the agent truthfully reports his type at every round and the principal commits to a (possibly randomized) supply decision as a function of this last report and of the entire history of reports. This is without loss of generality.

Formally, a direct mechanism or *policy* is a collection $(\mathbf{x}_n)_{n=0}^{\infty}$, with $\mathbf{x}_n : \{l, h\}^{n+1} \to [0, 1]$ mapping the agent's reports up to round n into a supply decision during round n.[10] Our definition exploits the fact that because preferences are time-separable, the policy may be considered independent of past realized supply decisions. A direct mechanism defines a decision problem for the agent who maximizes his utility. A reporting strategy is a collection $(\mathbf{m}_n)_{n=0}^{\infty}$, where $\mathbf{m}_n : \{l, h\}^n \times \{l, h\} \to \Delta(\{l, h\})$ maps previous reports and the value during round n into a report for that round.[11] The policy is *incentive compatible* if truth-telling (that is, reporting the current value truthfully, independent of past reports) is optimal.

Our first objective is to solve for the *optimal* policy that maximizes the principal's payoff subject to incentive compatibility. Second, we would like to find a simple indirect implementation. Finally, we explore how the agent's utility evolves under the optimal policy.

3 Independent Types

We begin with the simple case in which types are i.i.d. (*i.e.*, $\kappa = 0$). Section 4 is devoted to the general case $\kappa \geqslant 0$.

With independent types, it is well known that attention can be restricted to policies that are represented by a tuple of functions $p_h, p_l : [0, \mu] \to [0, 1]$, $U_h, U_l : [0, \mu] \to [0, \mu]$ mapping the agent's promised utility U into (i) the probabilities $p_h(U), p_l(U)$ of supplying the good in the current round given the current report and (ii) a continuation utility $U_h(U), U_l(U)$ starting from the next round.[12] These functions must be consistent in the sense that, given U, the probabilities of supplying the good and the continuation utilities yield U to the agent.

[10] For simplicity, we use the same symbols l, h for the possible agent reports as for the values of the good.

[11] Without loss of generality, we assume that this strategy does not depend on past values. Given past reports and the current value, the decision problem from round n onward does not depend on past values.

[12] The reader is referred to Spear and Srivastava (1987) and Thomas and Worrall (1990) for details. Not every policy can be represented in this fashion, as the principal does not need to treat two histories that lead to the same continuation utility identically. However, because they are equivalent from the agent's viewpoint, the principal's payoff must be maximized by some policy that does so.

This condition is called "promise keeping." We stress that U is the *ex ante* utility in a given round. It is computed before the agent's value is realized.

Because such a policy is Markovian with respect to the promised utility U, the principal's payoff is also a function of U only. Hence, solving for the optimal policy and the (principal's) value function $W : [0, \mu] \to \mathbf{R}$ amounts to a Markov decision problem. Given discounting, the optimality equation characterizes both the optimal policy and the value function. For any fixed $U \in [0, \mu]$, the optimality equation states the following:

$$W(U) = \sup_{\substack{p_h, p_l \in [0,1] \\ U_h, U_l \in [0,\mu]}} (1-\delta)\left(qp_h(h-c) + (1-q)p_l(l-c)\right) + \delta\left(qW(U_h) + (1-q)W(U_l)\right),$$

subject to incentive constraints that the agent reports his value truthfully and the promise-keeping constraint:

$$(1-\delta)p_h h + \delta U_h \geqslant (1-\delta)p_l h + \delta U_l, \tag{IC_h}$$

$$(1-\delta)p_l l + \delta U_l \geqslant (1-\delta)p_h l + \delta U_h, \tag{IC_l}$$

$$U = (1-\delta)\left(qp_h h + (1-q)p_l l\right) + \delta\left(qU_h + (1-q)U_l\right). \tag{PK}$$

3.1 Complete Information

We first consider the complete-information benchmark in which the agent's value is public information: that is, we solve the principal's problem without the incentive constraints but with the promise-keeping constraint. The principal chooses the supply probability during each round as a function of the realized values up to that round. The only constraint is that the agent's expected utility before round 0 has to be some fixed U. Since there is no private information, it is without loss to assume that p_h, p_l are constant over time.

If the principal supplies the good if and only if the value is high, then the agent's expected utility is qh. Hence, for any $U \neq qh$, the optimal policy cannot be efficient. To deliver $U < qh$, the principal scales down the probability with which the good is supplied when the value is high, maintaining $p_l = 0$. Similarly, for $U > qh$, the principal must supply the good with positive probability even when the value is low.[13] The lemma below follows readily.

[13]Given $U \neq qh$, there are many other nonconstant optimal policies. If $U < qh$, then efficiency only requires that the good is not supplied when the agent's type is low. The principal may vary the probability that the high type obtains the good across rounds, which results in a nonconstant optimal policy.

Lemma 1. *Under complete information, an optimal policy is*

$$(p_h, p_l) = \left(\frac{U}{qh}, 0\right) \text{ if } U \in [0, qh], \quad \text{and} \quad (p_h, p_l) = \left(1, \frac{U - qh}{(1-q)l}\right), \text{ if } U \in [qh, \mu].$$

The complete-information value function, denoted $\overline{W}$*, is*

$$\overline{W}(U) = \begin{cases} \frac{h-c}{h}U & \text{if } U \in [0, qh], \\ \frac{h-c}{h}qh + \frac{l-c}{l}(U - qh) & \text{if } U \in [qh, \mu]. \end{cases}$$

Hence, the optimal initial promise (which maximizes $\overline{W}$*) is* qh*.*

The value function $\overline{W}(U)$ is piecewise linear, as illustrated by the dashed line in Figure 1. It is an upper bound to the value function under incomplete information.

3.2 The Optimal Mechanism

We now solve for the optimal policy under incomplete information. We first provide an informal derivation of the solution, which follows from two observations (formally established below). First,

the efficient supply choice $(p_h, p_l) = (1, 0)$ *is made "as long as possible."*

To understand the qualification "as long as possible," note that if $U = 0$ (or $U = \mu$), promise keeping pins down the supply decision. The good cannot (or must) be supplied, independent of the reports. More generally, if $U \in [0, (1-\delta)qh)$, the principal cannot supply the good for certain if the value is high while satisfying promise keeping. In this utility range, the supply choice is as efficient as possible subject to the promise-keeping constraint. This implies that a high report leads to a continuation utility of 0, with the probability of the good being supplied adjusted accordingly. A symmetric argument applies if $U \in (\mu - (1-\delta)(1-q)l, \mu]$.

These two peripheral intervals vanish as δ goes to one and are ignored for the remainder of this discussion. For every other promised utility, we claim that it is optimal to make the "static" efficient supply choice. Intuitively, there is never a better time to redeem promised utility than when the value is high. During such rounds, the interests of the two parties are most aligned. Conversely, when the value is low, repaying the agent his promised utility is suboptimal, because tomorrow's value cannot be lower than today's.

As trivial as this observation may sound, it already implies that the dynamics of the inefficiencies must be quite different from those in models with transfers. Here, inefficiencies are backloaded.

Since the supply decision is efficient as long as possible, the low type must have a higher continuation utility than the high type does. The spread $U_l - U_h$ is chosen such that the low type is indifferent between reporting high and low:

The low type's incentive constraint always binds.

The reason for this is standard: the agent is risk neutral, and the principal's payoff is a concave function of U. (If the principal's payoff failed to be concave, she could offer the agent a fair lottery that would make her better off.) Concavity implies that there is no benefit from spreading continuation utilities U_h, U_l beyond what incentive compatibility requires.

Because we are left with two variables (U_h, U_l) and two constraints (PK and the binding IC_l), we can immediately solve for the continuation utilities. Algebra is not needed. The agent is always willing to report that his value is high, meaning that his utility can be computed *as if* he followed this reporting strategy, namely,

$$U = (1-\delta)\mu + \delta U_h, \quad \text{or} \quad U_h = \frac{U - (1-\delta)\mu}{\delta}.$$

Because U is a weighted average of U_h and $\mu \geqslant U$, it follows that $U_h \leqslant U$. Hence, promised utility always decreases after a high report.

The high type is unwilling to report a low value because he receives an incremental value $(1-\delta)(h-l)$ from obtaining the good relative to what would make him merely indifferent between the two reports. Hence, in expectation, the agent receives $q(1-\delta)(h-l)$ more than by claiming to be low. Thus, we let $\underline{U} := q(h-l)$ and obtain

$$U = (1-\delta)\underline{U} + \delta U_l, \quad \text{or} \quad U_l = \frac{U - (1-\delta)\underline{U}}{\delta}.$$

Because U is a weighted average of $\underline{U}$ and U_l, it follows that $U_l \leqslant U$ if and only if $U \leqslant \underline{U}$. In that case, even a low report leads to a decrease in the continuation utility, albeit a smaller decrease than after a high report. This is because the flow utility delivered by the efficient policy exceeds the annuity on the promised utility by so much that continuation utility must decrease, independent of the report.

The following theorem (proved in Appendix A; see appendices for all proofs) summarizes this discussion with the necessary adjustments on the peripheral intervals.

Theorem 1. *An optimal policy is*

$$p_h = \min\left\{1, \frac{U}{(1-\delta)\mu}\right\}, \quad p_l = \max\left\{0, 1 - \frac{\mu - U}{(1-\delta)l}\right\}.$$

Given these values of $(p_h(U), p_l(U))$, continuation utilities are

$$U_h = \frac{U - (1-\delta)p_h\mu}{\delta}, \quad U_l = \frac{U - (1-\delta)(p_l l + (p_h - p_l)\underline{U})}{\delta}.$$

This policy is uniquely optimal when $U > \underline{U}$.

We now turn to a discussion of the shape of the value function and the utility dynamics. When U is below $\underline{U}$, the continuation utility decreases after both reports and eventually drops to zero. The principal can deliver all the promised utility *only* when the agent's value is high. Hence, there is no efficiency loss due to incomplete information.

However, for $U \in (\underline{U}, \mu)$, the agent's utility will reach both 0 and μ with positive probabilities. Both inefficiencies—the good is not provided when the value is high and is provided when the value is low—occur with positive probabilities. Hence, W is strictly below $\overline{W}$. The solid curve in Figure 1 illustrates the value function W, which is strictly below $\overline{W}$ when $U \in (\underline{U}, \mu)$.

Finally, we observe that $W'(U)$ approaches the complete-information derivative $\overline{W}'(U)$ as U approaches μ. For U close to μ, promise keeping forces the principal to supply the good to the low-value agent with a high probability. For U arbitrarily close to μ, U_h is arbitrarily close to μ such that $W'(U)$ approaches $\overline{W}'(U)$.

We summarize the discussion in the following lemma.

Lemma 2. *The value function W is linear and equal to $\overline{W}$ on $[0, \underline{U}]$ and strictly concave on $[\underline{U}, \mu]$. Furthermore, it is continuously differentiable on $(0, \mu)$, with*

$$\lim_{U \downarrow 0} W'(U) = \frac{h-c}{h}, \quad \lim_{U \uparrow \mu} W'(U) = \frac{l-c}{l}.$$

Consider the following functional equation for W that we obtain from Theorem 1 (ignoring again the peripheral intervals for the sake of the discussion):

$$W(U) = (1-\delta)q(h-c) + \delta q W \underbrace{\left(\frac{U - (1-\delta)\mu}{\delta}\right)}_{U_h} + \delta(1-q) W \underbrace{\left(\frac{U - (1-\delta)\underline{U}}{\delta}\right)}_{U_l}.$$

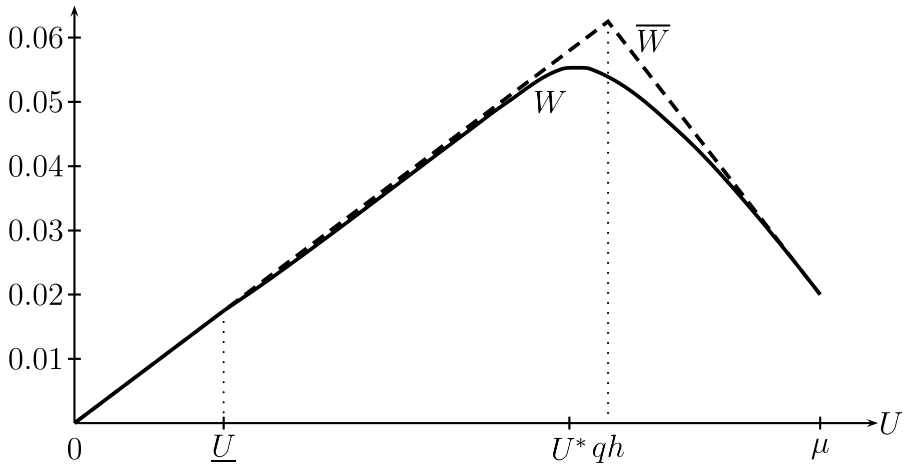

Figure 1: Value function for $(\delta, l, h, c, q) = (19/20, 2/5, 3/5, 1/2, 3/5)$.

Hence, taking for granted the differentiability of W stated in Lemma 2, we have

$$W'(U) = qW'(U_h) + (1-q)W'(U_l).$$

In probabilistic terms, $W'(U_n) = \mathbf{E}[W'(U_{n+1})]$ given the information at round n. This shows that W' is a bounded martingale, and thus, it converges.[14] This martingale is first mentioned in Thomas and Worrall (1990); we refer to it as the TW-martingale.

Because W is strictly concave on $(\underline{U}, \mu)$, but $U_h \neq U_l$ in this range, the convergence of $W'(U)$ implies that the process $\{U_n\}_{n=0}^{\infty}$ must eventually exit this interval. Since $U_h < U$ and $U_l \leqslant U$ on the interval $(0, \underline{U}]$, if we begin the process $\{U_n\}_{n=0}^{\infty}$ in $[0, \underline{U}]$, then the limit U_∞ must be 0. Therefore, U_n converges to either 0 or μ.

How likely is the process $\{U_n\}_{n=0}^{\infty}$ to converge to 0 or μ, given the optimal initial promise U^*? The answer is delivered by the TW-martingale. Indeed, U^* is characterized by $W'(U^*) = 0$. Hence,

$$0 = W'(U^*) = \mathbf{P}[U_\infty = 0 \mid U_0 = U^*]W'(0) + \mathbf{P}[U_\infty = \mu \mid U_0 = U^*]W'(\mu),$$

where $W'(0)$ and $W'(\mu)$ are the one-sided derivatives in Lemma 2. Hence,

$$\frac{\mathbf{P}[U_\infty = 0 \mid U_0 = U^*]}{\mathbf{P}[U_\infty = \mu \mid U_0 = U^*]} = -\frac{W'(\mu)}{W'(0)} = \frac{(c-l)/l}{(h-c)/h}. \tag{1}$$

[14] It is bounded because W is concave. Its derivative is bounded by its value at 0 and μ, given in Lemma 2.

We record this discussion in the next lemma.

Lemma 3. *The process* $\{U_n\}_{n=0}^{\infty}$ *(with* $U_0 = U^*$*) converges to* 0 *or* μ*, a.s., with probabilities given by* (1)*.*

The initial promise is set to yield this ratio of absorption probabilities. Remarkably, this ratio is independent of the discount factor and the invariant probability (although the step size of the random walk $\{U_n\}_{n=0}^{\infty}$ depends on δ and q). Both long-run outcomes 0 and μ are possible, irrespective of patience. Depending on the parameters, U^* can be above or below qh, the complete-information initial promise, as is easily verified. In Lemma 5 in Appendix A, we show that U^* decreases in c. Intuitively, this is because the random walk $\{U_n\}_{n=0}^{\infty}$ depends on c only through the choice of U^*, and it converges to 0 more often as c increases.

4 Persistent Types

We now return to the general model in which types are persistent rather than independent. As a prelminary, consider the case of perfect persistence $\rho_h = \rho_l = 0$. In that case, future allocations cannot be used as instruments to elicit truth-telling. The problem is equivalent to the static problem for which the solution is to either always provide the good (if $\mu \geqslant c$) or never do so.

This makes the role of persistence not entirely obvious. Because current types assign different probabilities of being a high type tomorrow, one might hope that tying promised future utility to current reports facilitates truth-telling. However, the case of perfect persistence also shows that correlation diminishes the scope for using future allocations as "transfers."

The techniques that served us well with independent types are no longer useful. Martingale techniques no longer allow us to solve for the absorption probabilities in closed-form, as in the i.i.d. case. In addition, *ex ante* utility is no longer a valid state variable. To understand why, note that with independent types, an agent of a given type can evaluate his utility based on only his current type, the probability of supply, and the promised continuation utility. However, if today's type is correlated with tomorrow's type, the agent cannot evaluate his continuation utility without knowing how the principal intends to implement it.

However, conditional on the agent's type tomorrow, his type today carries no information on future types, by the Markovian assumption. Hence, tomorrow's promised *interim* utilities suffice for the agent to compute his utility today regardless of whether he deviates. Of course, his type tomorrow is not directly observable. Therefore, we must use the utility he receives from tomorrow's report (assuming he tells the truth). That is, we must specify his promised utility tomorrow conditional on each possible report at that time.

This creates no difficulty in terms of his truth-telling incentives tomorrow: because the agent truthfully reports his type on path, he also does so after having lied in the previous round (conditional on his current type and his previous report, his previous type does not enter his decision problem). The one-shot deviation principle holds: when the agent considers lying now, there is no loss in assuming that he reports truthfully tomorrow.

Plainly, we are not the first to note that, with persistence, the appropriate state variables are the interim utilities. See Townsend (1982), Fernandes and Phelan (2000), Cole and Kocherlakota (2001), Doepke and Townsend (2006) and Zhang and Zenios (2008). However, this remains insufficient to evaluate the principal's payoff and apply dynamic programming to this problem. We must also specify the principal's belief. Let ϕ denote the probability that she assigns to the high type. This probability can take only three values, depending on whether this is the initial round or whether the last report was high or low. Nonetheless, it is just as convenient to treat ϕ as an arbitrary element in the unit interval.

4.1 The Program

As discussed above, the principal's optimization program requires three state variables: the belief of the principal $\phi = \mathbf{P}[v = h] \in [0, 1]$ and the pair of interim utilities $U = (U(h), U(l))$ that the principal delivers as a function of the current report.

We let $V \in \mathbf{R}^2$ denote the set of interim utility pairs that are supported by an incentive-compatible policy. (We formally define and characterize V in Section 4.2.) Let μ_h be the largest utility that can be given to an agent whose current type is high. This utility is delivered by always supplying the good. Similarly, μ_l is the largest utility that can be given to a current low type if the good is always supplied.[15] It follows that V is a subset of $[0, \mu_h] \times [0, \mu_l]$.

A policy is now a pair of functions $p_h, p_l : V \to [0, 1]$ mapping the current utility vector $U \in V$ onto the probability with which the good is supplied as a function of the report and a pair, $U_h, U_l : V \to V$, mapping U onto the continuation interim utilities, which are $U_h = (U_h(h), U_h(l))$ if the report is h and $U_l = (U_l(h), U_l(l))$ if the report is l. Here, $U_v(v')$

[15] The utility pair (μ_h, μ_l) solves

$$\mu_h = (1 - \delta)h + \delta(1 - \rho_h)\mu_h + \delta\rho_h\mu_l, \quad \mu_l = (1 - \delta)l + \delta(1 - \rho_l)\mu_l + \delta\rho_l\mu_h.$$

is the continuation utility if today's report is v and tomorrow's is v'.

