# Dynamic Allocation without Money*

Yingni Guo† Johannes Hörner‡

July 5, 2018

**Abstract**

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

**Keywords:** mechanism design, principal-agent, quota mechanism, token budget.

**JEL numbers:** C73, D82.

## 1 Introduction

This paper studies the dynamic allocation of resources when monetary transfers are not allowed, and the optimal use of the resources is the private information of a strategic agent.

We seek the social choice mechanism that maximizes efficiency. The societal decision not to allow monetary transfers in certain contexts – whether for economic, physical, legal, or ethical reasons – is taken as given. The good to be allocated is perishable.[1] Absent private

*Financial support from NSF Grant SES-1530608 is gratefully acknowledged. This paper was previously circulated under the title "Dynamic Mechanisms without Money."

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

‡Department of Economics, Yale University, 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 their time; which patients should receive scarce medical resources (*e.g.*, blood or treatments); how governments should distribute vouchers (*e.g.*, regarding food, health services, or schooling); or which investments and activities should be approved by a firm.

information, the allocation problem is trivial: the good should be provided if and only if its value exceeds its cost. However, in the presence of private information, and in the absence of transfers, linking future allocation decisions to current decisions is the only instrument to elicit truthful information. Our goal is to understand this link and the form of inefficiencies that will arise.

In each round, the agent privately learns his value for a good, which is either high or low. The agent always enjoys consumption, and enjoys it more when his value is high. However, supplying the good entails a fixed cost, so it is efficient only when the value is high. There is no statistical evidence concerning the agent's value, even *ex post*.

The agent's value evolves over time and exhibits persistence – the higher today's value is, the more likely tomorrow's value will be high. Persistence raises substantial technical challenges, but it is realistic in most applications. As we shall see, it has an impact on the structure and the performance of the optimal policy.

To characterize the optimal allocation policy, we cast this problem as an agency model. A principal with commitment power chooses when to supply the good, given the agent's reports. If transfers were possible, charging the agent the supply cost would induce him to reveal his value truthfully. Without transfers, inefficiencies are inevitable.

The optimal mechanism can be described as a *quantified entitlement*, corresponding to the number of units that the agent can claim *consecutively* with no questions asked. In every round, the agent's entitlement first drops by one unit, and he is given the option to consume today's unit. If the agent does so, the game moves on to the next round. If the agent forgoes today's unit, his entitlement is augmented by a certain amount. The increment is chosen such that the low-value agent would be indifferent between consuming today's unit and consuming those incremental units, *if he were to cash in his entitlement as quickly as possible, with those incremental units last.*

This mechanism results in straightforward dynamics. In the initial phase, the agent claims the unit if his value is high, and forgoes it if it is low. In the long run, his entitlement 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 initial phase of efficient allocations leads to one of these two inefficient outcomes. From the agent's perspective, polarization arises eventually, since he receives either his worst or his best outcome forever. After some histories the good is never provided again. In this sense, the scarcity of good provision is endogenously determined to elicit information.

The intuition for this structure is simple, 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, since the agent's future expected values cannot be higher than today's. Whenever the value is low, the interests of the two parties are the least aligned. Granting entitlement to future units is a way to deter the low-value agent from claiming to be high and consuming today. A key feature of the mechanism is that the principal adds these incremental units at the tail of the agent's current entitlement. Because the low-value agent anticipates his future values to go up on average, the later these incremental units accrue, the more aligned the two parties' interests become.

As mentioned above, the optimal allocation is implemented by a quantified entitlement. The updating of the entitlement when the agent forgoes the unit depends on his current entitlement level. Because the low-value agent expects his future values to go up, each incremental unit (attached to the tail of his entitlement) is especially valuable if his entitlement is already large, since his value is more likely to be high by the time he claims this incremental unit. Hence, when we control for impatience, an agent with a small entitlement must be granted more incremental units than an agent with a large one.

What is the impact of persistence? As values become more persistent, a low value today 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 forgo the current unit. As a result, the agent's entitlement is absorbed more quickly into the inefficient outcomes (*i.e.*, zero or infinity). Hence, persistence jeopardizes the principal's ability to use future allocations to induce truth-telling. Efficiency decreases with persistence.

Our findings contrast with those in studies of static design with multiple units (*e.g.*, Jackson and Sonnenschein, 2007), since the optimal mechanism is not a (discounted) quota mechanism as commonly studied in the literature: 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. By backloading inefficiencies, our mechanism exploits the agent's ignorance regarding his future values. In section 2, we use a simple, two-round example to illustrate that the benefit of using the optimal mechanism over the quota mechanism could be quite stark when the agent's value is highly persistent. Examining the continuous-time limit of our model, we show that our mechanism results in a higher rate of convergence to efficiency than does the quota mechanism.[2]

Backloading inefficiency contrasts with the outcome in dynamic mechanisms with trans-

[2]See Lemma 8 in online appendix E.

fers (*e.g.*, Battaglini, 2005; Pavan, Segal, and Toikka, 2014). 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 granted the unit either forever or never again. Hence, polarization is inevitable, 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).

Our insights extend to settings in which payments are one-sided, from the principal to the agent. We discuss this extension in section 5.5.

**Applications:** Allocation problems without transfers are plentiful. Our results inform the best practices on dynamic resource allocation policies. 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 regardless of how urgent their need is. This preference misalignment 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 institutions must confront (*e.g.*, Sendelbach, 2012). We suggest the best approach for trading off the risks of neglecting a patient in need and responding to one who simply "cries wolf."

Our results also shed light on the design of antipoverty programs (*e.g.*, Bardhan and Mookherjee, 2005). The poor have little liquidity left once subsistence requirements are met. The government can supply a good that the poor might value highly, but it must rely on information obtained from the poor themselves. If the program were one-shot, the supply would fail to respond to the local needs. However, if the program is long-lasting, our results imply that the optimal allocation policy is a quantified entitlement. Our mechanism improves the targeting of service delivery. Similarly, the results apply to other welfare programs for which the transfers from the citizens to the government are limited (*e.g.*, Alderman, 2002).