Given discounting, the optimality equation characterizes both the optimal policy and the value function. For any $U \in V$ and $\phi \in [0,1]$, the optimality equation states the following:

$$\begin{aligned} W(U,\phi) &= \sup \{\phi\left((1-\delta)p_h(h-c)+\delta W(U_h, 1-\rho_h)\right) \\ &+ (1-\phi)\left((1-\delta)p_l(l-c)+\delta W(U_l,\rho_l)\right)\}, \end{aligned}$$

over $p_h, p_l \in [0,1]$ and $U_h, U_l \in V$, subject to the promise-keeping and incentive-compatibility constraints:

$$\begin{aligned} U(h) &= (1-\delta)p_h h + \delta\left\{(1-\rho_h)U_h(h) + \rho_h U_h(l)\right\} && (PK_h) \\ &\geqslant (1-\delta)p_l h + \delta\left\{(1-\rho_h)U_l(h) + \rho_h U_l(l)\right\}, && (IC_h) \end{aligned}$$

and

$$\begin{aligned} U(l) &= (1-\delta)p_l l + \delta\left\{(1-\rho_l)U_l(l) + \rho_l U_l(h)\right\} && (PK_l) \\ &\geqslant (1-\delta)p_h l + \delta\left\{(1-\rho_l)U_h(l) + \rho_l U_h(h)\right\}. && (IC_l) \end{aligned}$$

Note that W is concave on V (by the linearity of the constraints on the utilities).[16]

4.2 Incentive-Feasible Utility Pairs

One difficulty in using interim utilities as state variables is that the dimensionality of the problem increases with the cardinality of the type space. A related difficulty is that it is not obvious which vectors of utilities are feasible given the incentive constraints. For instance, promising to assign all future units to the agent in the event that his current report is high while assigning none if the current report is low is simply not incentive compatible.

Our first step is to solve for the feasible, incentive-compatible (in short, incentive-feasible) utility pairs. These are interim utilities for which there are allocation probabilities and promised utility pairs tomorrow that make truth-telling incentive compatible. Moreover, these promised utility pairs tomorrow are also incentive feasible.

Definition 1. *The incentive-feasible set, $V \in \mathbf{R}^2$, is the set of interim utilities in round 0*

[16] The policy choice (p_h, p_l, U_h, U_l) is independent of ϕ because we can decompose the optimization problem into two sub-problems: (i) choose p_h, U_h to maximize $(1-\delta)p_h(h-c)+\delta W(U_h, 1-\rho_h)$ subject to PK_h and IC_l; (ii) choose p_l, U_l to maximize $(1-\delta)p_l(l-c)+\delta W(U_l,\rho_l)$ subject to PK_l and IC_h. This independence property applies to any finite-type environment.

that are obtained for some incentive-compatible policy.

It is standard to show that V is the largest bounded set such that for each $U \in V$, there exist $p_h, p_l \in [0,1]$ and two pairs $U_h, U_l \in V$ satisfying (PK_h)–(IC_l).[17] To obtain some intuition regarding V's structure, we enumerate some of its elements. Clearly, $0 \in V$ and $(\mu_h, \mu_l) \in V$. It suffices to never or always supply the unit, independent of the reports. More generally, for any integer $\nu \geqslant 0$, the principal can supply the unit for the first ν rounds, independent of the reports, and never supply the unit thereafter. We refer to such policies as pure *frontloading* policies because they deliver a given number of units as quickly as possible. Similarly, a pure *backloading* policy does not supply the unit for the first ν rounds but does so afterward, independent of the reports. A (mixed) frontloading policy is one that randomizes over two pure frontloading policies with consecutive integers. A (mixed) backloading policy is defined similarly.

The utility pairs corresponding to frontloading and backloading policies are immediate to define in parametric forms. Given $\nu \in \mathbf{N}$, let

$$f_h^\nu = \mathbf{E}\left[(1-\delta)\sum_{n=0}^{\nu-1}\delta^n v_n \mid v_0 = h\right], \quad f_l^\nu = \mathbf{E}\left[(1-\delta)\sum_{n=0}^{\nu-1}\delta^n v_n \mid v_0 = l\right], \tag{2}$$

and set $f^\nu := (f_h^\nu, f_l^\nu)$. This is the utility pair when the principal supplies the unit for the first ν rounds, independent of the reports. Second, for $\nu \in \mathbf{N}$, let

$$b_h^\nu = \mathbf{E}\left[(1-\delta)\sum_{n=\nu}^{\infty}\delta^n v_n \mid v_0 = h\right], \quad b_l^\nu = \mathbf{E}\left[(1-\delta)\sum_{n=\nu}^{\infty}\delta^n v_n \mid v_0 = l\right], \tag{3}$$

and set $b^\nu := (b_h^\nu, b_l^\nu)$.[18] This is the pair when the principal supplies the unit only from round ν onward. The sequence f^ν is increasing (in both arguments) as ν increases, with $f^0 = 0$ and $\lim_{\nu\to\infty} f^\nu = (\mu_h, \mu_l)$. On the contrary, b^ν is decreasing, with $b^0 = (\mu_h, \mu_l)$ and $\lim_{\nu\to\infty} b^\nu = 0$.

Fix a frontloading and a backloading policy such that the high type is indifferent between the two. Then, the low type prefers the backloading policy, as the low type's expected value for a unit in a given round ν increases with ν. For a fixed high type's utility, not only is

[17] Incentive feasibility is closely related to self-generation (see Abreu, Pearce and Stacchetti, 1990), although it pertains to the different types of a single agent rather than to different players. The distinction is not merely a matter of interpretation because a high type can become a low type and vice versa, which represents a situation with no analogue in repeated games. Nonetheless, the proof of this characterization is similar.

[18] Here and in Section B.1, we omit the obvious corresponding analytic expressions. See Appendix B.

backloading better than frontloading for the low type, but these policies also yield the highest and lowest utilities to the low type. In other words, these policies span the incentive-feasible set, as the next lemma establishes.

Lemma 4. *It holds that*

$$V = \mathrm{co}\{b^\nu, f^\nu : \nu \in \mathbf{N}\}.$$

That is, V is a polygon with a countable infinity of vertices (and two accumulation points). See Figure 2 for an illustration. We now discuss certain properties of V. Readers who are more interested in the optimal mechanism and its implementation may skip this part and proceed to Section 4.3.

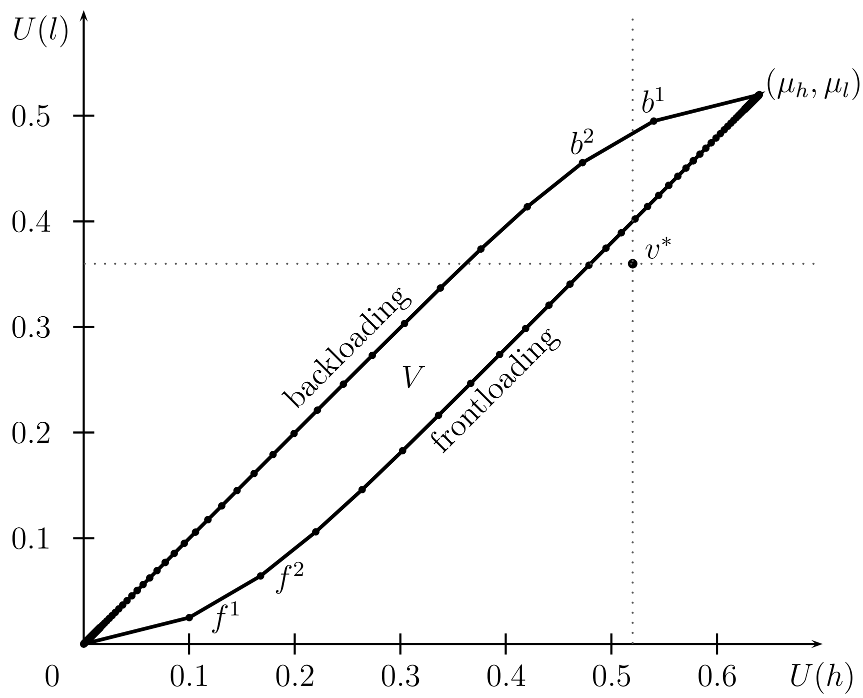

Figure 2: Incentive-feasible set V for $(\delta, \rho_h, \rho_l, l, h) = (9/10, 1/3, 1/4, 1/4, 1)$.

It is readily verified that

$$\lim_{\nu\to\infty} \frac{f_l^{\nu+1} - f_l^\nu}{f_h^{\nu+1} - f_h^\nu} = \lim_{\nu\to\infty} \frac{b_l^\nu - b_l^{\nu+1}}{b_h^\nu - b_h^{\nu+1}} = 1.$$

When the switching time ν is large, the change in the agent's utility from increasing this time is largely unaffected by his initial type. Hence, the slopes of the boundaries are less than and approach one as ν goes to ∞.

If the principal could see the agent's value, then she would supply the good if and only if the value is high. The agent's interim utility pair is then denoted by $v^* = (v_h^*, v_l^*)$, called the first-best utility.[19] It is easily verified that $(\mu_l - v_l^*)/(\mu_h - v_h^*) > 1$, and thus, the first-best utility v^* is outside V. Due to private information, the low type derives information rents: if the high type's utility is the first-best level, v_h^*, any incentive-compatible policy has to grant the low type strictly more than v_l^*.

Importantly, front- and backloading policies are not the only policies achieving boundary utilities. The frontloading boundary corresponds to policies that assign the maximum probability of supplying the good for high reports, while promising continuation utilities lying on the frontloading boundary for which IC_l bind. The backloading boundary corresponds to policies assigning the minimum probability of supplying the good for low reports, while promising continuation utilities on the backloading boundary for which IC_h bind. Front- and backloading policies are representative examples within each class.

Persistence affects the set V as follows. When values are i.i.d., the low type values the unit in round $\nu \geqslant 1$ the same as the high type does. Round 0 is the only exception. As a result, the vertices $\{f^\nu\}_{\nu=1}^\infty$ (or $\{b^\nu\}_{\nu=1}^\infty$) are aligned and V is a parallelogram with vertices $0, (\mu_h, \mu_l), f^1$ and b^1. As persistence increases, the imbalance between different types' utilities increases. The set V flattens. With perfect persistence, V becomes a straight line: the low type no longer cares about frontloading versus backloading, as no amount of time allows his type to change. Figure 3 illustrates the shape of V as persistence varies.

4.3 The Optimal Mechanism and Implementation

Not every incentive-feasible utility vector arises under the optimal policy. Irrespective of the sequence of reports, some vectors are never visited. While dynamic programming requires solving for the value function and the optimal policy on the entire domain V we focus on the subset of V that is relevant given the optimal initial promise and the resulting dynamics. We relegate the discussion of the entire set V to Appendix B.1.

This relevant subset is the frontloading boundary—the polygonal chain spanned by pure frontloading. Two observations from the i.i.d. case remain valid. First, the efficient choice is made as long as possible; second, promises are chosen such that the low type is indifferent

[19]The utility pair (v_h^*, v_l^*) solves

$$v_h^* = (1-\delta)h + \delta(1-\rho_h)v_h^* + \delta\rho_h v_l^*, \quad v_l^* = \delta(1-\rho_l)v_l^* + \delta\rho_l v_h^*.$$

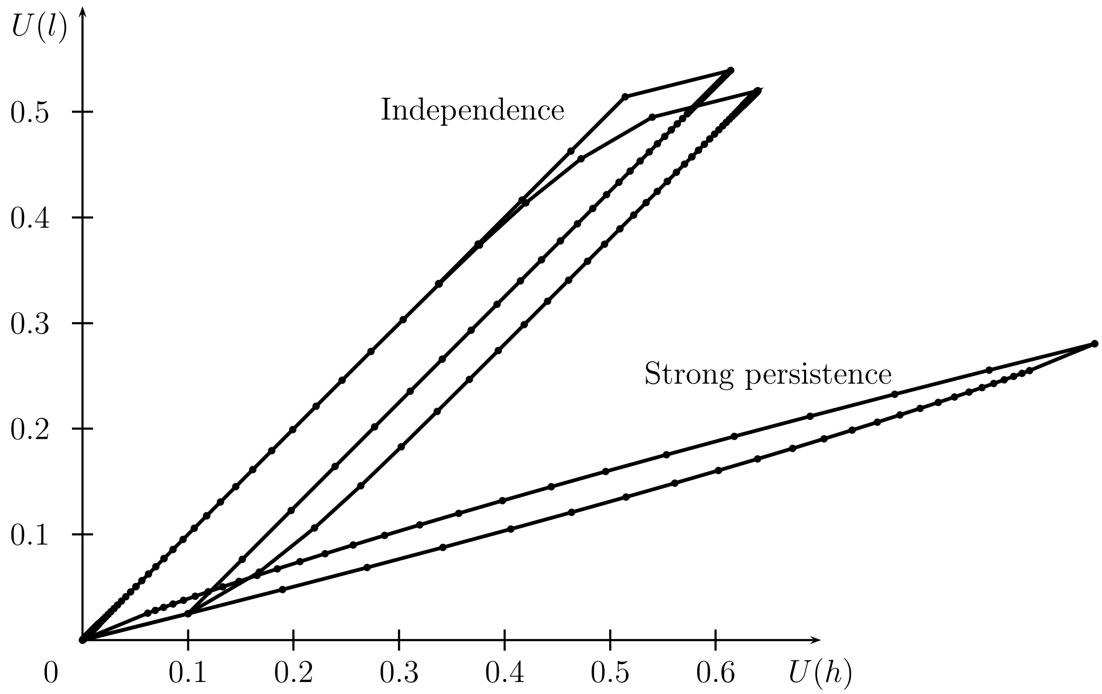

Figure 3: Impact of persistence, as measured by $\kappa \geqslant 0$.

between reporting high and low. To understand why such a policy yields utilities on the frontloading boundary (as mentioned in the second to last paragraph in Section 4.2), note that because the low type is indifferent between the two reports, the agent is willing to report a high value irrespective of his type. Because the good is then supplied, the agent's utilities can be computed as if frontloading prevailed.

From the principal's perspective, however, it matters that frontloading does not actually take place. The principal's payoff is higher under the optimal policy than under any frontloading policy, which simply supplies the good for some rounds without eliciting any information from the agent. The optimal policy implements efficiency as long as possible.

Hence, after a high report, continuation utility declines.[20] Specifically, U_h is computed as under frontloading as the solution to the following system, given U:

$$U(v) = (1-\delta)v + \delta \mathbf{E}[U_h \mid v], \quad v = l, h.$$

Here, $\mathbf{E}[U_h \mid v]$ is the expectation of the utility vector U_h provided that the current type is

[20]Because the frontloading boundary is upward sloping, the interim utilities of both types vary in the same way. Accordingly, we use terms such as "higher" or "lower" utility and write $U < U'$ for the component-wise order.

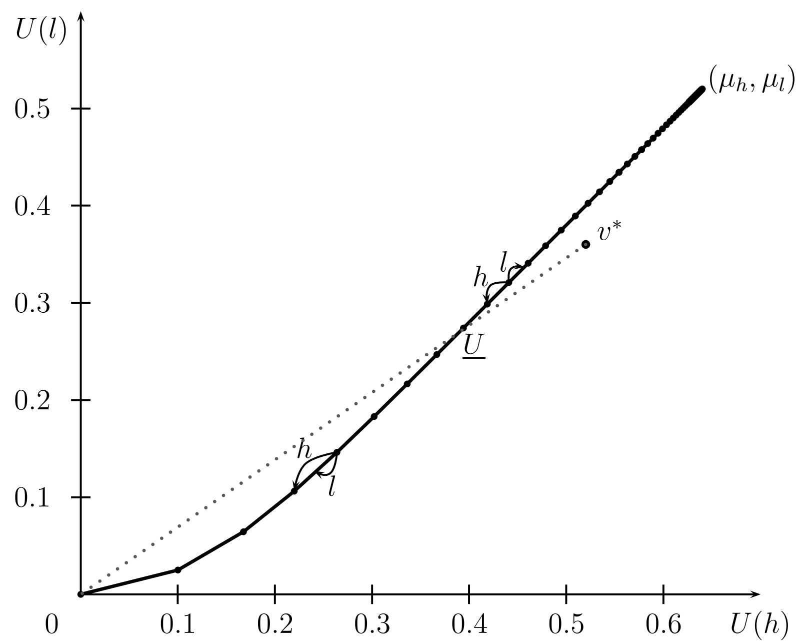

Figure 4: Dynamics of utility on the frontloading boundary.

v (*e.g.*, for $v = h$, $\mathbf{E}[U_h \mid v] = \rho_h U_h(l) + (1 - \rho_h) U_h(h)$).

The promised U_l does not admit such an explicit formula because it is pinned down by IC_l and the requirement that it lies on the frontloading boundary. In fact, U_l might be lower or higher than U (see Figure 4) depending on U. If U is high enough, U_l is higher than U; conversely, under certain conditions, U_l is lower than U if U is low enough. This condition has a simple geometric representation: let $\underline{U}$ denote the intersection of the half-open line segment $(0, v^*]$ with the frontloading boundary (see Figure 4).[21] Then, U_l is lower than U if and only if U lies below $\underline{U}$.[22] (If there is no such intersection, then U_l is always higher than U.) This intersection exists if and only if

$$\delta \rho_l (h - l) > l(1 - \delta), \tag{4}$$

which is satisfied when ρ_l is sufficiently high, that is, if the low-type persistence is sufficiently low. In this case, even a previous low type expects to have a high value today sufficiently often that, if U is low enough, then the continuation utility declines even after a low report. When the low-type persistence is high, this does not occur.[23] As in the i.i.d. case, the

[21] This line has the equation $U(l) = \frac{\delta \rho_l}{1 - \delta(1 - \rho_l)} U(h)$.

[22] With some abuse, we write $\underline{U} \in \mathbf{R}^2$ because it is the natural extension of the variable $\underline{U} \in \mathbf{R}$ introduced in Section 3. We set $\underline{U} = 0$ if the intersection does not exist.

[23] The condition (4) is satisfied in the i.i.d. case due to our assumption that $\delta > l/\mu$.

principal achieves the complete-information payoff if and only if $U \leqslant \underline{U}$ (or $U = (\mu_h, \mu_l)$). We summarize this discussion with the following theorem.

Theorem 2. *The optimal policy consists of the constrained-efficient policy*

$$p_h = \min\left\{1, \frac{U(h)}{(1-\delta)h}\right\}, \quad p_l = \max\left\{0, 1 - \frac{\mu_l - U(l)}{(1-\delta)l}\right\},$$

in addition to (i) choices (U_h, U_l) on the frontloading boundary of V such that IC_l always binds and (ii) a (specific) initial promise $U_0 > \underline{U}$ on this frontloading boundary.

While the implementation in the i.i.d. case can be described in terms of a "utility budget," inspired by the use of *ex ante* utility as a state variable, the analysis of the Markov case strongly suggests the use of a more concrete metric—the number of units that the agent is entitled to claim in a row with "no questions asked." The vectors on the frontloading boundary are parameterized by the number of rounds required to reach zero under frontloading. We denote such a policy by a number, $z \geqslant 0$, with the interpretation that the good is supplied for the first $\lfloor z \rfloor$ rounds and with probability $z - \lfloor z \rfloor$ also during round $\lfloor z \rfloor$. (Here, $\lfloor z \rfloor$ denotes the integer part of z.) The corresponding utility pair is written as $(U(h,z), U(l,z))$ such that

$$U(h,z) = \mathbf{E}\left[(1-\delta)\sum_{n=0}^{\lfloor z \rfloor - 1} \delta^n v_n \mid v_0 = h\right] + (z - \lfloor z \rfloor)\mathbf{E}\left[(1-\delta)\delta^{\lfloor z \rfloor} v_{\lfloor z \rfloor} \mid v_0 = h\right],$$

$$U(l,z) = \mathbf{E}\left[(1-\delta)\sum_{n=0}^{\lfloor z \rfloor - 1} \delta^n v_n \mid v_0 = l\right] + (z - \lfloor z \rfloor)\mathbf{E}\left[(1-\delta)\delta^{\lfloor z \rfloor} v_{\lfloor z \rfloor} \mid v_0 = l\right].$$

If $z = \infty$, the good is always supplied, yielding the utility vector (μ_h, μ_l).

We may think of the optimal policy as follows. During a given round n, the agent is promised z_n. If the agent asks for the unit (and this is feasible, *i.e.*, $z_n \geqslant 1$), the next promise $z_{n+1,h}$ equals $z_n - 1$. It is easy to verify that the following holds for both $v = l, h$:

$$U(v, z_n) = (1-\delta)v + \delta\mathbf{E}\left[U(v_{n+1}, z_n - 1) \mid v_n = v\right]. \tag{5}$$

If $z_n < 1$ and the agent asks for the unit, he receives the unit with probability z_n and obtains a continuation utility of 0.

If the agent claims to be low, this leads to the revised promise $z_{n+1,l}$ such that

$$U(l, z_n) = \delta \mathbf{E} \left[U(v_{n+1}, z_{n+1,l}) \mid v_n = l \right], \tag{6}$$

provided that there exists a (finite) $z_{n+1,l}$ that solves this equation.[24] Combining (5) and (6), we obtain the following:

$$(1 - \delta)l = \delta \mathbf{E} \left[U(v_{n+1}, z_{n+1,l}) - U(v_{n+1}, z_n - 1) \mid v_n = l \right].$$

Therefore, the promise after the low report $z_{n+1,l}$ is chosen such that the low type is indifferent between consuming the current unit and consuming the units between $z_n - 1$ and $z_{n+1,l}$.

Whenever the agent forgoes a unit, the principal compensates him through extra units at the tail of his entitled string. Adding units at the tail saves on the extra units needed, because the current low type becomes increasingly irrelevant for the expected value toward the tail. This implies that the higher the agent's current promise is, or the lower the persistence is, the fewer units are needed.

Given any initial choice of $U_0 > \underline{U}$, finitely many consecutive reports of l or h suffice for the promised utility to reach (μ_h, μ_l) or 0. By the Borel-Cantelli lemma, this implies that absorption occurs almost surely. Starting from the optimal initial promise, both long-run outcomes have strictly positive probability. However, unlike in the i.i.d. case, we cannot solve for the absorption probabilities beginning from the optimal initial promise in closed form.

4.4 The Role of Persistence

Because of the discrete nature of the framework, which prevents us from obtaining closed-form solutions, it is difficult to formally establish comparative statics.[25] However, it is straightforward to obtain them numerically.

We start with persistence. While there is no commonly agreed metric of persistence, it is clear that, comparing two Markov chains $\{v_n\}_{n=0}^{\infty}$ and $\{v'_n\}_{n=0}^{\infty}$ with transition matrices P and P', if it is the case that $P' = P^{\omega}$ for some integer $\omega > 1$, then the chain $\{v_n\}_{n=0}^{\infty}$ is more

[24]This is impossible if the promised z_n is too large. In that case, the good is provided with the probability $\tilde{q}$ that solves $U(l, z_n) - \tilde{q}(1 - \delta)l = \delta \mathbf{E} \left[U(v_{n+1}, \infty) \mid v_n = l \right]$.

[25]The entire analysis can also be performed in continuous time, using a Poisson formulation. This formulation is available in Section D of the Online Appendix. While this formulation is somewhat more technical, it has the benefit of yielding closed-form solutions both for the incentive-feasible set and the optimal value function.

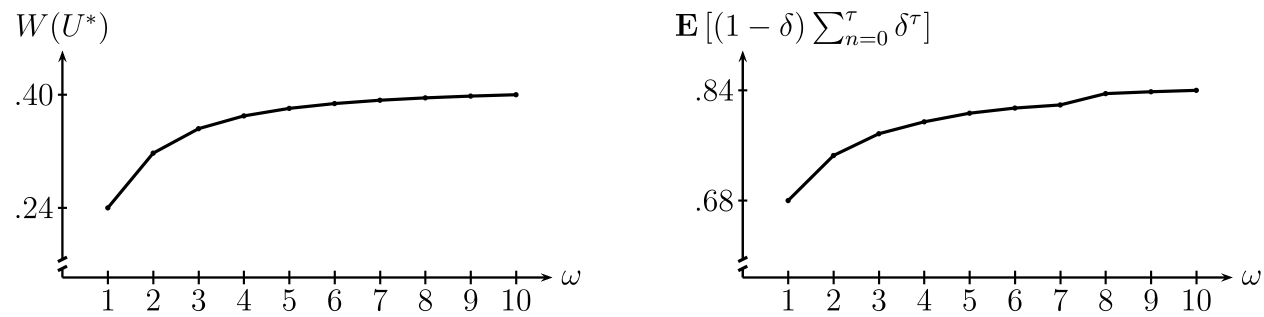

Figure 5: The value and the (discounted) time until absorption for $(\delta, l, h, c) = (.7, 1, 3, 2)$. The transition matrix is $((.9, .1), (.1, .9))^{\omega}$. The absorption time is denoted by τ.

persistent than $\{v'_n\}_{n=0}^{\infty}$, as the latter can be interpreted as the same process sampled less frequently. The i.i.d. case obtains in the limit, as $\omega = \infty$.[26]

As is plain from Figure 5 below, persistence hurts the principal. (The left-hand side shows how the value of $W(U^*)$ increases as ω increases.) It is not entirely obvious that persistence hurts the principal, as persistence means that types are correlated over time, giving the principal the ability to "cross-check" reports across rounds. However, with more persistent values, the low type has less to gain from future units, as values in the near future are also likely to be low. As a result, the principal must promise more of these future units to induce truth-telling, leading to faster absorption into one of the inefficient eventual outcomes.[27] On the right-hand side of Figure 5, we depict the expected (discounted) time until the process $\{U_n\}_{n=0}^{\infty}$ is absorbed at 0 or (μ_h, μ_l). This expected time until absorption measures how long the relationship stays in the efficient phase. Figure 5 shows that the less persistent the agent's value is, the longer the efficient allocation is implemented.

The impact of persistence on the likelihood of each of these eventual outcomes and on the agent's overall utility is ambiguious. Martingale methods do not allow a simple formula as in the i.i.d. case (see (1)), and indeed, unlike in the i.i.d. case, both the discount factor and the invariant distribution affect this likelihood, as is readily verified via examples. When values are highly persistent, the problem is close to the static one, in which the utility of the agent is either 0 or μ, according to whether granting the unit indiscriminately exceeds the

[26] That is, provided the chain is irreducible and aperiodic.

[27] Nonetheless, it is not persistence *per se* but positive correlation that is detrimental. It is tempting to think that any type of persistence is bad because the agent has private information that pertains not only to today's value but also to tomorrow's, and eliciting private information is often costly. However, with perfectly negatively correlated types, the complete-information payoff is easily achieved: offer the agent a choice between receiving the good in all odd or all even rounds. As $\delta > l/h$ (we assumed that $\delta > l/\mu$), truth-telling is optimal. Just as in the case of a lower discount rate, a more negative correlation (or less positive correlation) makes future promises more effective incentives because preference misalignment is short-lived.

cost of doing so. Hence, persistence increases or decreases the likelihood of absorption at 0, as well as the expected utility of the agent, depending on payoff parameters.

We have omitted a discussion of other comparative statics. Indeed, they align with intuition: the value $W(U^*)$ increases with the discount factor δ, with the probability ρ_l, as well as with the high value h. However, the value is not monotone in l due to the two counteracting forces. As l increases, it becomes more difficult to induce truth-telling from the low type. By contrast, the principal has less to lose from serving a low type. The impact on the value can be in either direction. (It is easy to analytically show that the value decreases in c because the cost has no impact on the agent's incentives.)

4.5 A Comparison with Dynamic Mechanisms with Transfers

In this section, we compare our results with those in the literature on dynamic mechanisms with transfers by discussing the difference between our results and those in Battaglini (2005). Our model differs from Battaglini (2005), but the comparison delineates both the role of transfers and the differences between the two problems.

The difference in results is striking. One of the main findings of Battaglini, "no distortion at the top," has no counterpart here. With transfers, efficient provision occurs *forever* once the agent first reports a high type. Further, even along the history in which efficiency is not achieved in finite time (namely, in his model, after an uninterrupted string of low reports), efficiency is asymptotically approached. As explained above, we eventually obtain (with probability one) an inefficient outcome, which can be implemented without further reports. Moreover, both long-run outcomes can arise. In summary, with transfers, inefficiencies are frontloaded to the greatest extent possible. Without transfers, they are backloaded.

The difference can be understood as follows. First, and importantly, Battaglini's results rely on revenue maximization being the objective. With transfers, efficiency is trivial: simply charge the cost whenever the good is supplied. When revenue is maximized, transfers reverse the incentive constraints: it is no longer the low type who would like to mimic the high type but the high type who would like to avoid paying his entire value for the good by claiming that his type is low. The high type's incentive constraint binds, and he must be given information rents. Ideally, the principal would like to charge for these rents *before* the agent has private information, when the expected value of these rents to the agent is still common knowledge. When types are i.i.d., this poses no difficulty, and these rents can be expropriated one round ahead of time. With correlation, however, different types value these rents differently, as their likelihood of being high in the future depends on their current types. However, when considering information rents sufficiently far in the future, the initial type exerts a minimal effect on the expected value of these rents. Hence, the value can "almost" be extracted. As a result, it is in the principal's best interest to maximize the surplus and offer a nearly efficient contract for all dates that are sufficiently far away.

Money plays two roles. First, because it is an instrument that allows promises to "clear" on the spot without allocative distortions, it prevents the occurrence of backloaded inefficiencies—a poor substitute for money in this regard. Even if payments could not be made "in advance," this would suffice to restore efficiency if that were the goal. Another role of money is that it allows value to be transferred before information becomes asymmetric. Hence, information rents no longer impede efficiency, at least with respect to the remote future. These future inefficiencies are eliminated altogether.

A plausible intermediate case arises when money is available but the agent is protected by limited liability, meaning that payments can only be made from the principal to the agent. The principal maximizes social surplus net of any payments.[28] In this case, we show in Appendix C.2 (see Lemma 7) that no transfers are made if (and only if) $c - l < l$. This condition can be interpreted as follows: $c - l$ is the cost to the principal of incurring one round of inefficiency (supplying the good when the type is low), whereas l is the cost to the agent of forgoing a low-value unit. Hence, if it is costlier to buy off the agent than to supply the good when the value is low, the principal still prefers to follow the optimal policy without money.

5 Discussion

Here, we discuss some extensions that have not yet been discussed.

Moral hazard. As discussed in Introduction, the model can be extended to incorporate moral hazard. Consider the following model of lending: the agent (the borrower) reports his type $v \in \{h, l\}$. The principal (the lender) decides whether to grant the desired short-term loan. If the loan is granted, the agent can either default or repay. If the agent repays, the principal's and the agent's payoffs are $v - c$ and v. If the agent defaults, the payoffs are $v - c - \eta$ and $v + \beta$. Here, $\eta, \beta > 0$.

We consider the first-best benchmark: the principal observes the agent's type and dictates

[28]If payments do not matter for the principal, efficiency is readily achieved because he could pay c to the agent if and only if the report is low and nothing otherwise.

his action. Let $\overline{W}(U)$ be the value function. The domain of U is now $[0, \mu + \beta]$, where $\mu = qh + (1-q)l$.

The first-best value function depends on three payoff rates. The first and the second are $\frac{h-c}{h}$ and $\frac{l-c}{l}$. These are the payoff rates from granting the loan to a high type and a low type, respectively. The third is $-\frac{\eta}{\beta}$, which is the payoff rate from letting the agent default. Let us assume that $\frac{l-c}{l} > -\frac{\eta}{\beta}$, such that granting the loan to a low type is more efficient than letting him default. In this case, the first-best value function, $\overline{W}(U)$ for $U \in [0, \mu]$, is the same as in the model without the default option. It is readily verified that the second-best value function, $W(U)$, is also the same as before if the discount factor is close enough to one. Intuitively, the agent is deterred from defaulting if his promised utility is large enough. However, for low promised utility $U < \underline{U}$, the principal has flexibility in when to grant the loan after high reports. She can grant the loan only when U is close enough to $\underline{U}$, meaning that the agent will not default.

Renegotiation-proofness. The optimal policy, as described in Sections 3 and 4, is clearly not renegotiation-proof. After a history of reports such that the promised utility is zero, both parties would be better off by reneging and starting afresh. There are several ways to define renegotiation-proofness. Strong-renegotiation (Farrell and Maskin, 1989) implies a lower bound on the utility vectors visited (except in the event that μ is so low that it makes the relationship altogether unprofitable, and thus, $U^* = 0$.) However, the structure of the optimal policy can still be derived from the same observations. The low-type incentive-compatibility condition and promise keeping specify the continuation utilities, unless a boundary is reached regardless of whether it is the lower boundary (that must serve as a lower reflecting boundary) or the upper absorbing boundary, μ.

More types. There is no conceptual difficulty in accommodating more types, but the higher dimensionality makes it difficult to obtain an explicit solution that can be easily described geometrically. Consider the case of three types, for instance. It is clear that frontloading is the worst policy for the low type, given some promised utility to the high type, and backloading is the best, but what maximizes a medium type's utility given a pair of utilities to the low and high types? It appears that the convex hull of utilities from frontloading and backloading policies trace out the lowest utility that a medium type can obtain for any such pair, but the set of incentive-feasible payoffs has full dimension. His maximum utility obtains when one of his incentive constraints binds, but there are two possibilities, according to the binding constraint, yielding two corresponding hypersurfaces.

The analysis of the two-type case suggests that the optimal policy follows a path of utility triples on such a boundary, although a formal proof is beyond the scope of this paper.

Incomplete information regarding the process. Thus far, we have assumed that the agent's type is drawn from a distribution that is common knowledge. This feature is obviously an extreme assumption. In practice, the agent might have superior information regarding the frequency with which high values arrive. If the agent knows the distribution from the beginning, the revelation principle applies, and it is a matter of revisiting the analysis from Section 3 with an incentive-compatibility constraint at time 0.

Alternatively, the agent does not initially possess such information but is able to determine the underlying distribution from successive arrivals. This is the more challenging case in which the agent himself is learning about q (or, more generally, the transition matrix) as time passes. In that case, the agent's belief might be private (in the event that he has deviated in the past). Therefore, it is necessary to enlarge the set of reports. A mechanism is now a map from the principal's belief (regarding the agent's belief), a report by the agent of this belief, and his report on his current type (h or l) onto a decision of whether to allocate the good and the promised continuation utility.

References

[1] Abdulkadiroğlu, A. and K. Bagwell (2012). "The Optimal Chips Mechanism in a Model of Favors," working paper, Duke University and Stanford University.

[2] Abreu, D., D. Pearce, and E. Stacchetti (1990). "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, **58**, 1041–1063.

[3] Albright S.C. (1977). "A Bayesian Approach to a Generalized House Selling Problem," *Management Science*, **24**, 432–440.

[4] Athey, S. and K. Bagwell (2001). "Optimal Collusion with Private Information," *RAND Journal of Economics*, 428–465.

[5] Battaglini, M. (2005). "Long-Term Contracting with Markovian Consumers," *American Economic Review*, **95**, 637–658.

[6] Battaglini, M. and R. Lamba (2017). "Optimal Dynamic Contracting: the First-Order Approach and Beyond," working paper, Cornell University and Penn State University.

[7] Benveniste, L.M. and J.A. Scheinkman (1979). "On the Differentiability of the Value Function in Dynamic Models of Economics," *Econometrica*, **41**, 727–733.

[8] Biais, B., T. Mariotti, G. Plantin and J.-C. Rochet (2007). "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications," *Review of Economic Studies*, **74**, 345–390.

[9] Boleslavsky, R. and D. Kelly (2014). "Dynamic regulation design without payments: The importance of timing," *Journal of Public Economics*, **120**, 169–180.

[10] Bulow J. and K. Rogoff (1989). "Sovereign Debt: Is to Forgive to Forget?" *American Economic Review*, **79**, 43–50.

[11] Casella, A. (2005). "Storable Votes," *Games and Economic Behavior*, **51**, 391–419.

[12] Cohn, Z. (2010). "A Note on Linked Bargaining," *Journal of Mathematical Economics*, **46**, 238–247.

[13] Cole, H.L. and N. Kocherlakota (2001). "Efficient Allocations with Hidden Income and Hidden Storage," *Review of Economic Studies*, **68**, 523–542.

[14] Derman, C., G.J. Lieberman, and S.M. Ross (1972). "A Sequential Stochastic Assignment Problem," *Management Science*, **18**, 349–355.

[15] Doepke, M. and R.M. Townsend (2006). "Dynamic Mechanism Design with Hidden Income and Hidden Actions," *Journal of Economic Theory*, **126**, 235–285.

[16] Fang, H. and P. Norman (2006). "To Bundle or Not To Bundle," *RAND Journal of Economics*, **37**, 946–963.

[17] Eilat, R. and A. Pauzner (2011). "Optimal Bilateral Trade of Multiple Objects," *Games and Economic Behavior*, **71**, 503–512.

[18] Farrell, J. and E. Maskin (1989). "Renegotiation in Repeated Games," *Games and Economic Behavior*, **1**, 327–360.

[19] Fernandes, A. and C. Phelan (2000). "A Recursive Formulation for Repeated Agency with History Dependence," *Journal of Economic Theory*, **91**, 223–247.

[20] Frankel, A. (2011). "Contracting over Actions," PhD Dissertation, Stanford University.

[21] Fries, B.E. (1975). "Optimal ordering policy for a perishable commodity with fixed lifetime," *Operations Research*, **23**, 46–61.

[22] Fu, S. and R.V. Krishna (2017). "Dynamic Financial Contracting with Persistence," working paper, University of Rochester and Florida State University.

[23] Garrett, D. and A. Pavan (2015). "Dynamic managerial compensation: A variational approach," *Journal of Economic Theory*, **159**, 775–818.

[24] Gershkov, A. and B. Moldovanu (2010). "Efficient Sequential Assignment with Incomplete Information," *Games and Economic Behavior*, **68**, 144–154.

[25] Halac, M. and P. Yared (2014). "Fiscal Rules and Discretion Under Persistent Shocks," *Econometrica*, **82**, 1557–1614.