As a third application, consider the problem of dynamic allocation schemes in (cloud) computing. Due to lack of resources, it is not possible for cloud providers to satisfy all computing requests. Hence, open-source resource lease managers (OpenNebula, Haizea,...) must manage scheduling policies to maximize efficiency (*e.g.*, Saraswathi, Kalaashri and Padmavathi, 2015). Virtual currencies are routinely use in the design of online mechanisms for such distributed systems (*e.g.*, Ng, 2011). While our environment is certainly stylized, we provide an exactly optimal policy, as opposed to the heuristic procedures used in the literature.

Our results readily extend to the case in which there is moral hazard, in addition to adverse selection. For instance, the agent is a borrower, who privately observes his credit needs, and the principal is a lender, who can grant the desired short-term loans. If granted a loan, the borrower can default or honor it – an observable choice. Such a standard model of lending (*e.g.*, Bulow and Rogoff, 1989) is readily accommodated. Our results apply when the main instrument to discipline the borrower is through future lending decisions. The only adjustment needed is that the principal freezes lending when the agent's entitlement to future loans is very small, and lets the agent's entitlement recharge. Once the entitlement is large enough, the principal is willing to grant loans again, as she no longer expects the agent to default. We discuss this extension in section 6.

**Related Literature:** Our paper is closely related to Li, Matouschek, and Powell (2017) and Lipnowski and Ramos (2018). Li, Matouschek, and Powell (2017) analyze the power dynamics within organizations, and have a benchmark model similar to our i.i.d. case. They show that the prospect of future power motivates the agent to make good use of his power today, and that the power evolution ends with either restricted or entrenched power. Lipnowski and Ramos (2018) analyze a repeated game in which the principal decides whether to delegate projects to the agent. Their principal cannot commit to future allocations. They show that the principal gives away projects whenever the agent's continuation utility is sufficiently high, and that the agent is eventually terminated. We characterize the optimal mechanism without transfers when the agent's value is persistent. Persistence implies an exogenous link between periods. To the best of our knowledge, we are the first to fully characterize the optimal mechanism without transfers when the agent has persistent information and the two parties interact repeatedly.

Our work is closely related to research on dynamic mechanisms with transfers when the agent has persistent private information (*e.g.*, Battaglini, 2005; Pavan, Segal, and Toikka, 2014; Battaglini and Lamba, 2017). Although dynamic mechanisms without transfers and those with transfers address different applications, comparing the results delineates both the role of transfers and the different forces that drive the optimal design. (Section 5.5 discusses the role of transfers.) The obvious benchmark is Battaglini (2005), who considers the exact same model as ours but allows for transfers. His results are diametrically opposed to ours. In Battaglini (2005), efficiency improves over time. In our setting, efficiency decreases over time, with an asymptotic outcome that is the outcome of the one-shot game. Krishna, Lopomo, and Taylor (2013) analyze a dynamic screening problem in which the agent has independent types and is protected by limited liability (although transfers are allowed).[3] They characterize how transfers and limited liability affect the optimal contract and the dynamics.

Our paper is also related to the literature on linking incentives. This refers to the notion that as the number of identical decision problems increases, linking them improves on each isolated problem. Radner (1981) and Rubinstein and Yaari (1983) show that linking incentives in moral hazard settings achieves efficiency asymptotically. Jackson and Sonnenschein (2007) show that the quota mechanism is asymptotically optimal as the number of linked problems approaches infinity.[4] In contrast to the bulk of this literature, we focus on the exactly optimal mechanism for a fixed discount factor. This allows us to (i) identify the form of inefficiencies and results (backloading of inefficiency, for instance) that need not hold for other asymptotically optimal mechanisms, and (ii) 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 (2016) characterizes the optimal mechanism in a related environment. It is assumed that the agent privately learns his cost of actions before time zero, and that his type in each round is payoff-relevant only for the principal but not for himself. He shows that a discounted quota mechanism is optimal. In our setting, the agent's type affects the payoffs of both parties. Our mechanism differs from a discounted quota mechanism.

Relatedly, the notion that token budgets can be used to incentivize agents with private information and observable actions has appeared in several papers. Within the context of games, Möbius (2001) shows that using the difference in the numbers of favors granted between two agents as a yardstick for granting new favors sustains cooperation in long-run relationships.[5] While his token rule is suboptimal, it has desirable properties: properly calibrated, it yields an efficient allocation as discounting vanishes. Hauser and Hopenhayn (2008) measure the optimality of a simple token rule (according to which each favor is

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

[4]Cohn (2010) links multiple bargaining problems, and constructs a mechanism that converges to efficiency faster than the simple quota mechanism. Eilat and Pauzner (2010) characterize the second-best mechanism when linking multiple bargaining problems. Escobar and Toikka (2013) analyze linking mechanisms when the agents' types follow a Markovian chain, and show that efficiency can be approximated as discounting vanishes.

[5]See also Abdulkadiroğlu and Bagwell (2012). Olszewski and Safronov (2018a, 2018b) show that the efficient outcome can be approximated by chip-strategy equilibria as discounting vanishes. In contrast, we characterize the optimal mechanism for a fixed discount factor.

weighted equally, independent of the history), showing that this rule might be too simple (the efficiency cost reaching up to 30% of surplus). Our model can be viewed as a game with one-sided incomplete information in which the production cost is known. There are several important differences, however. First, our principal has commitment and hence is not tempted to act opportunistically. Second, our agent's information is persistent, so he has private information not only about today's value but also about future values.[6]

Some principal-agent models find that simple capital-budgeting rules are exactly optimal in related models (*e.g.*, Malenko, 2018). Our results suggest how to extend such rules when values are persistent. Indeed, in the i.i.d. case, our mechanism admits an implementation in terms of a two-part tariff.[7] Only through the general persistent environment do we learn that it is more natural to interpret the budget as an entitlement for *consecutive* units that the agent can claim, rather than a budget with a fixed "currency" value.

More generally, it has long been recognized that allocation rights to other 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. Miralles (2012) solves a two-unit version, with both values being privately known at the outset. Boleslavsky and Kelly (2014) solve for the optimal regulation in a two-period environment in which the firm privately learns its current compliance cost.