[26] Hauser, C. and H. Hopenhayn (2008). "Trading Favors: Optimal Exchange and Forgiveness," working paper, UCLA.

[27] Hortala-Vallve, R. (2010). "Inefficiencies on Linking Decisions," *Social Choice and Welfare*, **34**, 471–486.

[28] Hylland, A. and R. Zeckhauser (1979). "The Efficient Allocation of Individuals to Positions," *Journal of Political Economy*, **87**, 293–313.

[29] Jackson, M.O. and H.F. Sonnenschein (2007). "Overcoming Incentive Constraints by Linking Decision," *Econometrica*, **75**, 241–258.

[30] Kalla, S.J. (2010). "Essays in Favor Trading," Dissertation, University of Pennsylvania.

[31] Kováč, E., D. Krähmer and T. Tatur (2014). "Optimal Stopping in a Principal-Agent Model With Hidden Information and No Monetary Transfers," working paper, University of Bonn.

[32] Krishna, R.V., G. Lopomo and C. Taylor (2013). "Stairway to Heaven or Highway to Hell: Liquidity, Sweat Equity, and the Uncertain Path to Ownership," *RAND Journal of Economics*, **44**, 104–127.

[33] Li, J., N. Matouschek and M. Powell (2017). "Power Dynamics in Organizations," *American Economic Journal: Microeconomics*, **9**, 217–241.

[34] Lipnowski, E. and J. Ramos (2017). “Repeated Delegation,” working paper, University of Chicago.

[35] Malenko, A. (2016). “Optimal Dynamic Capital Budgeting,” working paper, MIT Sloan.

[36] Miralles, A. (2012). “Cardinal Bayesian Allocation Mechanisms without Transfers,” *Journal of Economic Theory*, **147**, 179–206.

[37] Möbius, M. (2001). “Trading Favors,” working paper, Harvard University.

[38] Radner, R. (1981). “Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship,” *Econometrica*, **49**, 1127–1148.

[39] Radner, R. (1986). “Repeated moral hazard with low discount rates,” *Uncertainty, information, and communication Essays in honor of Kenneth J. Arrow*, Vol. III, edited by W.P. Heller, R.M. Starr and D.A. Starrett, 25–64, Cambridge University Press.

[40] Rubinstein, A. and M. Yaari (1983). “Repeated Insurance Contracts and Moral Hazard,” *Journal of Economic Theory*, **30**, 74–97.

[41] Sendelbach, S. (2012). “Alarm Fatigue,” *Nursing Clinics of North America*, **47**, 375–382.

[42] Spear, S.E. and S. Srivastava (1987). “On Repeated Moral Hazard with Discounting,” *Review of Economic Studies*, **54**, 599–617.

[43] Thomas, J. and T. Worrall (1990). “Income Fluctuations and Asymmetric Information: An Example of the Repeated Principal-Agent Problem,” *Journal of Economic Theory*, **51**, 367–390.

[44] Townsend, R.M. (1982). “Optimal Multiperiod Contracts and the Gain from Enduring Relationships under Private Information,” *Journal of Political Economy*, **90**, 1166–1186.

[45] Zhang, H. and S. Zenios (2008). “A Dynamic Principal-Agent Model with Hidden Information: Sequential Optimality Through Truthful State Revelation,” *Operations Research*, **56**, 681–696.

[46] Zhang, H. (2012). “Analysis of a Dynamic Adverse Selection Model with Asymptotic Efficiency,” *Mathematics of Operations Research*, **37**, 450–474.

A Proofs for Section 3

Proof of Theorem 1. We have argued that IC_l binds. Based on PK and the binding IC_l, we solve for U_h, U_l as a function of p_h, p_l and U:

$$U_h = \frac{U - (1-\delta)p_h(qh + (1-q)l)}{\delta}, \tag{7}$$

$$U_l = \frac{U - (1-\delta)(p_h q(h-l) + p_l l)}{\delta}. \tag{8}$$

We want to show that an optimal policy is such that (i) either U_h as defined in (7) equals 0 or $p_h = 1$; and (ii) either U_l as defined in (8) equals μ or $p_l = 0$. Write $W(U; p_h, p_l)$ for the payoff from using p_h, p_l as probabilities of assigning the good, and using promised utilities as given by (7)–(8). Substituting U_h and U_l into the optimality equation, we get that, from the fundamental theorem of calculus, for any fixed $p_h < p_h'$ such that the corresponding utilities U_h, U_l are interior,

$$W(U; p_h', p_l) - W(U; p_h, p_l) = \int_{p_h}^{p_h'} (1-\delta)q \left\{h - c - (1-q)(h-l)W'(U_l) - \mu W'(U_h)\right\} \mathrm{d}p_h.$$

This expression decreases (pointwise) in $W'(U_h)$ and $W'(U_l)$. Recall that $W'(U)$ is bounded from above by $1 - c/h$. Hence, plugging in the upper bound for W', we obtain that $W(U; p_h', p_l) - W(U; p_h, p_l) \geqslant 0$. It follows that there is no loss (and possibly a gain) in increasing p_h, unless feasibility prevents this. An entirely analogous reasoning implies that $W(U; p_h, p_l)$ is nonincreasing in p_l.

It is immediate that $U_h \leqslant U_l$ and that both U_h, U_l decrease in p_h, p_l. Therefore, either $U_h \geqslant 0$ binds or p_h equals 1. Similarly, either $U_l \leqslant \mu$ binds or p_l equals 0. The uniqueness for $U > \underline{U}$ follows from Lemma 2. ■

Proof of Lemma 2. We first prove that W is strictly concave on $(\underline{U}, \mu)$. We begin with some notation and preliminary remarks. First, given any interval $I = [a, b] \subset [0, \mu]$, we write

$$I_h := \left[\frac{a - (1-\delta)\mu}{\delta}, \frac{b - (1-\delta)\mu}{\delta}\right] \cap [0, \mu],$$

$$I_l := \left[\frac{a - (1-\delta)\underline{U}}{\delta}, \frac{b - (1-\delta)\underline{U}}{\delta}\right] \cap [0, \mu].$$

We also write $[a,b]_h, [a,b]_l$ for I_h, I_l. Furthermore, given some interval I, we write $\ell(I)$ for its length.

Second, let $\overline{U} = \mu - (1-\delta)(1-q)l$. We note that, for any interval $I \subset \left[\underline{U}, \overline{U}\right]$, for $U \in I$, it holds that

$$W(U) = (1-\delta)(qh-c) + \delta q W\left(\frac{U-(1-\delta)\mu}{\delta}\right) + \delta(1-q)W\left(\frac{U-(1-\delta)\underline{U}}{\delta}\right), \quad (9)$$

and hence, over this interval, it follows by differentiation that, a.e. on I,

$$W'(U) = qW'(U_h) + (1-q)W'(U_l).$$

Similarly, for any interval $I \subset \left[\overline{U}, \mu\right]$, for $U \in I$,

$$W(U) = (1-q)\left(U - c - (U-\mu)\frac{c}{l}\right) + (1-\delta)q(\mu - c) + \delta q W\left(\frac{U-(1-\delta)\mu}{\delta}\right), \quad (10)$$

and so a.e.,

$$W'(U) = (1-q)(1-c/l) + qW'(U_h).$$

That is, the slope of W at a point is an average of the slopes at U_h, U_l. This also holds on $\left[\overline{U}, \mu\right]$, with the convention that its slope at $U_l = \mu$ is given by $1 - c/l$. By weak concavity of W, if W is affine on I, then it must be affine on both I_h and I_l (with the convention that it is trivially affine at μ). We make the following observations.

1. For any $I \subseteq (\underline{U}, \mu)$ (of positive length) such that W is affine on I, $\ell(I_h \cap I) = \ell(I_l \cap I) = 0$. If not, then we note that, because the slope on I is the average of the other two, all three must have the same slope (since two intersect, and so have the same slope). But then the convex hull of the three has the same slope (by weak concavity). We thus obtain an interval $I' = \text{co}\{I_l, I_h\}$ of strictly greater length (note that $\ell(I_h) = \ell(I)/\delta$, and similarly $\ell(I_l) = \ell(I)/\delta$ unless I_l intersects μ). It must then be that I'_h or I'_l intersect I', and we can repeat this operation. This contradicts the fact the slope of W on $[0, \underline{U}]$ is $(1 - c/h)$, yet $W(\mu) = \mu - c$.

2. It follows that there is no interval $I \subseteq [\underline{U}, \mu]$ on which W has slope $(1 - c/h)$ (because then W would have this slope on $I' := \text{co}\{\{\underline{U}\} \cup I\}$, and yet I' would intersect I'_l.) Similarly, there cannot be an interval $I \subseteq [\underline{U}, \mu]$ on which W has slope $1 - c/l$.

3. It immediately follows from 2 that $W < \overline{W}$ on $(\underline{U}, \mu)$: if there is a $U \in (\underline{U}, \mu)$ such that

$W(U) = \overline{W}(U)$, then by concavity again (and the fact that the two slopes involved are the two possible values of the slope of $\overline{W}$), W must either have slope $(1 - c/h)$ on $[0, U]$, or $(1 - c/l)$ on $[U, \mu]$, both being impossible.

4. Next, suppose that there exists an interval $I \subset [\underline{U}, \mu)$ of length $\varepsilon > 0$ such that W is affine on I. There might be many such intervals; consider the one with the smallest lower extremity. Furthermore, without loss, given this lower extremity, pick I so that it has maximum length, that W is affine on I, but on no proper superset of I. Let $I := [a, b]$. We claim that $I_h \in [0, \underline{U}]$. Suppose not. Note that I_h cannot overlap with I (by point 1). Hence, either I_h is contained in $[0, \underline{U}]$, or it is contained in $[\underline{U}, a]$, or $\underline{U} \in (a, b)_h$. This last possibility cannot occur, because W must be affine on $(a, b)_h$, yet the slope on $(a_h, \underline{U})$ is equal to $(1 - c/h)$, while by point 2 it must be strictly less on $(\underline{U}, b_h)$. It cannot be contained in $[\underline{U}, a]$, because $\ell(I_h) = \ell(I)/\delta > \ell(I)$, and this would contradict the hypothesis that I was the lowest interval in $[\underline{U}, \mu]$ of length ε over which W is affine.

   It follows from 1 that I_l cannot intersect I. Assume $b \leqslant \overline{U}$. Hence, we have that I_l is an interval over which W is affine, and such that $\ell(I_l) = \ell(I)/\delta$. Let $\varepsilon' := \ell(I)/\delta$. By the same reasoning as before, we can find $I' \subset [\underline{U}, \mu)$ of length $\varepsilon' > 0$ such that W is affine on I', and such that $I'_h \subset [0, \underline{U}]$. Repeating the same argument as often as necessary, we conclude that there must be an interval $J \subset [\underline{U}, \mu)$ such that (i) W is affine on J, $J = [a', b']$, $b' \geqslant \overline{U}$, (ii) there exists no interval of equal or greater length in $[\underline{U}, \mu)$ over which W would be affine. By the same argument yet again, J_h must be contained in $[0, \underline{U}]$. Yet the assumption that $\delta > 1/2$ is equivalent to $\overline{U}_h = \frac{\overline{U} - (1-\delta)\mu}{\delta} > \underline{U}$, and so this is a contradiction. Hence, there exists no interval in $(\underline{U}, \mu)$ over which W is affine, and so W must be strictly concave.

Differentiability follows from an argument that follows Benveniste and Scheinkman (1979), based on some induction. We note that W is differentiable on $(0, \underline{U})$. Fix $U > \underline{U}$. Consider the following perturbation of the optimal policy. Fix $\varepsilon = (p - \bar{p})^2$, for some $\bar{p} \in (0, 1)$ to be determined. With probability $\varepsilon > 0$, the report is ignored, the good is supplied with probability $p \in [0, 1]$ and the next value is U_l (Otherwise, the optimal policy is implemented). Because this event is independent of the report, the IC constraints are still satisfied. Note that, for $p = 0$, this yields a strictly lower utility than U to the agent, while it yields a strictly higher utility for $p = 1$. As it varies continuously, there is some critical value—defined as $\bar{p}$—that makes the agent indifferent between both policies. By varying p, we may thus generate all utilities within some interval $(U-\nu, U+\nu)$, for some $\nu > 0$, and the payoff $\tilde{W}$ that we obtain in this fashion is continuously differentiable in $U' \in (U-\nu, U+\nu)$. It follows that the concave function W is above a continuously differentiable function $\tilde{W}$—hence, it must be as well. ■

Lemma 5. *W has a unique maximizer, $U^* = \arg\max_U W(U)$. U^* is nonincreasing in c.*

Proof. Uniqueness of U^* follows from strict concavity of W on $[\underline{U}, \mu]$. Consider now (9)–(10): replace the function W that appears on the right-hand side with a function W_n, define the left-hand side W_{n+1} iteratively, and set $W_0 = 0$ identically. By induction, the function W_n admits a cross-partial in (U, c) a.e., which is nonpositive (note that the "flow payoff" in (10) involves a negative cross-partial term $-(1-q)/l$. It follows by convergence of W_n to W that W has a nonpositive cross-partial (a.e.) as well, implying that U^* is nonincreasing in c. ■

B Proofs for Section 4

Proof of Lemma 4. Given an arbitrary $A \subset [0, \mu_h] \times [0, \mu_l]$, we first define the operator $\mathcal{B}$:

$$\mathcal{B}(A) := \{U \in [0, \mu_h] \times [0, \mu_l] : p_h, p_l \in [0,1], U_h, U_l \in A, \text{ subject to } (PK_h)\text{–}(IC_l)\}.$$

Substituting the agent's expected value in each round into (2) and (3), we obtain the formula for frontloading and backloading policies:

$$f_h^\nu = (1-\delta^\nu)\mu_h + \delta^\nu(1-q)(\mu_h-\mu_l)(1-\kappa^\nu), \quad f_l^\nu = (1-\delta^\nu)\mu_l - \delta^\nu q(\mu_h-\mu_l)(1-\kappa^\nu),$$
$$b_h^\nu = \delta^\nu\mu_h - \delta^\nu(1-q)(\mu_h-\mu_l)(1-\kappa^\nu), \qquad b_l^\nu = \delta^\nu\mu_l + \delta^\nu q(\mu_h-\mu_l)(1-\kappa^\nu).$$

Let X denote the set $\text{co}\{b^\nu, f^\nu : \nu \geqslant 0\}$. The point b^0 is supported by $(p_h, p_l) = (1,1), U_h = U_l = (\mu_h, \mu_l)$. For $\nu \geqslant 1$, b^ν is supported by $(p_h, p_l) = (0,0), U_h = U_l = b^{\nu-1}$. The point f^0 is supported by $(p_h, p_l) = (0,0), U_h = U_l = (0,0)$. For $\nu \geqslant 1$, f^ν is supported by $(p_h, p_l) = (1,1), U_h = U_l = f^{\nu-1}$. Therefore, we have $X \subset \mathcal{B}(X)$. This implies that $\mathcal{B}(X) \subset V$.

We define two sequences of utility vectors as follows. First, for $\nu \geqslant 0$, let

$$\overline{w}_h^\nu = \delta^\nu (1-\kappa^\nu)(1-q)\mu_l, \quad \overline{w}_l^\nu = \delta^\nu (1-q+\kappa^\nu q)\mu_l,$$

and set $\overline{w}^\nu = (\overline{w}_h^\nu, \overline{w}_l^\nu)$. Second, for $\nu \geqslant 0$, let

$$\underline{w}_h^\nu = \mu_h - \delta^\nu (1-\kappa^\nu)(1-q)\mu_l, \quad \underline{w}_l^\nu = \mu_l - \delta^\nu (1-q+\kappa^\nu q)\mu_l,$$

and set $\underline{w}^\nu = (\underline{w}_h^\nu, \underline{w}_l^\nu)$. The sequence $\overline{w}^\nu$ starts at $\overline{w}^0 = (0, \mu_l)$ with $\lim_{\nu\to\infty} \overline{w}^\nu = 0$. Similarly, $\underline{w}^\nu$ starts at $\underline{w}^0 = (\mu_h, 0)$ and $\lim_{\nu\to\infty} \underline{w}^\nu = (\mu_h, \mu_l)$. For any $\nu \geqslant 1$, $\overline{w}^\nu$ is supported by $(p_h, p_l) = (0,0), U_h = U_l = \overline{w}^{\nu-1}$, and $\underline{w}^\nu$ is supported by $(p_h, p_l) = (1,1), U_h = U_l = \underline{w}^{\nu-1}$. We define a set sequence $\{X^\nu\}_{\nu\geqslant 0}$ as follows:

$$X^\nu = \mathrm{co}\left(\{b^k, f^k : 0 \leqslant k \leqslant \nu\} \cup \{\overline{w}^\nu, \underline{w}^\nu\}\right).$$

It is obvious that $X^0 = [0,\mu_h]\times[0,\mu_l]$ so $V \subset \mathcal{B}(X^0) \subset X^0$. To prove that $V = X$, it suffices to show that (i) $X^\nu = \mathcal{B}(X^{\nu-1})$, and (ii) $\lim_{\nu\to\infty} X^\nu = X$. The latter is obvious given that $\lim_{\nu\to\infty} \overline{w}^\nu = 0$ and $\lim_{\nu\to\infty} \underline{w}^\nu = (\mu_h, \mu_l)$. We next prove (i).