We formulate the optimal mechanism design in a repeated hidden-information environment as a dynamic programming problem, following Green (1987) and Thomas and Worrall (1990). Thomas and Worrall (1990) characterize the optimal insurance contract, showing that the agent's future utility becomes arbitrarily negative with probability one. In our environment, polarization rather than immiseration results in the long run. The dynamic programming approach is also used to study repeated moral hazard problems (*e.g.*, Spear and Srivastava (1987) and Clementi and Hopenhayn (2006)). Unlike the bulk of the literature, we assume that the agent's values are serially correlated and explore how persistence affects the performance of the optimal mechanism. Building on the recursive method by Fernandes and Phelan (2000), we fully characterize the set of interim utilities and the optimal mechanism, and show that higher persistence hurts the principal.[8]

[6] This results in different limiting models in continuous time. Our model corresponds to the Markovian case in which flow values *switch* according to a Poisson process. In Hauser and Hopenhayn (2008), the lump-sum value *arrives* according to a Poisson process.

[7] If the agent wants to consume the unit, he pays a fixed utility budget.

[8] Fu and Krishna (2017) study the optimal financial contract when the agent's type is persistent. Since transfers are allowed, they show that all utility vectors above the 45 degree line are feasible interim utilities. In contrast, our principal can only use future allocations to incentivize truth-telling, leading to a completely

In section 2, we illustrate the essence of our problem and solution with a two-round example. Section 3 introduces the model. Section 4 examines the i.i.d. case, introducing many of the ideas of the paper, while section 5 examines the general model and develops an implementation for the optimal mechanism. Section 6 discusses extensions.

# 2 An Example

A principal (she) and an agent (he) interact for two rounds: today and tomorrow. There is no discounting, so each party maximizes the sum of today's and tomorrow's payoffs. In each round, the agent privately observes his value for a good, and the principal can supply the good at a cost $c = \frac{9}{4}$. The agent's value is either high, $h = 3$, or low, $l = 1$, and it changes over time. Today's value is equally likely to be high or low. Tomorrow's value stays the same as today's value with probability $\kappa \in \left[\frac{1}{2}, 1\right)$. A higher value of $\kappa$ corresponds to stronger persistence.

When the good is supplied, the principal's payoff is the agent's value minus the cost. Hence, the principal wants to supply the good only when the agent has a high value. However, the principal must rely on the agent's reports. Given that the agent is strategic, the principal knows that truthful reporting does not come for free. If the interaction was one-shot, the principal would not supply the good, as the expected value falls short of the cost. Similarly, if she knew that today's value is low, she would prefer not to supply the good tomorrow, since the current low type's expected value tomorrow, $v_l := \kappa l + (1 - \kappa)h$, would be even lower than the expected value in the first round (and hence $v_l < c$).

Under a quota mechanism, the principal decides on how many units she will supply *in total*. The agent allocates these units whenever he likes. In this example, the optimal quota number will be either one or zero. If the agent has a quota of one unit, a current high type will use it today while a current low type will save it for tomorrow. The principal's payoff is given by:

$$\frac{1}{2}(h - c) - \frac{1}{2}(c - v_l) = \frac{3}{4} - \kappa.$$

The expected value $v_l$ decreases in $\kappa$, because a current low type is unlikely to switch to high when persistence is strong. The principal's payoff from offering a quota of one unit is positive if $\kappa < \frac{3}{4}$. Otherwise, the principal prefers to never supply the good.

Our mechanism offers a number of units that the agent can claim *consecutively*. Whenever the agent forgoes a unit, this number is adjusted so that the current low type is indifferent

different characterization of the set of interim utilities and the optimal policy.

between forgoing the unit and not doing so. In this example, the principal either offers no unit at all, or offers to supply today's unit. If the principal offers today's unit, but the agent forgoes it, the principal will supply tomorrow's unit with probability $\frac{l}{v_l}$. Since the low type's expected value tomorrow is exactly $v_l$, he is indifferent between consuming today, and consuming tomorrow with probability $\frac{l}{v_l}$. The principal's payoff from providing one unit is:

$$\frac{1}{2}(h-c) - \frac{1}{2}(c-v_l)\frac{l}{v_l} = \frac{6-7\kappa}{12-8\kappa}.$$

This is always higher than the payoff from offering a quota of one unit, due to the multiplier $\frac{l}{v_l} < 1$. This payoff is positive if and only if $\kappa < \frac{6}{7}$.

The quota mechanism manages to align the two parties' incentives to some extent, but is much less efficient than the optimal mechanism. The performance difference is quite stark when the value is highly persistent. When $\kappa \in \left[\frac{1}{2}, \frac{3}{4}\right)$, both mechanisms give the principal a positive payoff. However, the principal's payoff difference across the two mechanisms increases as the value is more persistent (*i.e.*, as $\kappa$ increases). When $\kappa \in \left[\frac{3}{4}, \frac{6}{7}\right]$, our mechanism still gives the principal a positive payoff, while the quota mechanism never supplies the good.

This example shows how the optimal mechanism differs from the quota mechanism, and how it cranks up the reward for forgoing a unit as persistence increases. However, it fails to illustrate how this reward is distributed intertemporally – as explained, with persistence, it is optimal to append the awarded claims at the tail of the agent's current entitlement. Also, the example cannot convey what happens in the long run. For these reasons, we now turn to the infinite-horizon problem.

# 3 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 can supply a unit of a good at cost $c$. 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 always enjoys getting the unit. The value follows a Markov chain. For all $n$, the value switches from one round to the next according to the following:

$$\mathbf{P}[v_{n+1} = h \mid v_n = h] = 1 - \rho_h, \quad \mathbf{P}[v_{n+1} = l \mid v_n = l] = 1 - \rho_l, \quad \text{for some } \rho_h, \rho_l \in [0,1].$$

We assume that the initial value is drawn according to the invariant distribution $\mathbf{P}[v_0 = h] = q := \rho_l/(\rho_h + \rho_l)$. The expected value of the good is $\mu := \mathbf{E}[v] = qh + (1-q)l$.

We focus on nonnegative correlation (*i.e.*, $1 - \rho_h \geqslant \rho_l$). The value is more likely to be $h$, conditional on the previous value being $h$ rather than $l$.[9] Two cases of special interest occur when $1 - \rho_h = \rho_l$ and $\rho_h = \rho_l = 0$, corresponding to i.i.d. values and perfectly persistent values.[10]

At the beginning of each round, only the agent is informed of the value. The two parties share a common discount factor $\delta < 1$, where $\delta > l/\mu$ and $\delta > 1/2$.[11] The supply decision in round $n$ is denoted by $x_n \in \{0,1\}$; $x_n = 1$ means that the good is supplied. The principal internalizes both the cost of supplying the good and its value to the agent. Thus, given an infinite history $\{v_n, x_n\}_{n=0}^{\infty}$, 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.

Given the principal's commitment, we may focus on mechanisms in which the agent truthfully reports his type at every round and the principal commits to a possibly random supply decision as a function of the entire history of reports. This is without loss of generality. (Because preferences are time-separable, the policy may be taken as independent of past supply decisions.)

Formally, a direct mechanism, or *policy*, is a collection $(\mathbf{x}_n)_{n=0}^{\infty}$, $\mathbf{x}_n : \{l,h\}^{n+1} \to [0,1]$, mapping the agent's reports up to round $n$ onto a probability of supplying the unit in round $n$. A direct mechanism induces a Markov decision problem for the agent. A reporting strategy is a collection $(\mathbf{r}_n)_{n=0}^{\infty}$, where $\mathbf{r}_n : \{l,h\}^n \times \{l,h\} \to \Delta(\{l,h\})$ maps previous reports and the value in round $n$ onto a report for that round.[12] The policy is *incentive-compatible*

[9] We comment on the case in which values are negatively correlated in footnote 31.

[10] There are well-known examples in the literature in which persistence changes results more drastically than it does here; see, for instance, Halac and Yared (2015).

[11] It is important that the discount factor be common. Here, this is a natural assumption, because we view our principal as a social planner trading off the agent's utility with the social cost of providing the good.

[12] Given past reports and the current value, the decision problem from round $n$ onward does not depend

if reporting the current value truthfully is always optimal.

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

We now comment on three assumptions maintained throughout.

- There is no *ex post* signal regarding the realized value of the agent. In particular, the principal does not observe the agent's realized utility. Plainly, this does not fit all applications. In some cases, a signal about the true value obtains whether the good is supplied or not. In other cases, a signal occurs only if the good is supplied. It is also possible that this signal occurs only from *not* supplying the good, if supplying it averts a risk. 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 token budgets, which typically assumes no commitment on either side and solves for the equilibria of the game. In section 6, 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).

# 4 Independent Types

We begin with the benchmark case in which types are i.i.d. Section 5 is devoted to the general case.

With independent types, it is well known that attention can be restricted to policies that are represented by a tuple of functions $p_h, p_l : [0, \mu] \to [0, 1]$ and $U_h, U_l : [0, \mu] \to [0, \mu]$, mapping the agent's promised utility $U$ onto (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.[13] These functions must be consistent in the sense

on past values. Hence, without loss of generality, this strategy does not depend on past values.

[13]See Spear and Srivastava (1987) and Thomas and Worrall (1990). Not every policy can be summarized this way, as the principal need not treat two histories leading to the same continuation utility identically. However, the principal's payoff must be maximized by some policy that does so.

that, given $U$, the probabilities of supplying the good and the continuation utilities yield $U$ to the agent. 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:

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

subject to the incentive constraints for truth-telling, and the promise-keeping constraint:

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

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

$$U = (1-\delta)\left(q p_h h + (1-q) p_l l\right) + \delta \left(q U_h + (1-q) U_l\right). \qquad (PK)$$

## 4.1 Complete Information

We first review the benchmark in which the agent's value is public; 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 history of past values. The only constraint is that the agent's expected utility at the start has to be some fixed $U$. Since there is no private information, it is without loss to pick $p_h, p_l$ as 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 $p_l = 0$. For $U > qh$, the principal must supply the good with positive probability even when the value is low.[14] The lemma below follows readily.

[14]Given $U \neq qh$, there are many other nonconstant optimal policies. For instance, if $U < qh$, then efficiency requires only that the good not be 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 by* $\overline{W}$, *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.

## 4.2 The Optimal Mechanism

For the optimal policy under incomplete information, we begin with an informal discussion that follows from two observations. 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 still 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 zero, 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 $\delta$ 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 choice $(p_h, p_l) = (1, 0)$. 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 differ from those in models with transfers. Here, inefficiencies are backloaded.

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

*the low type's incentive constraint always binds.*

This is because the principal's payoff is concave in $U$. (If the principal's payoff failed to be concave, she could offer the risk-neutral agent a fair lottery that would make her better off.) Concavity implies that there is no benefit from spreading continuation utilities 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 immediately obtain the continuation utilities. 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, which means that:

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

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 $(1-\delta)q(h-l)$ 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 drop in the continuation utility, albeit a smaller one 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*

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

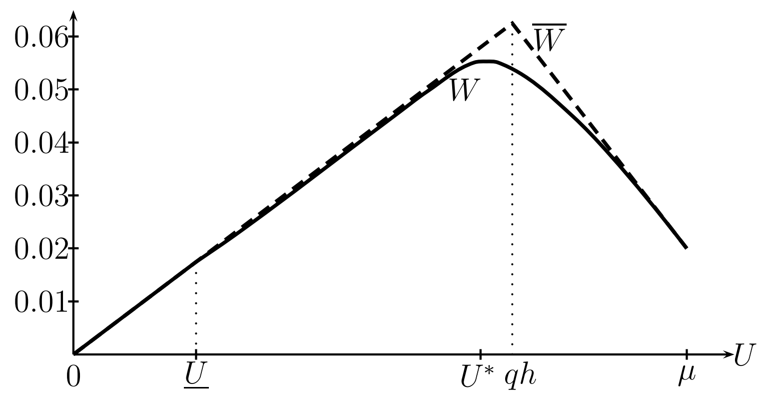


Figure 1: Value function for $(\delta, l, h, c, q) = (19/20, 2/5, 3/5, 1/2, 3/5)$.