For any $\nu \geqslant 1$, we define the supremum score in direction (λ_1, λ_2) given $X^{\nu-1}$ as $K((\lambda_1,\lambda_2), X^{\nu-1}) = \sup_{p_h,p_l,U_h,U_l}(\lambda_1 U(h) + \lambda_2 U(l))$, subject to (PK_h)–(IC_l), $p_h, p_l \in [0,1]$, and $U_h, U_l \in X^{\nu-1}$. The set $\mathcal{B}(X^{\nu-1})$ is given by

$$\bigcap_{(\lambda_1,\lambda_2)\in\mathbf{R}^2} \left\{(U(h),U(l)) : \lambda_1 U(h) + \lambda_2 U(l) \leqslant K((\lambda_1,\lambda_2), X^{\nu-1})\right\}.$$

Without loss of generality, we focus on directions $(1,-\lambda)$ and $(-1,\lambda)$ for all $\lambda \geqslant 0$. We define three sequences of slopes as follows:

$$\begin{aligned}
\lambda_1^\nu &= \frac{(1-q)(\delta\kappa-1)\kappa^\nu(\mu_h-\mu_l) - (1-\delta)(q\mu_h+(1-q)\mu_l)}{q(1-\delta\kappa)\kappa^\nu(\mu_h-\mu_l) - (1-\delta)(q\mu_h+(1-q)\mu_l)},\\
\lambda_2^\nu &= \frac{1-(1-q)(1-\kappa^\nu)}{q(1-\kappa^\nu)},\\
\lambda_3^\nu &= \frac{(1-q)(1-\kappa^\nu)}{q\kappa^\nu+(1-q)}.
\end{aligned}$$

It is easy to verify that

$$\lambda_1^\nu = \frac{b_h^\nu - b_h^{\nu+1}}{b_l^\nu - b_l^{\nu+1}} = \frac{f_h^\nu - f_h^{\nu+1}}{f_l^\nu - f_l^{\nu+1}}, \quad \lambda_2^\nu = \frac{b_h^\nu - \overline{w}_h^\nu}{b_l^\nu - \overline{w}_l^\nu} = \frac{f_h^\nu - \underline{w}_h^\nu}{f_l^\nu - \underline{w}_l^\nu}, \quad \lambda_3^\nu = \frac{\overline{w}_h^\nu - 0}{\overline{w}_l^\nu - 0} = \frac{\underline{w}_h^\nu - \mu_h}{\underline{w}_l^\nu - \mu_l}.$$

When $(\lambda_1, \lambda_2) = (-1, \lambda)$, the supremum score as we vary λ is

$$K((-1,\lambda), X^{\nu-1}) = \begin{cases} (-1,\lambda)\cdot(0,0) & \text{if } \lambda \in [0, \lambda_3^\nu] \\ (-1,\lambda)\cdot \overline{w}^\nu & \text{if } \lambda \in [\lambda_3^\nu, \lambda_2^\nu] \\ (-1,\lambda)\cdot b^\nu & \text{if } \lambda \in [\lambda_2^\nu, \lambda_1^{\nu-1}] \\ (-1,\lambda)\cdot b^{\nu-1} & \text{if } \lambda \in [\lambda_1^{\nu-1}, \lambda_1^{\nu-2}] \\ \cdots & \\ (-1,\lambda)\cdot b^0 & \text{if } \lambda \in [\lambda_1^0, \infty). \end{cases}$$

Similarly, when $(\lambda_1, \lambda_2) = (1, -\lambda)$, we have

$$K((1,-\lambda), X^{\nu-1}) = \begin{cases} (1,-\lambda)\cdot(\mu_h,\mu_l) & \text{if } \lambda \in [0, \lambda_3^\nu] \\ (1,-\lambda)\cdot \underline{w}^\nu & \text{if } \lambda \in [\lambda_3^\nu, \lambda_2^\nu] \\ (1,-\lambda)\cdot f^\nu & \text{if } \lambda \in [\lambda_2^\nu, \lambda_1^{\nu-1}] \\ (1,-\lambda)\cdot f^{\nu-1} & \text{if } \lambda \in [\lambda_1^{\nu-1}, \lambda_1^{\nu-2}] \\ \cdots & \\ (1,-\lambda)\cdot f^0 & \text{if } \lambda \in [\lambda_1^0, \infty). \end{cases}$$

Therefore, we have $X^\nu = \mathcal{B}(X^{\nu-1})$. Note that this method only works when parameters are such that $\lambda_3^\nu \leqslant \lambda_2^\nu \leqslant \lambda_1^{\nu-1}$ for all $\nu \geqslant 1$. If $\rho_l/(1-\rho_h) \geqslant l/h$, the proof stated above applies. Otherwise, the following proof applies.

We define two sequences of utility vectors as follows. First, for $0 \leqslant m \leqslant \nu$, let

$$\begin{aligned} \overline{w}_h(m,\nu) &= \delta^{\nu-m}\left(q\mu_h\left(1-\delta^m\right)+(1-q)\mu_l\right) - (1-q)(\delta\kappa)^{\nu-m}\left(\mu_h\left((\delta\kappa)^m-1\right)+\mu_l\right), \\ \overline{w}_l(m,\nu) &= \delta^{\nu-m}\left(q\mu_h\left(1-\delta^m\right)+(1-q)\mu_l\right) + q(\delta\kappa)^{\nu-m}\left(\mu_h\left((\delta\kappa)^m-1\right)+\mu_l\right), \end{aligned}$$

and set $\overline{w}(m,\nu) = (\overline{w}_h(m,\nu), \overline{w}_l(m,\nu))$. Second, for $0 \leqslant m \leqslant \nu$, let

$$\begin{aligned} \underline{w}_h(m,\nu) &= \frac{(1-q)\delta^\nu\kappa^\nu\left(\mu_h\left(\delta^m\kappa^m-1\right)+\mu_l\right)+\kappa^m\left(\mu_h\delta^m-\delta^\nu\left(q\mu_h\left(1-\delta^m\right)+(1-q)\mu_l\right)\right)}{\delta^m\kappa^m}, \\ \underline{w}_l(m,\nu) &= \frac{-q\delta^\nu\kappa^\nu\left(\mu_h\left(\delta^m\kappa^m-1\right)+\mu_l\right)+\kappa^m\left(\mu_l\delta^m-\delta^\nu\left(q\mu_h\left(1-\delta^m\right)+(1-q)\mu_l\right)\right)}{\delta^m\kappa^m}, \end{aligned}$$

and set $\underline{w}(m,\nu) = (\underline{w}_h(m,\nu), \underline{w}_l(m,\nu))$. Fixing ν, the sequence $\overline{w}(m,\nu)$ is increasing (in both its arguments) as m increases, with $\lim_{\nu\to\infty} \overline{w}(\nu-m,\nu) = b^m$. Similarly, fixing ν,

$\underline{w}(m,\nu)$ is decreasing as m increases, $\lim_{\nu\to\infty}\underline{w}(\nu-m,\nu)=f^m$.

Let $\overline{Y}^\nu=\{\overline{w}(m,\nu):0\leqslant m\leqslant\nu\}$ and $\underline{Y}^\nu=\{\underline{w}(m,\nu):0\leqslant m\leqslant\nu\}$. We define a set sequence $\{Y^\nu\}_{\nu\geqslant 0}$ as follows:

$$Y^\nu=\operatorname{co}\left(\{(0,0),(\mu_h,\mu_l)\}\cup\overline{Y}^\nu\cup\underline{Y}^\nu\right).$$

It is obvious that $Y^0=[0,\mu_h]\times[0,\mu_l]$ so $V\subset\mathcal{B}(Y^0)\subset Y^0$. To prove that $V=X$, it suffices to show that (i) $Y^\nu=\mathcal{B}(Y^{\nu-1})$ and (ii) $\lim_{\nu\to\infty}Y^\nu=X$. The rest of the proof is similar to the first part and hence omitted. ■

B.1 The General Solution

Theorem 2 follows from the analysis of the optimal policy on the entire domain V. Here, we solve the program in Section 4.1 for the entire V.

We further divide V into subsets and introduce two sequences of utility vectors. First, given $\underline{U}$, define the sequence $\{\underline{v}^\nu\}_{\nu\geqslant 0}$ by

$$\underline{v}_h^\nu=\mathbf{E}\left[\delta^\nu\underline{U}_{v_\nu}\mid v_0=h\right],\quad \underline{v}_l^\nu=\mathbf{E}\left[\delta^\nu\underline{U}_{v_\nu}\mid v_0=l\right]. \tag{11}$$

Intuitively, this is the payoff from waiting for ν rounds from initial value h or l, and then getting $\underline{U}$. (See Footnote 21 for the definition of $\underline{U}$.) Let $\underline{V}$ be the payoff vectors in V that lie below the graph of the set of points $\{\underline{v}^\nu\}_{\nu\geqslant 0}$. Figure 6 illustrates this construction. Note that $\underline{V}$ has a nonempty interior if and only if ρ_l is sufficiently large (see (4)). This set is the domain of utilities for which the complete-information payoff can be achieved, as stated below.

Lemma 6. *For all $U\in\underline{V}\cup\{(\mu_h,\mu_l)\}$ and all ϕ,*

$$W(U,\phi)=\overline{W}(U,\phi).$$

Conversely, if $U\notin\underline{V}\cup\{(\mu_h,\mu_l)\}$, then $W(U,\phi)<\overline{W}(U,\phi)$ for all $\phi\in(0,1)$ when $\delta>\underline{\delta}$, where $\underline{\delta}$ is given by the condition that $b_h^1=v_h^*$.

To understand Lemma 6, we first observe that if the agent is promised $\underline{U}$, his future promised utility after a low report is exactly $\underline{U}$ under the optimal policy. Therefore, for any U that is on the frontloading boundary of V and lies below $\underline{U}$, the complete-information payoff can be achieved. Second, for any $\nu\geqslant 1$, the utility $\underline{v}^\nu$ can be delivered by not supplying the unit in the current round and setting the future utility to be $\underline{v}^{\nu-1}$, regardless of the report. The agent's promised utility becomes $\underline{U}(=\underline{v}^0)$ after ν rounds. From this point on, the optimal policy as specified in Theorem 2 is implemented. Clearly, under this policy, the unit is never supplied when the agent's value is low. Therefore, the complete-information payoff is achieved.

Second, we define $\hat{u}^\nu := (\hat{u}_h^\nu, \hat{u}_l^\nu)$, $\nu \geqslant 0$ as follows:

$$\hat{u}_h^\nu = \mathbf{E}\left[(1-\delta)\sum_{n=1}^{\nu}\delta^n v_n \mid v_0 = h\right], \quad \hat{u}_l^\nu = \mathbf{E}\left[(1-\delta)\sum_{n=1}^{\nu}\delta^n v_n \mid v_0 = l\right]. \tag{12}$$

This is the utility vector if the principal supplies the unit from round 1 to round ν, independently of the reports. Note that the unit is not supplied in round 0. We note that $\hat{u}^0 = 0$ and $\hat{u}^\nu$ is an increasing sequence (in both coordinates) contained in V, where $\lim_{\nu\to\infty}\hat{u}^\nu = b^1$. The sequence $\{\hat{u}^\nu\}_{\nu\geqslant 0}$ defines a polygonal chain P that divides $V \setminus \underline{V}$ into two subsets, V_b and V_f, consisting of those points in $V \setminus \underline{V}$ that lie above or below P. We also let P_f, P_b be the (closure of the) polygonal chains defined by $\{f^\nu\}_{\nu\geqslant 0}$ and $\{b^\nu\}_{\nu\geqslant 0}$ that correspond to the frontloading and backloading boundaries of V.

We now define a policy (which, as we will see below, is optimal) ignoring for the present the choice of the initial promise.

Definition 2. *For all $U \in V$, set*

$$p_h = \min\left\{1, \frac{U(h)}{(1-\delta)h}\right\}, \quad p_l = \max\left\{0, 1 - \frac{\mu_l - U(l)}{(1-\delta)l}\right\}, \tag{13}$$

and

$$U_h \in P_f, \quad U_l \in \begin{cases} P_f & \text{if } U \in V_f \\ P_b & \text{if } U \in P_b. \end{cases}$$

Furthermore, if $U \in V_b$, U_l is chosen such that IC_h binds.

For any U, the current allocation is the efficient one (as long as possible), and the future utilities U_h and U_l are chosen so that PK_h and PK_l are satisfied. We show that the closer U is to the frontloading boundary of V, the higher the principal's payoff is. Therefore, U_h and U_l shall be as close to P_f as IC_l and IC_h permit. One can always choose U_h to be on P_f, because moving U_h toward P_f makes IC_l less binding. However, one cannot always choose U_l to be on P_f without violating IC_h, because the high type might have an incentive to mimic the low type as we increase $U_l(h)$ and decrease $U_l(l)$. This is where the definition of P plays the role. If the promised utility is $(\hat{u}_h^\nu, \hat{u}_l^\nu) \in P$, the high type is promised the expected utility from foregoing the unit in round 0 and consuming the unit from round 1 to ν. Therefore, if the low type consumes no unit in the current round and is promised a future utility that is on the frontloading boundary, IC_h holds with equality. Hence, for all utilities above P, U_l is chosen such that IC_h binds, whereas for all utilities below P, one chooses U_l to be on the frontloading boundary. It is readily verified that the policy and choices of U_l, U_h also imply that IC_l binds for all $U \in P_f$.

This policy is independent of the principal's belief. That is, the belief regarding the agent's type today does not enter the optimal policy, for a fixed promised utility. However, the belief affects the optimal initial promise and the payoff.

Figure 6 illustrates the dynamics. Given any promised utility vector, the vector $(p_h, p_l) = (1, 0)$ is used (unless U is too close to 0 or (μ_h, μ_l)), and promised utilities depend on the report. A report of l shifts the utility to the right (toward higher values), whereas a report of h shifts it to the left and toward the frontloading boundary. Below the interior polygonal chain, utility jumps to the frontloading boundary after an l report; above it, the jump is determined by IC_h. If the utility starts from the frontloading boundary, the continuation utility after an l report stays there.

For completeness, we also define the subsets over which promise keeping prevents the efficient choices $(p_h, p_l) = (1, 0)$ from being made. Let $V_h = \{U : U \in V, U(l) \geqslant b_l^1\}$ and $V_l = \{U : U \in V, U(h) \leqslant f_h^1\}$. It is easily verified that $(p_h, p_l) = (1, 0)$ is feasible at U given promise keeping if and only if $U \in V \setminus (V_h \cup V_l)$.

Theorem 3. *For any $U \in V$, the policy in Definition 2 is optimal. The initial promise U^* is in $P_f \cap (V \setminus \underline{V})$, with U^* increasing in the principal's prior belief.*

Furthermore, the value function $W(U(h), U(l), \phi)$ is weakly increasing in $U(h)$ along the rays $x = \phi U(h) + (1-\phi)U(l)$ for any $\phi \in \{1-\rho_h, \rho_l\}$.

Given that $U^* \in P_f$ and given the structure of the optimal policy, the promised utility vector never leaves P_f. It is also simple to verify that, as in the i.i.d. case (and by the same arguments), the (one-sided) derivative of W approaches the derivative of $\overline{W}$ as U approaches either (μ_h, μ_l) or the set $\underline{V}$. As a result, the initial promise U^* is in $P_f \cap (V \setminus \underline{V})$.

Proof of Lemma 6. It will be useful in this proof and those that follow to define the operator $\mathcal{B}_{ij}$ for $i, j \in \{0, 1\}$. Given an arbitrary $A \subset [0, \mu_h] \times [0, \mu_l]$, let

$$\mathcal{B}_{ij}(A) := \{U \in [0, \mu_h] \times [0, \mu_l] : U_h, U_l \in A \text{ solving } (PK_h)\text{–}(IC_l) \text{ for } (p_h, p_l) = (i, j)\},$$

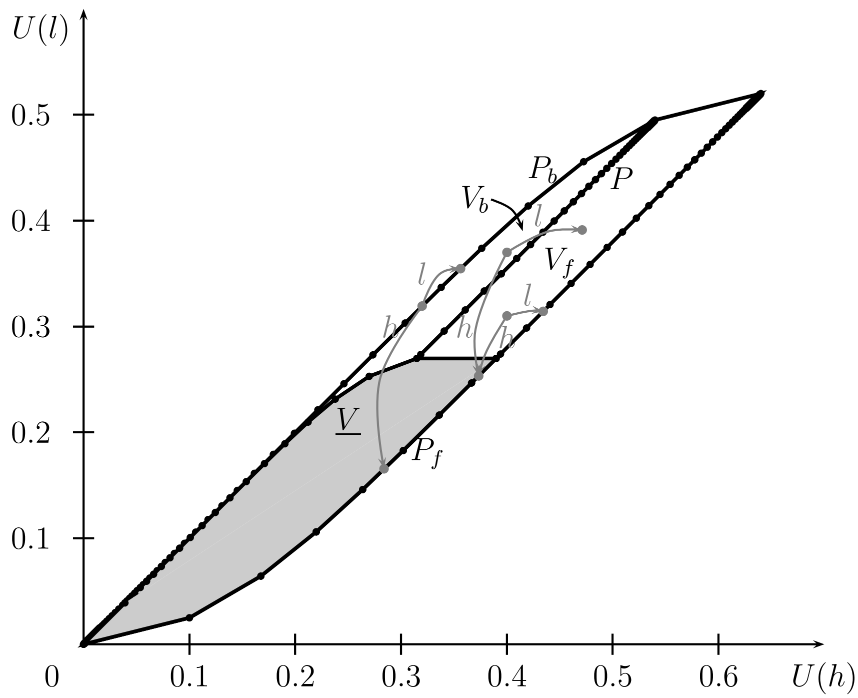

Figure 6: Incentive-feasible set V and the optimal policy for $(\delta, \rho_h, \rho_l, l, h) = (9/10, 1/3, 1/4, 1/4, 1)$.

and similarly $\mathcal{B}_{i\cdot}(A), \mathcal{B}_{\cdot j}(A)$ when only p_h or p_l is constrained.

The first step is to compute V_0, the largest set such that $V_0 \subset \mathcal{B}_{\cdot 0}(V_0)$. Plainly, this is a proper subset of V, because any promise $U(l) \in (b_l^1, \mu_l]$ requires that p_l be strictly positive.

We first show that $\underline{V} \subset V_0$. Substituting the probability of being h in round ν conditional on the current type into (11), we obtain the analytic expression of $\{\underline{v}^\nu\}_{\nu \geqslant 0}$:

$$\underline{v}_h^\nu = \delta^\nu \left((1-q)\underline{U}_l + q\underline{U}_h + (1-q)\kappa^\nu(\underline{U}_h - \underline{U}_l)\right), \underline{v}_l^\nu = \delta^\nu \left((1-q)\underline{U}_l + q\underline{U}_h - q\kappa^\nu(\underline{U}_h - \underline{U}_l)\right).$$

Note that the sequence $\{\underline{v}^\nu\}_{\nu \geqslant 0}$ solves the system of equations, for all $\nu \geqslant 0$:

$$\underline{v}_h^{\nu+1} = \delta(1-\rho_h)\underline{v}_h^\nu + \delta\rho_h\underline{v}_l^\nu, \quad \underline{v}_l^{\nu+1} = \delta(1-\rho_l)\underline{v}_l^\nu + \delta\rho_l\underline{v}_h^\nu,$$

and $\underline{v}_l^1 = \underline{v}_l^0$ (which is easily verified given the conditions that $\underline{v}^0 = \underline{U}$ and that $\underline{U}$ lies on the line $U(l) = \frac{\delta\rho_l}{1-\delta(1-\rho_l)}U(h)$). First, for any $\nu \geqslant 1$, $\underline{v}^\nu$ can be supported by setting $p_h = p_l = 0$ and $U_h = U_l = \underline{v}^{\nu-1}$. Therefore, $\underline{v}^\nu$ can be delivered with p_l being 0 and continuation utilities in $\underline{V}$. Second, we show that $\underline{v}^0$ can itself be obtained with continuation utilities in $\underline{V}$. This one is obtained by setting $(p_h, p_l) = (1, 0)$, setting IC_l as a binding constraint, and $U_l = \underline{v}^0$ (again one can check that U_h is in $\underline{V}$ and that IC_h holds). Third, for any $U \in P_f \cap \underline{V}$, U can be delivered with p_l being 0 and continuation utilities on $P_f \cap \underline{V}$. This shows that $\underline{V} \subseteq V_0$, because the extreme points of $\underline{V}$ can be supported with p_l being 0 and continuation utilities in $\underline{V}$, and all other utility vectors in $\underline{V}$ can be written as a convex combination of these extreme points.

The proof that $V_0 \subset \underline{V}$ follows the similar lines as determining the boundaries of V in the proof of Lemma 4: one considers a sequence of programs, setting $Z^0 = V$ and defining recursively the supremum score in direction (λ_1, λ_2) given $Z^{\nu-1}$ as $K((\lambda_1, \lambda_2), Z^{\nu-1}) = \sup_{p_h, p_l, U_h, U_l} \lambda_1 U(h) + \lambda_2 U(l)$, subject to (PK_h)–(IC_l), $p_l = 0$, $p_h \in [0,1]$, $U_h, U_l \in Z^{\nu-1}$, $\lambda \cdot U_h \leqslant \lambda \cdot U$ and $\lambda \cdot U_l \leqslant \lambda \cdot U$. Let the set $\mathcal{B}(Z^{\nu-1})$ be given by

$$\bigcap_{(\lambda_1,\lambda_2)} \left\{(U(h), U(l)) \in V : \lambda_1 U(h) + \lambda_2 U(l) \leqslant K((\lambda_1, \lambda_2), Z^{\nu-1})\right\},$$

and the set $Z^\nu = \mathcal{B}(Z^{\nu-1})$ obtains by considering an appropriate choice of λ_1, λ_2. More precisely, we always set $\lambda_2 = 1$, and for $\nu = 1$, pick $\lambda_1 = 0$. This gives $Z^1 = V \cap \{U : U(l) \leqslant \underline{v}_l^0\}$. We then pick (for every $\nu \geqslant 2$) as direction λ the vector $(-(\underline{v}_l^{\nu-1} - \underline{v}_l^\nu)/(\underline{v}_h^{\nu-1} - \underline{v}_h^\nu), 1)$, and as a result obtain that

$$V_0 \subseteq Z^\nu = Z^{\nu-1} \cap \left\{U : U(l) - \underline{v}_l^\nu \leqslant \frac{\underline{v}_l^{\nu-1} - \underline{v}_l^\nu}{\underline{v}_h^{\nu-1} - \underline{v}_h^\nu}(U(h) - \underline{v}_h^\nu)\right\}.$$

It follows that $V_0 \subset \underline{V}$.

Next, we argue that the complete-information payoff is achieved in $\underline{V}$. It is clear that any policy that never gives the unit to the low type must be optimal. This is a feature of the policies that we have described to obtain the boundary of $\underline{V}$.

Finally, one must show that the complete-information payoff cannot be achieved for $U \notin \underline{V} \cup \{(\mu_h, \mu_l)\}$.

1. Since $\underline{V}$ is the largest fixed point of $\mathcal{B}_{\cdot 0}$, starting from any utility $U \notin \underline{V}$, there is a positive probability that the unit is given (after some history that has positive probability) to the low value. This implies that the complete-information payoff cannot be achieved for $U \leqslant v^*$, $U \notin \underline{V}$.

2. For $U \geqslant v^*$, achieving the complete-information payoff requires that the agent gets the unit whenever he reports h. This means that the agent can report h and get the unit all the time. The agent's interim utility is effectively (μ_h, μ_l). Therefore, the complete-information payoff is achieved only at $U = (\mu_h, \mu_l)$.

3. We are left with $U \in V$ such that $U(l) > v_l^*$ and $U(h) < v_h^*$. Given that $U(l) > v_l^*$, the complete-information requires that if the agent reports low today, then he gets the unit after the high report from tomorrow on. This implies that the agent will get at least b^1 as his interim utilities, contradicting the presumption that $U(h) < v_h^*$ since $v_h^* \leqslant b_h^1$ when δ is high enough.

■

Proof of Theorem 2 and 3. **Part I:** We first examine a relaxed problem in which IC_h is dropped. We define the function $\tilde{W} : V \times \{\rho_l, 1-\rho_h\} \to \mathbf{R} \cup \{-\infty\}$, that solves the following program: for any $U \in V, \phi \in \{\rho_l, 1-\rho_h\}$,

$$\begin{aligned} \tilde{W}(U,\phi) &= \sup_{p_h,p_l,U_h,U_l} \phi\left((1-\delta)p_h(h-c)+\delta\tilde{W}(U_h,1-\rho_h)\right) \\ &+ (1-\phi)\left((1-\delta)p_l(l-c)+\delta\tilde{W}(U_l,\rho_l)\right), \end{aligned}$$

over $p_h, p_l \in [0,1]$ and $U_h, U_l \in V$, subject to PK_h, PK_l, IC_l. Given that IC_h is dropped, this is a relaxed version of the original problem. It holds that $\tilde{W}(U,\phi) \geqslant W(U,\phi)$ for any $U \in V, \phi \in \{\rho_l, 1-\rho_h\}$.

We want to show that the value function $\tilde{W}$ satisfies two properties (I) and (II):

(I) $\tilde{W}(U(h),U(l),1-\rho_h)$ is weakly increasing in $U(h)$ along the rays $x = (1-\rho_h)U(h) + \rho_h U(l)$, and $\tilde{W}(U(h),U(l),\rho_l)$ is weakly increasing in $U(h)$ along the rays $y = \rho_l U(h) + (1-\rho_l)U(l)$;

(II) Given $\tilde{W}$, we define two functions $\tilde{w}_h(x)$ and $\tilde{w}_l(y)$. Let $(U_{h1}(x), U_{l1}(x))$ be the intersection of P_f and the line $x = (1-\rho_h)U(h) + \rho_h U(l)$. We define $\tilde{w}_h(x) := \tilde{W}(U_{h1}(x), U_{l1}(x), 1-\rho_h)$ on the domain $[0, (1-\rho_h)\mu_h + \rho_h\mu_l]$. Similarly, $(U_{h2}(y), U_{l2}(y))$ denotes the intersection of P_f and the line $y = \rho_l U(h) + (1-\rho_l)U(l)$. We define $\tilde{w}_l(y) := \tilde{W}(U_{h2}(y), U_{l2}(y), \rho_l)$ on the domain $[0, \rho_l\mu_h + (1-\rho_l)\mu_l]$. For any U, let $X(U) = (1-\rho_h)U(h) + \rho_h U(l)$ and $Y(U) = \rho_l U(h) + (1-\rho_l)U(l)$. We want to show that (II.a) $\tilde{w}_h(\cdot)$ and $\tilde{w}_l(\cdot)$ are concave; (II.b) $\tilde{w}_h', \tilde{w}_l'$ are in $[1-c/l, 1-c/h]$ (throughout this proof, derivatives have to be understood as either right- or left-derivatives, depending on the inequality); and (II.c) for any U on P_f

$$\tilde{w}_h'(X(U)) \geqslant \tilde{w}_l'(Y(U)). \tag{14}$$

If $\tilde{W}$ satisfies (I) and (II), then it follows that for the relaxed problem, for any U:

1. It is optimal to choose U_h and U_l that lie on P_f. First, for any $U \in V$, it is feasible to choose a U_h on P_f since the intersection of IC_l and PK_h is above P_f. It is also feasible to choose a U_l on P_f since IC_h is dropped. Given property (I), the principal prefers to have U_h, U_l as close to P_f as possible.

2. It is optimal to choose p_h, p_l as in (13), *i.e.*, to choose p_h and p_l as efficiently as possible. This is due to property (II.b) that w_h', w_l' are in $[1 - c/l, 1 - c/h]$.

Next, we show that $\tilde{W}$ satisfies (I) and (II) by an argument of value function iterations. Let $\mathcal{W}$ be the set of all functions $\hat{W}(U, 1-\rho_h)$ and $\hat{W}(U, \rho_l)$ that satisfy (I) and (II). More specifically,

1. $\hat{W}(U(h), U(l), 1-\rho_h)$ is weakly increasing in $U(h)$ along the rays $x = (1-\rho_h)U(h) + \rho_h U(l)$, and $\hat{W}(U(h), U(l), \rho_l)$ is weakly increasing in $U(h)$ along the rays $y = \rho_l U(h) + (1-\rho_l)U(l)$;

2. We define $\hat{w}_h$ and $\hat{w}_l$ given $\hat{W}(U, 1-\rho_h)$ and $\hat{W}(U, \rho_l)$ in the same way as in (II). The functions $\hat{w}_h$ and $\hat{w}_l$ satisfy (II.a)-(II.c).

We next show that, if we use some $\hat{W} \in \mathcal{W}$ as the principal's continuation value function, then the new value function, denoted by $\hat{\hat{W}}$, still satisfies properties (I) and (II). On the other hand, properties (I) and (II) are closed under value function iterations. This proves that the value function of the relaxed problem, $\tilde{W}$, is in $\mathcal{W}$.

1. We define $\hat{\hat{w}}_h$ and $\hat{\hat{w}}_l$ given $\hat{\hat{W}}(U, 1-\rho_h)$ and $\hat{\hat{W}}(U, \rho_l)$ in the same way as in (II). We first show that $\hat{\hat{w}}_h$ and $\hat{\hat{w}}_l$ satisfy properties (II.a)-(II.c). For any $U \in P_f \setminus (\underline{V} \cup V_h)$, a high report leads to U_h such that $(1-\rho_h)U_h(h) + \rho_h U_h(l) = (U(h) - (1-\delta)h)/\delta$ and U_h is lower than U. Also, a low report leads to U_l such that $\rho_l U_l(h) + (1-\rho_l)U_l(l) = U(l)/\delta$ and U_l is higher than U given that $U \in P_f \setminus (\underline{V} \cup V_h)$. We thus have

$$\hat{\hat{W}}(U, 1-\rho_h) = (1-\rho_h)\Big((1-\delta)(h-c) + \delta\hat{W}(U_h, 1-\rho_h)\Big) + \rho_h\delta\hat{W}(U_l, \rho_l),$$
$$\hat{\hat{W}}(U, \rho_l) = \rho_l\Big((1-\delta)(h-c) + \delta\hat{W}(U_h, 1-\rho_h)\Big) + (1-\rho_l)\delta\hat{W}(U_l, \rho_l).$$

Let $x = X(U)$ and $y = Y(U)$. It follows that

$$(U_{h1}(x), U_{l1}(x)) = (U_{h2}(y), U_{l2}(y)) = (U(h), U(l)). \tag{15}$$

Given the definition of $\hat{\hat{w}}_h, \hat{\hat{w}}_l, \hat{w}_h, \hat{w}_l$, we have

$$\begin{aligned}
\hat{\hat{w}}_h'(x) &= (1-\rho_h)U_{h1}'(x)\hat{w}_h'\left(\frac{U_{h1}(x)-(1-\delta)h}{\delta}\right) + \rho_h U_{l1}'(x)\hat{w}_l'\left(\frac{U_{l1}(x)}{\delta}\right), \\
\hat{\hat{w}}_l'(y) &= \rho_l U_{h2}'(y)\hat{w}_h'\left(\frac{U_{h2}(y)-(1-\delta)h}{\delta}\right) + (1-\rho_l)U_{l2}'(y)\hat{w}_l'\left(\frac{U_{l2}(y)}{\delta}\right).
\end{aligned} \tag{16}$$

Given that $\hat{w}_h'(X(U)) \geqslant \hat{w}_l'(Y(U))$ for any U, that $\hat{w}_h, \hat{w}_l$ are concave, and that U_h is lower than U_l on P_f, we have that

$$\hat{w}_h'\left(\frac{U(h)-(1-\delta)h}{\delta}\right) \geqslant \hat{w}_l'\left(\frac{U(l)}{\delta}\right). \tag{17}$$

We want to show that for any $U \in P_f$ and $x = X(U), y = Y(U)$

$$(1-\rho_h)U_{h1}'(x) + \rho_h U_{l1}'(x) = \rho_l U_{h2}'(y) + (1-\rho_l)U_{l2}'(y) = 1, \tag{18}$$

$$(1-\rho_h)U_{h1}'(x) - \rho_l U_{h2}'(y) \geqslant 0. \tag{19}$$

This can be shown by assuming that U is on some line segment $U(h) = aU(l) + b$. (Recall that P_f is piecewise linear.) For any $a > 0$, the conditions (18) and (19) hold.

Given (16), (18), and $\hat{w}_h', \hat{w}_l' \in [1-c/l, 1-c/h]$, we obtain that $\hat{\hat{w}}_h', \hat{\hat{w}}_l' \in [1-c/l, 1-c/h]$. Given (15) to (19), we obtain that $\hat{\hat{w}}_h'(X(U)) \geqslant \hat{\hat{w}}_h'(Y(U))$ for any $U \in P_f \setminus (\underline{V} \cup V_h)$.

The concavity of $\hat{\hat{w}}_h, \hat{\hat{w}}_l$ can be shown by taking the derivative of (16):

$$\begin{aligned}
\hat{\hat{w}}_h''(x) &= (1-\rho_h)U_{h1}'(x)\hat{w}_h''\left(\frac{U_{h1}(x)-(1-\delta)h}{\delta}\right)\frac{U_{h1}'(x)}{\delta} + \rho_h U_{l1}'(x)\hat{w}_l''\left(\frac{U_{l1}(x)}{\delta}\right)\frac{U_{l1}'(x)}{\delta}, \\
\hat{\hat{w}}_l''(y) &= \rho_l U_{h2}'(y)\hat{w}_h''\left(\frac{U_{h2}(y)-(1-\delta)h}{\delta}\right)\frac{U_{h2}'(y)}{\delta} + (1-\rho_l)U_{l2}'(y)\hat{w}_l''\left(\frac{U_{l2}(y)}{\delta}\right)\frac{U_{l2}'(y)}{\delta}.
\end{aligned}$$

Here, we use the fact that $U_{h1}(x), U_{l1}(x)$ (or $U_{h2}(y), U_{l2}(y)$) are piece-wise linear in x (or y).

The analysis for $U \in P_f \cap V_h$ is similar. The only difference is that U_l equals (μ_h, μ_l) and that the good is supplied with positive probability after the low report. We omit the details for this case.

To sum up, if $\hat{w}_h, \hat{w}_l$ satisfy properties (II.a), (II.b) and (II.c), then $\hat{\hat{w}}_h, \hat{\hat{w}}_l$ also satisfy these properties.

2. We next show that $\hat{\hat{W}}$ satisfies (I). We first focus on the region $U \in V \setminus (\underline{V} \cup V_h \cup V_l)$. We need to show, for $\phi \in \{1-\rho_h, \rho_l\}$, that:

$$\hat{\hat{W}}(U(h)+\varepsilon, U(l), \phi) - \hat{\hat{W}}(U, \phi) \geqslant \hat{\hat{W}}\left(U(h), U(l) + \frac{\phi}{1-\phi}\varepsilon, \phi\right) - \hat{\hat{W}}(U, \phi), \text{ where } \varepsilon > 0. \tag{20}$$

The left-hand side of (20) equals

$$\delta\phi\left(\hat{W}\left(\tilde{U}_h(h), \tilde{U}_h(l), 1-\rho_h\right) - \hat{W}(U_h(h), U_h(l), 1-\rho_h)\right), \tag{21}$$

where $\tilde{U}_h$ and U_h are on P_f and

$$(1-\delta)h + \delta\left((1-\rho_h)\tilde{U}_h(h) + \rho_h\tilde{U}_h(l)\right) = U(h) + \varepsilon,$$
$$(1-\delta)h + \delta\left((1-\rho_h)U_h(h) + \rho_h U_h(l)\right) = U(h).$$

The right-hand side of (20) equals

$$\delta(1-\phi)\left(\hat{W}\left(\tilde{U}_l(h), \tilde{U}_l(l), \rho_l\right) - \hat{W}(U_l(h), U_l(l), \rho_l)\right), \tag{22}$$

where $\tilde{U}_l$ and U_l are on P_f and

$$\delta\left(\rho_l\tilde{U}_l(h) + (1-\rho_l)\tilde{U}_l(l)\right) = U(l) + \frac{\phi}{1-\phi}\varepsilon,$$
$$\delta\left(\rho_l U_l(h) + (1-\rho_l)U_l(l)\right) = U(l).$$

We need to show that (21) is greater than (22). Note that $U_h, \tilde{U}_h, U_l, \tilde{U}_l$ are on P_f, so only the properties of $\hat{w}_h, \hat{w}_l$ are needed. Taking the limit as ε goes to 0, we obtain that (20) is equivalent to

$$\hat{w}_h'\left(\frac{U(h)-(1-\delta)h}{\delta}\right) \geqslant \hat{w}_l'\left(\frac{U(l)}{\delta}\right), \quad \forall U \in V \setminus (\underline{V} \cup V_h \cup V_l). \tag{23}$$

Given that $\hat{w}_h, \hat{w}_l$ are concave, $\hat{w}_h', \hat{w}_l'$ are decreasing. Therefore, we only need to show that inequality (23) holds when U is on P_f, since any $U \in P_f$ gives the highest $U(h)$ (among vectors in V) for a fixed $U(l)$. The inequality (23) holds given that (i) $\hat{w}_h, \hat{w}_l$ are concave; (ii) inequality (14) holds; (iii) $(U(h)-(1-\delta)h)/\delta$ corresponds to a lower point on P_f than $U(l)/\delta$ does. When $U \in V_h$, the right-hand side of (20) is given by

$\phi\varepsilon(1-c/l)$. Inequality (20) is equivalent to $\hat{w}_h'((U(h)-(1-\delta)h)/\delta) \geqslant 1-c/l$, which is obviously true. Similar analysis applies to the case in which $U \in V_l$.

This shows that the value function for the relaxed problem $\tilde{W}$ is in $\mathcal{W}$. If the initial promise is on P_f, then the continuation utility never leaves P_f. Moreover, for $U \in P_f$, IC_h never binds under the optimal policy of the relaxed problem. This implies that $W = \tilde{W}$ for any $U \in P_f$.

Part II: We turn to the original optimization problem by adding back the constraint IC_h. The first observation is that we can decompose the optimization problem into two sub-problems: (i) choose p_h, U_h to maximize $(1-\delta)p_h(h-c)+\delta W(U_h, 1-\rho_h)$ subject to PK_h and IC_l; (ii) choose p_l, U_l to maximize $(1-\delta)p_l(l-c)+\delta W(U_l, \rho_l)$ subject to PK_l and IC_h.

1. We want to show that the policy in Definition 2 with respect to p_h, U_h is the optimal solution to the first sub-problem. First, we have shown at the beginning of this proof that it is feasible to choose U_h on P_f so that PK_h and IC_l are satisfied. If we take $\tilde{W}$ as the continuation value function, the policy in Definition 2 is optimal. Given that (i) $\tilde{W}$ is weakly higher than W pointwise, and (ii) $\tilde{W}$ coincides with W on P_f, the policy in Definition 2 with respect to p_h, U_h solves the first sub-problem.

2. For the second sub-problem, if the principal chooses p_l, U_l subject to PK_l so that $U_l \in P_f$ and p_l is as low as possible, then IC_h will not bind for $U \in V_f$ and will bind for $U \in V_b$. (This follows from the definition of P, V_f, V_b.) This shows that for any $U \in V_f$, the policy in Definition 2 is optimal, because (i) $\tilde{W}$ is weakly higher than W pointwise, and (ii) $\tilde{W}$ coincides with W on P_f. This also implies that for any $U \in V_b$, IC_h must bind. We want to show that for $U \in V_b$, it is optimal to choose the lowest possible p_l.