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

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

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

We now turn to the shape of the value function and the utility dynamics. For any $U \leqslant \underline{U}$, the continuation utility decreases after both reports and eventually drops to zero. The principal is able to deliver all the promised utility without *ever* supplying the unit when the value is low. Hence, there is no efficiency loss (*i.e.*, $W(U) = \overline{W}(U)$). However, for $U \in (\underline{U}, \mu)$, the agent's utility reaches 0 and $\mu$ with positive probability. Both inefficiencies eventually occur – the good is denied when the value is high and provided when the value is low. Therefore, the value function $W$ is strictly below $\overline{W}$ for $U \in (\underline{U}, \mu)$. The solid curve in figure 1 illustrates the function $W$.

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

**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 $W$ obtained from Theorem 1 (ignoring again the peripheral intervals for the sake of the discussion):

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

Since $W$ is differentiable, we have $W'(U) = qW'(U_h) + (1-q)W'(U_l)$, or:

$$W'(U_n) = \mathbf{E}[W'(U_{n+1})].$$

This bounded martingale is first mentioned in Thomas and Worrall (1990), and is useful for understanding the process of utilities $\{U_n\}_{n=0}^{\infty}$, which must converge to 0 or $\mu$.[15] The optimal initial promise $U^*$ 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. Therefore,

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

**Lemma 3.** *The process $\{U_n\}_{n=0}^{\infty}$ (with $U_0 = U^*$) converges to 0 or $\mu$, 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 both the discount factor and the invariant probability (although the step size of the random walk $\{U_n\}_{n=0}^{\infty}$ depends on $\delta$ and $q$). 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 zero more often as $c$ increases.

[15] Since $W$ is concave, so $W'(U)$ is bounded by its value at 0 and $\mu$, given in Lemma 2. Convergence of $W'(U)$ implies convergence of the utility. 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 $U_\infty$ must be 0. Therefore, $U_n$ converges to either 0 or $\mu$.

# 5 Persistent Types

As a preliminary, consider the case of perfect persistence. The problem is equivalent to a static one for which the principal either always provides the good (if $\mu \geqslant c$) or never does so. In this case, future allocations cannot be used as instruments to elicit truth-telling. This makes the role of persistence not entirely clear. Because current types assign different probabilities of being a high type tomorrow, one might hope that tying promised future utility to current reports would facilitate truth-telling. However, the case of perfect persistence shows that correlation diminishes the scope for using future allocations as "transfers."

With persistent types, *ex ante* utility is no longer a sufficient state variable. To understand why, note that with independent types, the agent of a given type can evaluate his utility based on his current type, the probability of getting the current unit, and the promised continuation utility. However, if the agent has private information about his future types, he cannot evaluate his continuation utility without knowing how it is implemented.

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 is irrelevant for his decision problem.) The one-shot deviation principle holds: when the agent considers lying now, there is no loss of generality in assuming that he will report truthfully tomorrow.

We are not the first to note that, with persistence, interim utilities must be included as state variables.[16] Yet, to the best of our knowledge, we are the first to fully characterize the set of interim utilities for a dynamic allocation problem with persistent types. The characterization in section 5.2 is quite intuitive and potentially useful for other dynamic allocation problems with persistent types.

For the principal's problem, we must also specify her belief that the agent's type is high. This belief takes only three values: the invariant probability $q$ for the initial round, $1-\rho_h$ if

[16] See Fernandes and Phelan (2000), Doepke and Townsend (2006), and Zhang and Zenios (2008).

the previous report is high, and $\rho_l$ if low. Yet it is just as convenient to treat this belief as an arbitrary element in the unit interval.

## 5.1 The Program

As discussed above, the principal's 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. Let $V$ denote the set of interim utility pairs supported by some incentive-compatible policy. (Section 5.2 characterizes $V$.) Let $\mu_h$ and $\mu_l$ be the largest utilities that can be promised when the current type is, respectively, high and low, delivered by always supplying the good.[17] Hence, $V$ is a subset of $[0, \mu_h] \times [0, \mu_l]$.

A policy is now (i) a pair of functions $p_h, p_l : V \to [0, 1]$, mapping the current utility pair $U$ onto the supply probability given the report, and (ii) a pair of functions $U_h, U_l : V \to V$, mapping $U$ onto the continuation utility pair, which is $U_h = (U_h(h), U_h(l))$ if the report is $h$ and $U_l = (U_l(h), U_l(l))$ if it is $l$. Here, $U_v(v')$ is the continuation utility if the current report is $v$ and tomorrow's report is $v'$. For any $U$ and $\phi$, the optimality equation states that:

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

subject to the following promise-keeping and incentive 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 $V$.[18]

[17] The utility pair $(\mu_h, \mu_l)$ solves

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

[18] The optimal policy $(p_h, p_l, U_h, U_l)$ will be independent of $\phi$, because we can decompose the optimization

## 5.2 Incentive-Compatible Utility Pairs

When the agent has private information about his future values, it is not obvious which vectors of interim utilities are generated by incentive-compatible policies. For instance, promising to supply all future units in the event that the agent's current report is high while supplying none if it is low is not incentive-compatible. Our first step is to solve for the incentive-compatible utility pairs. These are interim utilities supported by some incentive-compatible policy.[19]

**Definition 1.** *The incentive-compatible set, $V \subset \mathbf{R}^2$, is the set of interim utilities in round* 0 *that are obtained by some incentive-compatible policy.*

To get some intuition regarding the structure of $V$, we enumerate some of its elements. Clearly, both $(0,0)$ and $(\mu_h, \mu_l)$ are in $V$; they can be generated by either never or always supplying the good, respectively. More generally, for any integer $m \geqslant 0$, the principal can supply the good for the first $m$ rounds, independent of the reports, and never supply the good 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 good for the first $m$ rounds but does supply it afterward, independent of the reports. A (mixed) frontloading policy randomizes over two pure frontloading policies with consecutive integers. A backloading policy is defined analogously.

The utility pairs corresponding to frontloading and backloading policies can be immediately defined in parametric forms. Given $m \in \mathbf{N}$, let

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

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

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

problem into two subproblems: (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 $IC_l$; and (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 $IC_h$.

[19] The incentive-compatible set differs from the feasible set in Guo (2016). Here, the agent has private information in every round, so an incentive-compatible policy may respond to the agent's report in every round. In Guo (2016), the agent has private information only before round zero and a policy's resource pair is independent of the agent's information. This conceptual difference leads to different characterizations and proof methods.

and set $b^m := (b_h^m, b_l^m)$.[20] This is the pair when the principal supplies the unit only from round $m$ onward. The sequence $\{f^m\}_{m\geqslant 0}$ is increasing as $m$ increases, with $f^0 = (0,0)$ and $\lim_{m\to\infty} f^m = (\mu_h, \mu_l)$. In contrast, $\{b^m\}_{m\geqslant 0}$ is decreasing, with $b^0 = (\mu_h, \mu_l)$ and $\lim_{m\to\infty} b^m = (0,0)$.

Now consider a frontloading policy and a backloading one such that the high type is indifferent between the two; the low type must then prefer the backloading one. The low type's expected value in a given round $m$ increases with $m$, and backloading allows him to consume later rather than earlier. For a fixed high type's utility, not only does the low type prefer backloading to frontloading, but these policies also give the low type the highest and lowest utilities. In other words, these policies span the incentive-compatible set.

**Lemma 4.** *It holds that*

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

As shown in figure 2, $V$ is a polygon with countably infinite vertices. Importantly, front- and backloading policies are not the only policies that achieve boundary utilities of $V$. The frontloading boundary includes policies that assign the maximum probability of supplying the good for high reports, while promising continuation utilities lying on the frontloading boundary for which $IC_l$ binds.[21] Among the policies of this group are frontloading policies. The backloading boundary includes policies assigning the minimum probability of supplying the good for low reports, while promising continuation utilities on the backloading boundary for which $IC_h$ binds. Among the policies of this group are backloading policies.

We now discuss several properties of $V$. Readers who are more interested in the optimal mechanism and its implementation may skip this part and proceed to section 5.3. We can readily verify that:

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

When the switching time $m$ 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 one, and approach one as $m$ goes to infinity.

If the principal could see the agent's value, then she would supply the good if and only if that value is high. The agent's interim utility pair is then denoted by $v^* = (v_h^*, v_l^*)$, called

[20] Here and in section B.1, we omit the corresponding analytic expressions.

[21] Under such policies, a low type is always willing to report high, and the good is supplied after high reports as long as the promised utility permits. Hence, the agent's utility pair is the same as that given by some frontloading policy.

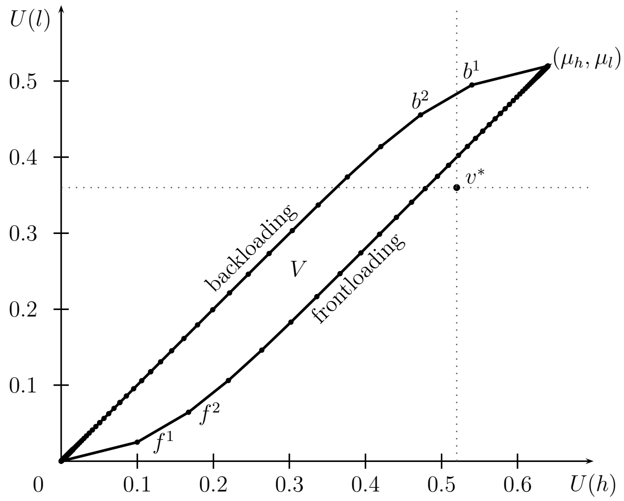


Figure 2: Incentive-compatible set $V$ for $(\delta, \rho_h, \rho_l, l, h) = (9/10, 1/3, 1/4, 1/4, 1)$.

the first-best utility.[22] We can easily verify that $(\mu_l - v_l^*)/(\mu_h - v_h^*) > 1$, so 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, $v_h^*$, then any incentive-compatible policy has to grant the low type strictly more than $v_l^*$. Figure 3 illustrates how persistence affects the

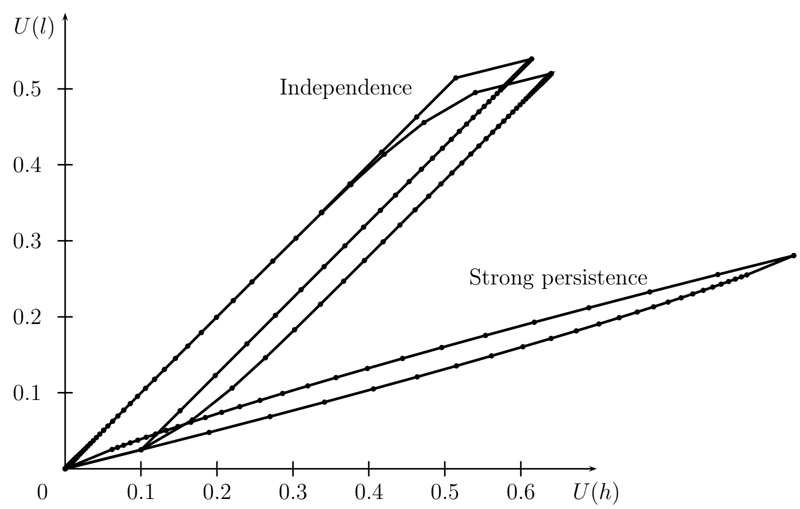


Figure 3: Impact of persistence on $V$.

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

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

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

## 5.3 The Optimal Mechanism and Implementation

Not every incentive-compatible 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.