   For a fixed $U \in V_b$, PK_l and a binding IC_h determines $U_l(h), U_l(l)$ as a function of p_l. Let γ_h, γ_l denote the derivative of $U_l(h), U_l(l)$ with respect to p_l

   $$\gamma_h = \frac{(1-\delta)(l\rho_h - h(1-\rho_l))}{\delta(1-\rho_h-\rho_l)}, \quad \gamma_l = \frac{(1-\delta)(h\rho_l - l(1-\rho_h))}{\delta(1-\rho_h-\rho_l)}.$$

   Given that $\rho_l + \rho_h < 1$, it follows that $\gamma_h < 0$. We want to show that it is optimal to set p_l as low as possible. That is, the principal's payoff from the low type, $(1-\delta)p_l(l-c)+\delta W(U_l, \rho_l)$, decreases in p_l. The principal's payoff decreases in p_l if and only if:

   $$\gamma_h \frac{\partial W(U, \rho_l)}{\partial U(h)} + \gamma_l \frac{\partial W(U, \rho_l)}{\partial U(l)} \leqslant \frac{(1-\delta)(c-l)}{\delta}, \tag{24}$$

where the left-hand side is the directional derivative of $W(U, \rho_l)$ along the vector (γ_h, γ_l). We first show that (24) holds for all $U \in P_f \setminus (V_h \cup V_l)$. For any fixed $U \in P_f \setminus (V_h \cup V_l)$, we have

$$W(U, \rho_l) = \rho_l \left((1-\delta)(h-c) + \delta \tilde{w}_h \left(\frac{U(h) - (1-\delta)h}{\delta} \right) \right) + (1-\rho_l)\delta \tilde{w}_l \left(\frac{U(l)}{\delta} \right).$$

It follows that $\partial W / \partial U(h) = \rho_l \tilde{w}_h'$ and $\partial W / \partial U(l) = (1-\rho_l)\tilde{w}_l'$. We have shown that $\tilde{w}_h' \geqslant \tilde{w}_l'$ and $\tilde{w}_h', \tilde{w}_l' \in [1 - c/l, 1 - c/h]$. Given the linearity of (24) in $\tilde{w}_h'$ and $\tilde{w}_l'$, it is sufficient to show that (24) holds when either (i) $\tilde{w}_l' = \tilde{w}_h'$, or (ii) $\tilde{w}_l' = 1 - c/l$. In case (i), the left-hand side of (24) equals $-(1-\delta)l\tilde{w}_h'/\delta$, so (24) holds for any $\tilde{w}_h' \in [1 - c/l, 1 - c/h]$. In case (ii), the left-hand side of (24) decreases in $\tilde{w}_h'$ and is maximized when $\tilde{w}_h' = 1 - c/l$. Substituting $\tilde{w}_l' = \tilde{w}_h' = 1 - c/l$, we obtain that (24) holds. (Using similar arguments, we can show that (24) holds for all $U \in P_f \cap V_h$ or $U \in P_f \cap V_l$.)

Lastly, we want to show that (24) holds for $U \in V_b$. Note that $W(U, \rho_l)$ is concave on V. Therefore, its directional derivative along the vector (γ_h, γ_l) is monotone. For any fixed U on P_f, we have

$$\begin{aligned}
&\lim_{\varepsilon \to 0} \frac{\gamma_h \frac{\partial W(U(h)+\gamma_h \varepsilon, U(l)+\gamma_l \varepsilon, \rho_l)}{\partial U(h)} + \gamma_l \frac{\partial W(U(h)+\gamma_h \varepsilon, U(l)+\gamma_l \varepsilon, \rho_l)}{\partial U(l)} - \left(\gamma_h \frac{\partial W(U, \rho_l)}{\partial U_h} + \gamma_l \frac{\partial W(U, \rho_l)}{\partial U(l)} \right)}{\varepsilon} \\
=&{\gamma_h}^2 \frac{\rho_l}{\delta} \tilde{w}_h'' \left(\frac{U(h) - (1-\delta)h}{\delta} \right) + {\gamma_l}^2 \frac{1-\rho_l}{\delta} \tilde{w}_l'' \left(\frac{U(l)}{\delta} \right) \leqslant 0.
\end{aligned}$$

The last inequality follows as $\tilde{w}_h, \tilde{w}_l$ are concave. Given that (γ_h, γ_l) points towards the interior of V, (24) holds within V. This part of the proof also shows that W coincides with $\tilde{W}$ on V_f. Since W is concave on V, W also satisfies property (I).

Part III: We next show that the optimal initial promise is on P_f and increases in the prior belief of the high type. For any $x \in [0, (1-\rho_h)\mu_h + \rho_h \mu_l]$, let $z(x)$ be $\rho_l U_{h1}(x) + (1 - \rho_l)U_{l1}(x)$. The function $z(x)$ is piecewise linear with z' being positive and increasing in x. Let ϕ_0 denote the prior belief of the high type. We want to show that the maximum of $\phi_0 W(U, 1-\rho_h) + (1-\phi_0)W(U, \rho_l)$ is achieved on P_f for any prior ϕ_0. Suppose not. Suppose $\tilde{U} \in V \setminus P_f$ achieves the maximum. Let U^0 (or U^1) denote the intersection of P_f and $(1-\rho_h)U(h) + \rho_h U(l) = (1-\rho_h)\tilde{U}(h) + \rho_h \tilde{U}(l)$ (or $\rho_l U(h) + (1-\rho_l)U(l) = \rho_l \tilde{U}(h) + (1-\rho_l)\tilde{U}(l)$). Given that $\rho_l < 1 - \rho_h$, it follows that $U^0 < U^1$. Given that $\tilde{U}$ achieves the maximum, it must be true that

$$W(U^1, 1-\rho_h) - W(U^0, 1-\rho_h) < 0,$$
$$W(U^1, \rho_l) - W(U^0, \rho_l) > 0.$$

We show that this is impossible by arguing that for any $U^0, U^1 \in P_f$ and $U^0 < U^1$, $W(U^1, 1-\rho_h) - W(U^0, 1-\rho_h) < 0$ implies that $W(U^1, \rho_l) - W(U^0, \rho_l) < 0$. It is without loss to assume that U^0, U^1 are on the same line segment $U(h) = aU(l) + b$. Hence,

$$W(U^1, 1-\rho_h) - W(U^0, 1-\rho_h) = \int_{s^0}^{s^1} \tilde{w}_h'(s)\mathrm{d}s,$$
$$W(U^1, \rho_l) - W(U^0, \rho_l) = z'(s) \int_{s^0}^{s^1} \tilde{w}_l'(z(s))\mathrm{d}s,$$

where $s^0 = (1-\rho_h)U^0(h) + \rho_h U^0(l)$, $s^1 = (1-\rho_h)U^1(h) + \rho_h U^1(l)$. Given that $\tilde{w}_h'(s) \geqslant \tilde{w}_l'(z(s))$ and $z'(s) > 0$, $\int_{s^0}^{s^1} \tilde{w}_h'(s)\mathrm{d}s < 0$ implies $z'(s) \int_{s^0}^{s^1} \tilde{w}_l'(z(s))\mathrm{d}s < 0$. This shows the optimal initial promise is on P_f.

The optimal U^* is chosen such that $X(U^*)$ maximizes $\phi_0 \tilde{w}_h(x) + (1-\phi_0)\tilde{w}_l(z(x))$ which is concave in x. Thus, at $x = X(U^*)$,

$$\phi_0 \tilde{w}_h'(X(U^*)) + (1-\phi_0)\tilde{w}_l'(z(X(U^*)))z'(X(U^*)) = 0.$$

According to (14), we know that $\tilde{w}_h'(X(U^*)) \geqslant 0 \geqslant \tilde{w}_l'(z(X(U^*)))$. Therefore, the derivative above is weakly positive for any $\phi_0' > \phi_0$. Hence, the optimal initial promise given ϕ_0' must be higher than U^*, the optimal initial promise given ϕ_0. ■

Online appendix

"Dynamic Mechanisms without Money"
by Yingni Guo and Johannes Hörner

C Omitted Proofs

C.1 Complete Information

We review the solution under complete information by dropping (IC_h) and (IC_l).

If we ignore promises, the efficient policy is to supply the good if and only if the value is h. The corresponding utility pair, denoted by (v_h^*, v_l^*), satisfies

$$v_h^* = (1-\delta)h + \delta\left\{(1-\rho_h)v_h^* + \rho_h v_l^*\right\}, \quad v_l^* = \delta\left\{(1-\rho_l)v_l^* + \rho_l v_h^*\right\},$$

which yields

$$v_h^* = \frac{h(1-\delta(1-\rho_l))}{1-\delta(1-\rho_h-\rho_l)}, \quad v_l^* = \frac{\delta h \rho_l}{1-\delta(1-\rho_h-\rho_l)}.$$

For a fixed $U = (U(h), U(l))$, the problems of delivering $U(h)$ to a current high type and $U(l)$ to a current low type are uncoupled. If $U(h) \leqslant v_h^*$, the good is supplied only when the value is high. When $U(h) \geqslant v_h^*$, the good is always supplied when the value is high, and sometimes supplied when the value is low. We proceed analogously to deliver $U(l)$. In summary, the resulting value function, $\overline{W}(U,\phi)$ is given by:

$$\begin{cases} \phi\frac{U(h)(h-c)}{h} + (1-\phi)\frac{U(l)(h-c)}{h} & \text{if } U \in [0, v_h^*] \times [0, v_l^*], \\ \phi\frac{U(h)(h-c)}{h} + (1-\phi)\left(\frac{v_l^*(h-c)}{h} + \frac{(U(l)-v_l^*)(l-c)}{l}\right) & \text{if } U \in [0, v_h^*] \times [v_l^*, \mu_l], \\ \phi\left(\frac{v_h^*(h-c)}{h} + \frac{(U(h)-v_h^*)(l-c)}{l}\right) + (1-\phi)\frac{U(l)(h-c)}{h} & \text{if } U \in [v_h^*, \mu_l] \times [0, v_l^*], \\ \phi\left(\frac{v_h^*(h-c)}{h} + \frac{(U(h)-v_h^*)(l-c)}{l}\right) + (1-\phi)\left(\frac{v_l^*(h-c)}{h} + \frac{(U(l)-v_l^*)(l-c)}{l}\right) & \text{if } U \in [v_h^*, \mu_l] \times [v_l^*, \mu_l]. \end{cases}$$

The derivative of $\overline{W}$ (differentiable except at $U(h) = v_h^*$ or $U(l) = v_l^*$) is in the interval $[1-c/l, 1-c/h]$, as expected. The latter corresponds to the most efficient allocation, whereas the former corresponds to the most inefficient allocation.

C.2 Transfers with Limited Liability

Here, we consider the case in which transfers are allowed but the agent is protected by limited liability. Therefore, only the principal can pay the agent. The principal maximizes his payoff net of payments. The following lemma shows that transfers occur on the equilibrium path when $c - l$ is higher than l.

Lemma 7. *The principal makes transfers on path if and only if $c - l > l$.*

Proof. We first show that the principal makes transfers if $c-l > l$. Suppose not. The optimal mechanism is the same as the one characterized in Theorem 1. When U is sufficiently close to μ, we want to show that it is "cheaper" to provide incentives using transfers. Given the optimal allocation (p_h, U_h) and (p_l, U_l), if we reduce U_l by ε and make a transfer of $\delta\varepsilon/(1-\delta)$ to the low type, the IC_l/PK constraints are satisfied. When U_l is sufficiently close to μ, the principal's payoff increment is close to $\delta(c/l-1)\varepsilon - \delta\varepsilon = \delta(c/l-2)$, which is strictly positive if $c - l > l$. This contradicts the fact that the allocation (p_h, U_h) and (p_l, U_l) is optimal. Therefore, the principal makes transfers if $c - l > l$.

If $c - l \leqslant l$, we first show that the principal never makes transfers if $U_l, U_h < \mu$. With abuse of notation, let t_m denote the current-round transfer after m report. Suppose $u_m < \mu$ and $t_m > 0$. We can increase u_m ($m = l$ or h) by ε and reduce t_m by $\delta\varepsilon/(1 - \delta)$. This adjustment has no impact on $IC_h/IC_l/PK$ constraints and strictly increases the principal's payoff given that $W'(U) > 1 - c/l$ when $U < \mu$.[29] Suppose $U_l = \mu$ and $t_l > 0$. We can always replace p_l, t_l with $p_l + \varepsilon, t_l - \varepsilon l$. This adjustment has no impact on $IC_h/IC_l/PK$ and (weakly) increases the principal's payoff. If $U_l = \mu, p_l = 1$, we know that the promised utility to the agent is at least μ. The optimal scheme is to provide the unit forever. ∎

D Continuous Time

We have opted for a discrete time framework because it embeds the case of independent values—a natural starting point for which there is no counterpart in continuous time. This choice comes at a price. With Markovian types, closed-forms (and comparative statics) are not available in discrete time. As we show, they are in continuous time. We consider the continuous-time limit of the problem, by scaling the transition probabilities according to the

[29] It is easy to show that the principal's complete-information payoff, if $U \in [0, \mu]$ and $c - l \leqslant l$, is the same as $\overline{W}$ in Lemma 1.

usual Poisson limit. We start with an overview of this model and its results, relegating some of the details and proofs to Section D.1.

The round length is $\Delta > 0$. The transition probabilities are set to $\rho_h = \lambda_h \Delta + o(\Delta)$, $\rho_l = \lambda_l \Delta + o(\Delta)$. The agent's utility from consuming the good is $h\Delta$ if his value is h and $l\Delta$ if his value is l. The supply cost is $c\Delta$. We take Δ to 0 and obtain the limiting stochastic process of the agent's value. Formally, this is a continuous-time Markov chain $(v_t)_{t\geqslant 0}$ (by definition, a right-continuous process) with state space $\{l, h\}$, transition matrix $((-\lambda_l, \lambda_l), (\lambda_h, -\lambda_h))$, and initial probability $q = \lambda_l/(\lambda_h + \lambda_l)$ of h. Let $T_0 = 0$ and $T_1, T_2, \ldots$ be the random times at which the value switches. We set $T_1 = 0$ if the initial value is h such that, by convention, $v_t = l$ on any interval $[T_{2k}, T_{2k+1})$ and $v_t = h$ on $[T_{2k+1}, T_{2k+2})$, for any $k \in \mathbf{N}$.

The incentive-feasible set V is characterized by the frontloading and backloading policies. By taking the limit (as $\Delta \to 0$) of the formulas for $\{f^\nu, b^\nu\}_{\nu \in \mathbf{N}}$, we obtain the boundaries of V. In particular, the frontloading boundary is given in parametric form by[30]

$$f_h^z = \mathbf{E}\left[\int_0^z e^{-rt} v_t \mathrm{d}t \mid v_0 = h\right], \quad f_l^z = \mathbf{E}\left[\int_0^z e^{-rt} v_t \mathrm{d}t \mid v_0 = l\right],$$

where $z \geqslant 0$ can be interpreted as the length of time that the agent is entitled to claim the good with "no questions asked."[31]

Similar to the discrete-time case, the principal might achieve the complete-information payoff for some utility pairs on the frontloading boundary. Let $\underline{U}$ be the largest intersection of the frontloading boundary $\{(f_h^z, f_l^z)\}_{z\geqslant 0}$ with the line $U(l) = \frac{\lambda_l}{\lambda_l + r} U(h)$. It is immediately verifiable that $\underline{U} > 0$ if and only if (cf. (4))

$$\frac{h - l}{l} > \frac{r}{\lambda_l}.$$

When the low type is too persistent, *i.e.*, λ_l is too small, the complete-information payoff cannot be achieved (except for $0, (\mu_h, \mu_l)$). It can be achieved for some promised utility pairs when the agent is sufficiently patient.

[30]Here and in what follows, we omit the obvious corresponding analytic expressions. See Appendix D.1.

[31]The utility pair from getting the good forever is $(\mu_h, \mu_l) = \lim_{z\to\infty}(f_h^z, f_l^z)$. The backloading boundary is given in parametric forms by

$$b_h^z = \mathbf{E}\left[\int_z^\infty e^{-rt} v_t \mathrm{d}t \mid v_0 = h\right], \quad b_l^z = \mathbf{E}\left[\int_z^\infty e^{-rt} v_t \mathrm{d}t \mid v_0 = l\right].$$

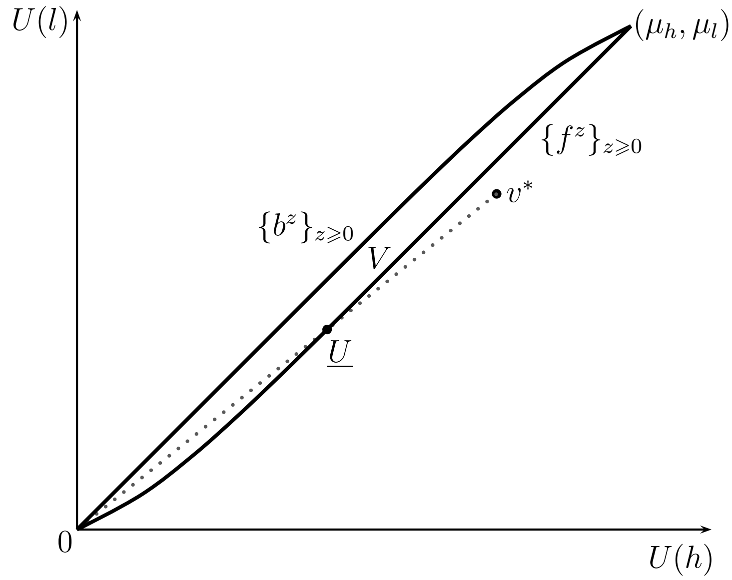

Figure 7: Incentive-feasible set V for $(\lambda_l, \lambda_h, r, l, h) = (1, 1, 1/4, 1, 5/2)$.

The optimal policy is given by a continuous-time process $(z_t)_{t\geqslant 0}$, where z_t is the promised length of provision at time t. On any interval $[T_{2k+1}, T_{2k+2})$ over which h is continuously reported, $(z_t)_{t\geqslant 0}$ evolves deterministically, with increments

$$\frac{\mathrm{d}z_t}{\mathrm{d}t} = -1.$$

Instead, on any interval $[T_{2k}, T_{2k+1})$ over which l is continuously reported, the evolution is given by

$$\frac{\mathrm{d}z_t}{\mathrm{d}t} = \frac{l}{\mathbf{E}[e^{-rz_t}v_{z_t} \mid v_0 = l]} - 1.$$

This increment $\mathrm{d}z_t$ is positive or negative depending upon whether z_t maps onto a utility vector above or below $\underline{U}$. The interpretation is as in discrete time: whether the agent asks for the unit or not, he must pay a fixed cost of 1 in terms of z_t. However, conditional on reporting l and foregoing the unit, he is compensated by a term that equals the rate of substitution between the opportunity flow cost of giving up the current unit, l, and the value of getting it at the future time z_t. The agent is compensated at the end of the time z_t, because his expected value, conditional on being low today, increases in time. The interests of the principal and the low type grow more congruent in time.

The optimal policy defines a tuple of continuous-time processes that follow deterministic trajectories over any interval $[T_k, T_{k+1})$. First, the belief $(\phi_t)_{t\geqslant 0}$ of the principal takes values in the set $\{0, 1\}$. Namely, $\phi_t = 0$ over any interval $[T_{2k}, T_{2k+1})$, and $\phi_t = 1$ otherwise.