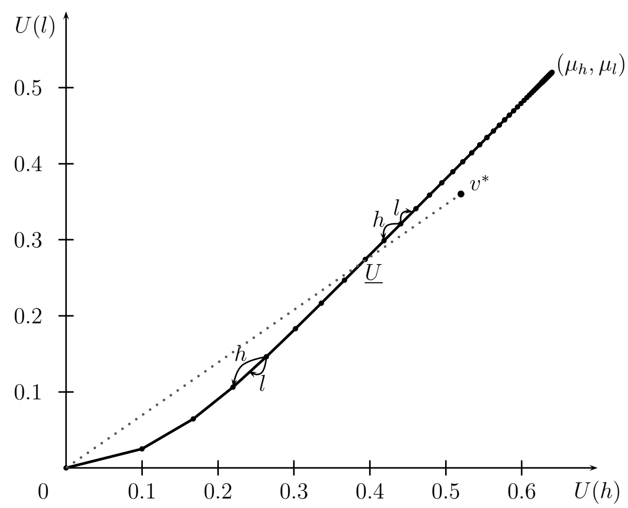


Figure 4: Dynamics of utility on the frontloading boundary.

This relevant subset is the frontloading boundary. Two observations from the i.i.d. case remain valid. First, the efficient choice $(p_h, p_l) = (1, 0)$ is made as long as possible; second, promised continuation utilities are chosen such that the low type is indifferent between reporting high and low. To understand why such a policy yields utilities on the frontloading boundary (as mentioned in the paragraph after Lemma 4), 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 had 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. In contrast to a frontloading policy which induces no information, the optimal policy implements efficiency as long as possible.

Figure 4 depicts the utility dynamics. Specifically, $U_h$ is chosen from the frontloading boundary to deliver the interim utility $U(h)$ to the high type:

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

Here, $\mathbf{E}[U_h \mid v]$ is the expectation of the utility vector $U_h$, provided that the current type is $v$ (*e.g.*, for $v = h$, $\mathbf{E}[U_h \mid v] = \rho_h U_h(l) + (1-\rho_h)U_h(h)$). Since the low type is indifferent between reporting high and low, $U_h$ also satisfies the following:

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

The promised $U_l$ is chosen from the frontloading boundary to deliver $U(l)$ to the low type:

$$U(l) = \delta\mathbf{E}[U_l \mid l].$$

After a high report, $U_h$ is lower than $U$.[23] However, $U_l$ might be lower or higher than $U$. If $U$ is high enough, $U_l$ is higher than $U$; if $U$ is low enough, $U_l$ is lower than $U$ under a certain condition. This condition has a simple geometric representation: let $\underline{U}$ denote the intersection of the half-open line segment $((0,0), v^*]$ with the frontloading boundary.[24] Then, $U_l$ is lower than $U$ if and only if $U$ lies below $\underline{U}$.[25] This intersection exists if and only if

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

which is satisfied when $\rho_l$ is sufficiently high, that is, when 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

[23]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.

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

[25]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 4. We set $\underline{U} = 0$ if the intersection does not exist.

low report. When the low-type persistence is high, this does not occur.[26] As in the i.i.d. case, the principal achieves the complete-information payoff if and only if $U \leqslant \underline{U}$ or $U = (\mu_h, \mu_l)$.

We summarize this discussion in the following theorem with the adjustments for the peripheral regions where $U$ is close to $(0,0)$ or $(\mu_h, \mu_l)$.

**Theorem 2.** *The optimal policy consists of the following allocation probabilities:*

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

*in addition to (i) continuation utility pairs $U_h$ and $U_l$ on the frontloading boundary of $V$ such that $IC_l$ always binds, and (ii) a (specific) initial promise $U^* > \underline{U}$ 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 persistent case 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 utility pairs on the frontloading boundary are indexed by the number of units that the agent can claim in a row. We represent each pair by a number, $z \geqslant 0$, with the interpretation that the agent can claim the good for the first $\lfloor z \rfloor$ rounds and can claim it with probability $z - \lfloor z \rfloor$ also during round $\lfloor z \rfloor$. (Here, $\lfloor z \rfloor$ denotes the integer part of $z$.) The interim utilities for the high and low types, given the entitlement $z$, are written as $(U(h,z), U(l,z))$:

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

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

We can implement the optimal policy as follows. Suppose that, at the beginning of round $n$, the agent is entitled to $z_n$ units in a row. If the agent asks for the unit and $z_n \geqslant 1$, he gets today's unit and his entitlement tomorrow drops to $(z_n - 1)$. It is easy to verify that the following holds both when $v$ equals $h$ and when it equals $l$:

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

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

If the agent asks for the unit and $z_n < 1$, he receives the unit with probability $z_n$ and his entitlement tomorrow drops to zero.

If the agent claims to be low and thus forgoes today's unit, his entitlement tomorrow is revised to $z_{n+1,l}$ 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.[27] Combining (5) and (6), we obtain the following:

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

Thus, this $z_{n+1,l}$ is chosen such that the low type is indifferent between consuming today's unit and consuming the additional entitlement between $z_{n+1,l}$ and $(z_n - 1)$ from tomorrow's perspective.[28]

Whenever the agent forgoes today's unit when his entitlement is $z_n$, the principal compensates him by adding units, $z_{n+1,l} - (z_n - 1)$, at the tail of his entitlement tomorrow. Adding these units at the very tail saves on the units needed for deterring the low type from consuming today, because the current low value becomes increasingly irrelevant for the expected value toward the tail. Therefore, when we control for impatience, the higher the agent's entitlement or the lower the persistence, the fewer units are needed.

Given any choice of $U$, finitely many consecutive reports of $l$ or $h$ suffice for the promised utility to reach $(\mu_h, \mu_l)$ or $(0, 0)$.[29] By the Borel-Cantelli lemma, this implies that absorption almost surely occurs. When we start from the optimal initial promise $U^* > \underline{U}$, both long-run outcomes have strictly positive probability. However, unlike in the i.i.d. case, we cannot solve for the absorption probabilities in closed form.

[27] This is impossible if the promised $z_n$ is too large. In that case, today's unit is provided with the probability $\tilde{q}$ that solves $U(l, z_n) = \tilde{q}(1-\delta)l + \delta \mathbf{E}\left[U(v_{n+1}, \infty) \mid v_n = l\right]$. The implementation can be extended to this case. We omit the details.