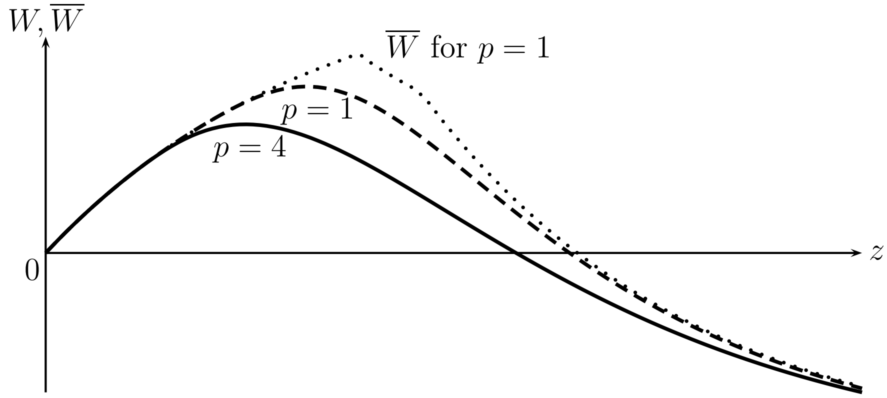

Figure 8: Value function and complete-information payoff as a function of z, for $(\lambda_l, \lambda_h, r, l, h, c) = (1/p, 1/p, 1/4, 1, 5/2, 2)$ and $p = 1, 4$.

Second, the utilities of the agent $(U_t(h),\, U_t(l))_{t\geqslant 0}$ are functions of his type and the promise z_t. Finally, the expected payoff of the principal, $(W_t)_{t\geqslant 0}$, is computed according to his belief ϕ_t and the promise to the agent.

Continuous time allows to explicitly solve for the principal's value function. Because her belief is degenerate, except at the initial instant, we write $W_h(z)$ (resp., $W_l(z)$) for the payoff when (she assigns probability one to the event that) the agent's value is currently high (resp. low). See Appendix D.1 for a derivation of the solution. This closed-form solution allows for explicit comparative statics. Here, we will restrict ourselves with illustrations.

Figure 8 illustrates the value function for two levels of persistence $p = 1, 4$ and compares it to the complete-information payoff $\overline{W}$ evaluated along the frontloading boundary when $p = 1$.[32]

What about the agent's utility? Persistence plays an ambiguous role in determining the agent's utility: perfect persistence is his favorite outcome if $\mu > c$, because always providing the good is the principal's best option. Conversely, perfect persistence is the agent's worst outcome if $\mu < c$. Hence, persistence tends to improve the agent's situation when $\mu > c$. However, this convergence is not necessarily monotone, which is easy to verify via examples.

As $r \to 0$, the principal's value converges to the complete-information payoff $q(h - c)$. We conclude with a rate of convergence which is higher than the convergence rate of token mechanisms as in Jackson and Sonnenschein (2007).

Lemma 8. *It holds that*

$$|\max_z W(z) - q(h - c)| = \mathcal{O}(r).$$

[32]The value function is concave on the agent's promised utility pair but not on z.

Lastly, we examine the impact of varying h or l on the value function $W(z)$, for any fixed λ_h, λ_l, r and c. We make three observations. First, we fix the expected value of the unit (*i.e.*, $\mu = qh + (1-q)l$) and vary h. Simulations show that the value function $W(z)$ increases pointwise in h. (See Figure 9.) As we increase the high type's value h, the low type's value l decreases. The agent's preferences become more congruent with the principal's. Hence, for any promised length of provision z, the principal's payoff is higher.

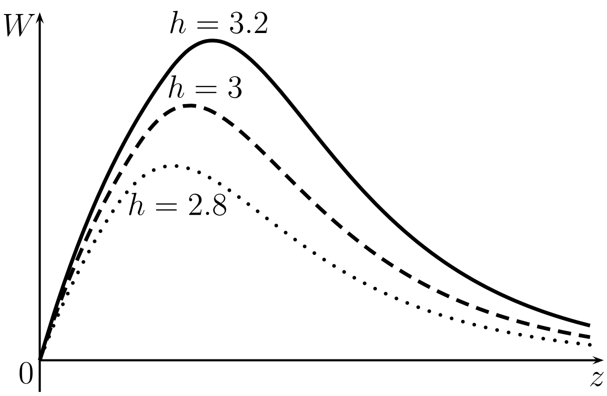

Figure 9: Value function for $(\lambda_l, \lambda_h, r, c) = (1, 1, 1/2, 2)$ and $\mu = 2$.

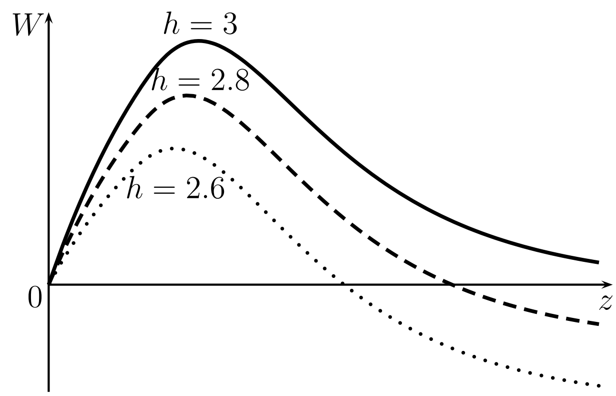

Figure 10: Value function for $(\lambda_l, \lambda_h, r, c) = (1, 1, 1/2, 2)$ and $l = 1$.

Second, we fix l and vary h. The value function $W(z)$ increases pointwise in h. This is shown in Figure 10. Increasing h while fixing l has two effects. First, a unit is on average worth more to the agent, so the payoff to the principal from providing the unit also increases. Second, the agent's preferences become more aligned with the principal's. Both effects suggest that the principal's value shall be higher. This is confirmed by the simulations.

Third, we fix h and vary l. The principal's maximal payoff is not monotone in l. This can be seen from the example in Figure 11. Increasing l has two counteracting effects. On the one hand, a unit is on average worth more to the agent. On the other, the low type is more tempted to claim the unit. The example shows that, among three low values, the principal's payoff is the highest when $l = .2$. It decreases as l increases to 1, and then increases as l increases to 1.8.

D.1 Continuous Time: Details and Proofs

We first derive the formulas for the frontloading boundary by considering the Poisson limit. Write $\rho_h = \lambda_h \Delta, \rho_l = \lambda_l \Delta$, $\delta = e^{-r\Delta}$. We fix the total length of provision z and take Δ to

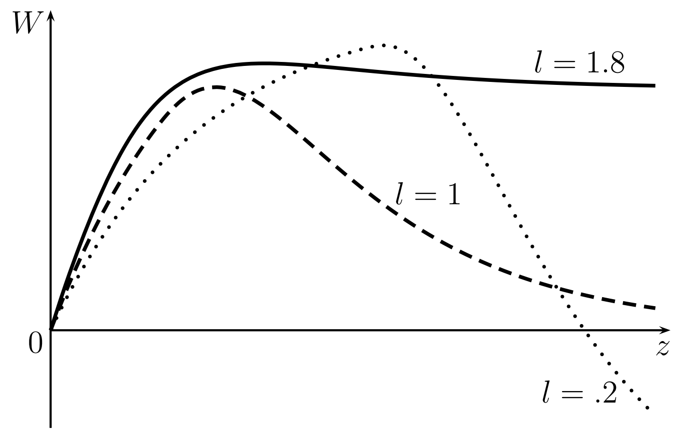

Figure 11: Value function for $(\lambda_l, \lambda_h, r, c) = (1, 1, 1/2, 2)$ and $h = 3$.

zero. The frontloading boundary is smooth, and is given in parametric forms by

$$f_h^z = (1 - e^{-rz})\mu_h + e^{-rz}(1 - e^{-(\lambda_h+\lambda_l)z})(1 - q)(\mu_h - \mu_l),$$
$$f_l^z = (1 - e^{-rz})\mu_l - e^{-rz}(1 - e^{-(\lambda_h+\lambda_l)z})q(\mu_h - \mu_l),$$

where $q = \lambda_l/(\lambda_h + \lambda_l)$ is the invariant probability of h, and (μ_h, μ_l) is the agent's utility pair from getting the good forever:

$$(\mu_h, \mu_l) = \left(h - \frac{\lambda_h(h-l)}{\lambda_h + \lambda_l + r}, l + \frac{\lambda_l(h-l)}{\lambda_h + \lambda_l + r}\right).$$

If the good is supplied if and only if the agent's value is high, the agent's utility pair is

$$(v_h^*, v_l^*) = \left(\frac{h\,(\lambda_l + r)}{\lambda_h + \lambda_l + r}, \frac{h\lambda_l}{\lambda_h + \lambda_l + r}\right).$$

Given the optimal policy, let $\mathrm{d}z_l$ and $\mathrm{d}z_h$ be the incremental change in z_t when the agent's report is low and high, respectively. Define the function

$$g(z) := q(h-l)e^{-(\lambda_h+\lambda_l)z} + le^{rz} - \mu,$$

so that upon direct calculation,

$$\mathrm{d}z_l := \frac{g(z)}{\mu - q(h-l)e^{-(\lambda_h+\lambda_l)z}}\mathrm{d}t,$$

where $\mu = qh + (1-q)l$, as before. If $\underline{V}$ has a nonempty interior, we can identify the value of z that is the intersection of the line $U(l) = \frac{\lambda_l}{\lambda_l + r} U(h)$ and the frontloading boundary $\{f^z\}_{z \geqslant 0}$; call this value $\underline{z}$, which is simply the positive root (if any) of g. Otherwise, set $\underline{z} = 0$. Also, recall that $\mathrm{d}z_h = -1$.

We now motivate the derivation of the differential equations solved by W_h, W_l. Given the optimal policy, the value functions solve the following paired system of equations:

$$W_h(z) = r\mathrm{d}t(h-c) + \lambda_h \mathrm{d}t W_l(z) + (1 - r\mathrm{d}t - \lambda_h \mathrm{d}t) W_h(z + \mathrm{d}z_h) + \mathcal{O}(\mathrm{d}t^2),$$

and

$$W_l(z) = \lambda_l \mathrm{d}t W_h(z) + (1 - r\mathrm{d}t - \lambda_l \mathrm{d}t) W_l(z + \mathrm{d}z_l) + \mathcal{O}(\mathrm{d}t^2).$$

Assume for now (as will be verified) that the functions W_h, W_l are twice differentiable. We then obtain the differential equations[33]

$$(r + \lambda_h) W_h(z) = r(h-c) + \lambda_h W_l(z) - W_h'(z),$$

and

$$(r + \lambda_l) W_l(z) = \lambda_l W_h(z) + \frac{g(z)}{\mu - q(h-l)e^{-(\lambda_h + \lambda_l)z}} W_l'(z).$$

We directly work with the expected payoff $W(z) = qW_h(z) + (1-q)W_l(z)$. Using the system of differential equations, we get

$$\begin{aligned} & g(z)((r + \lambda_h)W'(z) + W''(z)) \\ =& (h-l)q\lambda_h e^{-(\lambda_h + \lambda_l)z} W'(z) + \mu(r(\lambda_h + \lambda_l)W(z) + \lambda_l W'(z) - r\lambda_l(h-c)). \end{aligned} \tag{25}$$

The value function W must satisfy the following boundary conditions. First, at $z = \underline{z}$, the value must coincide with the complete-information payoff $\overline{W}$ in that range. We let $\overline{W}_1$ denote the complete-information payoff for those z such that $(f_h^z, f_l^z) \in [0, v_h^*] \times [0, v_l^*]$. Substituting f_h^z, f_l^z and simplifying, we obtain the formula for $\overline{W}_1(z)$:

$$\overline{W}_1(z) = qf_h^z \frac{h-c}{h} + (1-q)f_l^z \frac{h-c}{h} = (1 - e^{-rz})(1 - c/h)\mu.$$

We must have $W(\underline{z}) = \overline{W}(\underline{z}) = \overline{W}_1(\underline{z})$. Second, as $z \to \infty$, it must hold that the payoff

[33] To be clear, these are not HJB equations, as there is no need to verify the optimality of the policy that is being followed. This fact has already been established. The functions must satisfy these simple recursive equations.

$\mu - c$ is approached. Hence,

$$\lim_{z\to\infty} W(z) = \mu - c.$$

Let $\hat{z}$ denote the positive root of

$$\chi(z) := \mu e^{-rz} - (1-q)l.$$

As is easy to see, this root always exists and is strictly above $\underline{z}$, with $\chi(z) > 0$ iff $z < \hat{z}$. Finally, let

$$\psi(z) := r - (\lambda_h + \lambda_l)\frac{\chi(z)}{g(z)} e^{rz}.$$

It is then straightforward to verify (though not quite as easy to obtain) that:[34]

Proposition 1. *The value function of the principal is given by*

$$W(z) = \begin{cases} \overline{W}_1(z) & \text{if } z \in [0, \underline{z}), \\ \overline{W}_1(z) - \chi(z)\frac{h-l}{hl} cr\mu \frac{\int_{\underline{z}}^{z} \frac{e^{-\int_{\hat{z}}^{t}\psi(s)ds}}{\chi^2(t)}dt}{\int_{\underline{z}}^{\infty}\frac{\lambda_h+\lambda_l}{g(t)} e^{2rt - \int_{\hat{z}}^{t}\psi(s)ds}dt} & \text{if } z \in [\underline{z}, \hat{z}), \\ \overline{W}_1(z) + \chi(z)\frac{h-l}{hl} c\left(1 + r\mu \frac{\int_{z}^{\infty} \frac{e^{-\int_{\hat{z}}^{t}\psi(s)ds}}{\chi^2(t)}dt}{\int_{\underline{z}}^{\infty}\frac{\lambda_h+\lambda_l}{g(t)} e^{2rt - \int_{\hat{z}}^{t}\psi(s)ds}dt}\right) & \text{if } z \geqslant \hat{z}, \end{cases}$$

where

$$\overline{W}_1(z) := (1 - e^{-rz})(1 - c/h)\mu.$$

Proof of Lemma 8. The proof is divided into two steps. First we show that the difference in payoffs between $W(z)$ and the complete-information payoff computed at the same level of utility f^z converges to 0 at a rate linear in r, for all z. Second, we show that the distance between the closest point on the graph of $\{f^z\}_{z\geqslant 0}$ and the complete-information payoff maximizing utility pair (which is v^*) converges to 0 at a rate linear in r. Given that the complete-information payoff is piecewise affine in utilities, the result follows from the triangle inequality.

1. We first note that the complete-information payoff along the graph of $\{f^z\}_{z\geqslant 0}$ is at most $\max\left\{\overline{W}_1(z), \overline{W}_2(z)\right\}$, where $\overline{W}_1$ is defined in Proposition 1 and $\overline{W}_2$ is the complete-information payoff for those z such that $(f_h^z, f_l^z) \in [v_h^*, \mu_h] \times [v_l^*, \mu_l]$. Substituting f_h^z, f_l^z

[34] As $z \to \hat{z}$, the integrals entering in the definition of W diverge, although not W itself, given that $\lim_{z\to\hat{z}} \chi(z) \to 0$. As a result, $\lim_{z\to\hat{z}} W(z)$ is well-defined, and strictly below $\overline{W}_1(\hat{z})$.

and simplifying, we obtain the formula for $\overline{W}_2(z)$:

$$\begin{aligned}\overline{W}_2(z) &= q\left(v_h^*\frac{h-c}{h} + (f_h^z - v_h^*)\frac{l-c}{l}\right) + (1-q)\left(v_l^*\frac{h-c}{h} + (f_l^z - v_l^*)\frac{l-c}{l}\right)\\ &= (1-e^{-rz})(1-c/l)\mu + q(h/l-1)c.\end{aligned}$$

The complete-information payoff given promise z is at most $\max\left\{\overline{W}_1(z), \overline{W}_2(z)\right\}$, because $\overline{W}_1(z), \overline{W}_2(z)$ are two of the four affine maps whose lower envelope defines $\overline{W}$. (See Section C.1.)

Fix $y = rz$ (note that as $r \to 0$, $\underline{z} \to \infty$, so that changing variables is necessary to compare limiting values as $r \to 0$), and fix y such that $le^y > \mu$ (that is, such that $g(y/r) > 0$ and hence $y \geqslant r\underline{z}$ for small enough r). Algebra gives

$$\lim_{r\to 0} \psi(y/r) = \frac{(e^y - 1)\lambda_h l - \lambda_l h}{le^y - \mu},$$

and similarly

$$\lim_{r\to 0} \chi(y/r) = (qh - (e^y - 1)(1-q)l)e^{-y},$$

as well as

$$\lim_{r\to 0} g(y/r) = le^y - \mu.$$

Hence, fixing y and letting $r \to 0$ (so that $z \to \infty$), it follows that $\frac{\chi(z)\int_{\underline{z}}^{z} \frac{e^{-\int_{\hat{z}}^{t}\psi(s)\mathrm{d}s}}{\chi^2(t)}\mathrm{d}t}{\int_{\underline{z}}^{\infty} \frac{\lambda_h+\lambda_l}{g(t)} e^{2rt - \int_{\hat{z}}^{t}\psi(s)\mathrm{d}s}\mathrm{d}t}$ converges to a well-defined limit. (Note that the value of $\hat{z}$ is irrelevant to this quantity, and we might as well use $r\hat{z} = \ln(\mu/((1-q)l))$, a quantity independent of r.) Denote this limit η. Hence, for $y < r\hat{z}$, because

$$\lim_{r\to 0} \frac{\overline{W}_1(y/r) - W(y/r)}{r} = \frac{h-l}{hl}c\eta,$$

it follows that $W(y/r) = \overline{W}_1(y/r) + \mathcal{O}(r)$. On $y > r\hat{z}$, it is immediate to check from the formula of Proposition 1 that

$$W(z) = \overline{W}_2(z) + \chi(z)\frac{h-l}{hl}cr\mu\frac{\int_{\underline{z}}^{z} \frac{e^{-\int_{\hat{z}}^{t}\psi(s)\mathrm{d}s}}{\chi^2(t)}\mathrm{d}t}{\int_{\underline{z}}^{\infty} \frac{\lambda_h+\lambda_l}{g(t)} e^{2rt - \int_{\hat{z}}^{t}\psi(s)\mathrm{d}s}\mathrm{d}t}.$$

By definition of $\hat{z}$, $\chi(z)$ is now negative. By the same steps it follows that $W(y/r) =$

$\overline{W}_2(y/r) + \mathcal{O}(r)$ on $y > r\hat{z}$. Because $W = \overline{W}_1$ for $z < \underline{z}$, this concludes the first step.

2. For the second step, note that the utility pair maximizing complete-information payoff is given by $v^* = \left(\frac{r+\lambda_l}{r+\lambda_l+\lambda_h}h, \frac{\lambda_l}{r+\lambda_l+\lambda_h}h\right)$. We evaluate $f^z - v^*$ at a particular choice of z, namely

$$z^* = \frac{1}{r} \ln \frac{\mu}{(1-q)l}.$$

It is immediate to check that

$$\frac{f_l^{z^*} - v_l^*}{qr} = -\frac{f_h^{z^*} - v_h^*}{(1-q)r} = \frac{l + (h-l)\left(\frac{(1-q)l}{\mu}\right)^{\frac{r+\lambda_l+\lambda_h}{r}}}{r + \lambda_l + \lambda_h} \to \frac{l}{\lambda_l + \lambda_h},$$

and so $\|f^{z^*} - v^*\| = \mathcal{O}(r)$. It is also easily verified that this gives an upper bound on the order of the distance between the polygonal chain and the point v^*. This concludes the second step. ■