[28] Note that $z_{n+1,l}$ is higher than $z_n - 1$, but might be lower than $z_n$. Also, if $z_n < 1$, the number $z_{n+1,l}$ is chosen such that the low type is indifferent between consuming today's unit with probability $z_n$ and consuming the entitlement $z_{n+1,l}$ from tomorrow on.

[29] There exists a finite $N$ such that for every $U$, if the current promise is $U$, either $N$ consecutive $l$ reports suffice for the promise utility to reach $(\mu_h, \mu_l)$ or $N$ consecutive $h$ reports suffice for the promise utility to reach $(0, 0)$.

## 5.4 The Role of Persistence

Due to the high dimensionality of the principal's program, it is difficult to establish comparative statics analytically. However, it is straightforward to obtain them numerically.

We start with persistence, by using the following measure of persistence. Consider 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 persistent than $\{v'_n\}_{n=0}^{\infty}$, since the latter can be interpreted as the same process as the former but sampled less frequently. The i.i.d. case obtains in the limit when $\omega = \infty$.[30]

As shown in figure 5, persistence hurts the principal. (The left-hand side shows that the principal's payoff, $W(U^*)$, increases in $\omega$.) It is not entirely obvious that persistence hurts the principal, since 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.[31] 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 $(0,0)$ or $(\mu_h, \mu_l)$. 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 ambiguous. Martingale methods do not allow a simple formula as in the i.i.d. case; 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 $\mu$ or 0, according to whether granting the unit indiscriminately exceeds the cost of doing so. Hence, persistence might either increase or decrease the likelihood of absorption at zero, as well as the expected utility of the agent, depending on the payoff parameters.

We have omitted a discussion of other comparative statics. Indeed, they align with

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

[31] 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. Since $\delta > l/h$, truth-telling is optimal. Just as in the case of a lower discount rate, a more negative correlation makes future promises into more effective incentives because preference misalignment is short-lived.

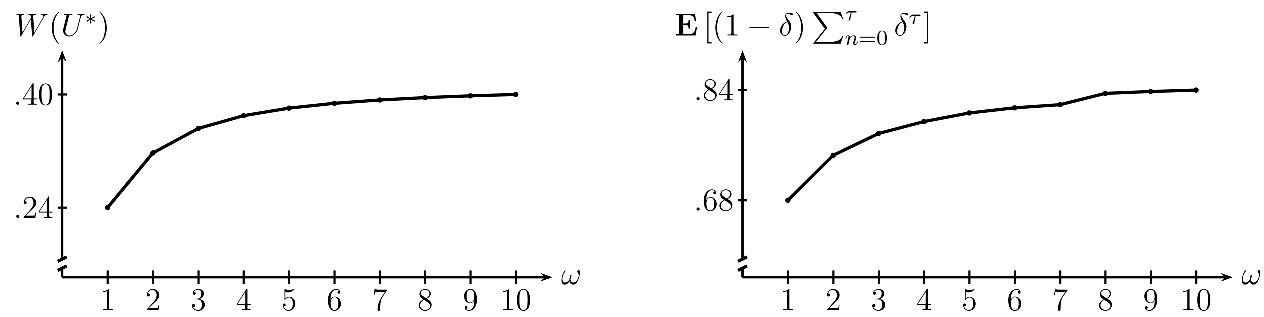


Figure 5: The value and the (discounted) time until absorption for $(\delta, l, h, c) = (.7, 1, 3, 2)$. The transition matrix is $((.9, .1), (.1, .9))^{\omega}$. The absorption time is denoted by $\tau$.

intuition: the value $W(U^*)$ increases with the discount factor $\delta$, with probability $\rho_l$, as well as with the high value $h$. However, the value is not monotone in $l$ due to the following two counteracting forces. On the one hand, as $l$ increases, it becomes more difficult to induce truth-telling by the low type. On the other, the principal has less to lose from serving a low type. The impact on the value can be in either direction. Lastly, because the cost has no impact on the agent's incentives, it is easy to show analytically that the value $W(U^*)$ decreases in $c$.[32]

## 5.5 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). We emphasize that dynamic mechanisms without transfers address different applications than do those with transfers. Yet, the comparison delineates both the role of transfers in dynamic mechanism design, and the differences between these two strands of literature.

One of the main findings of Battaglini (2005), "no distortion at the top," has no counterpart here. With transfers, efficient provision occurs *forever* once the agent first reports a high type. Furthermore, 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 communication. Moreover, both long-run outcomes can arise. In summary, with transfers, inefficiencies are frontloaded

[32]Based on the results in the discrete-time model, we can extend the results to the continuous-time limit, using a Poisson formulation. This formulation is available in section E of the online appendix. We choose the discrete-time model because the continuous-time model admits no counterpart to the i.i.d. case and we derive almost all our results in the discrete-time model.

to the greatest extent possible. Without transfers, they are backloaded.

The difference can be understood as follows. First, and importantly, results in Battaglini (2005) rely on having revenue maximization be 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 rather, 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, since 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*, allowing for payments on the spot would suffice to restore efficiency if efficiency 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 be made only from the principal to the agent. The principal maximizes the social surplus net of any payments.[33] In this case, Lemma 7 in appendix C shows 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 (*i.e.*, 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.

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

## 6 Discussion

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

**Moral hazard.** As discussed in the introduction, the model can be extended to incorporate moral hazard. Consider the following model of lending (see, *e.g.*, Bulow and Rogoff, 1989): 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 $v + \beta$ with $\eta, \beta > 0$. Here, $\eta$ is the loss to the principal if the agent defaults, and $\beta$ is the agent's benefit from defaulting.

We consider the first-best benchmark: the principal observes the agent's type and dictates the agent's 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. We can readily verify that the second-best value function, $W(U)$, is exactly the same as before, as long as the discount factor exceeds some constant. Intuitively, the agent is deterred from defaulting if his promised utility is large enough. At the same time, for $U < \underline{U}$, the principal has flexibility in when to grant the loan after high reports. In particular, if $U$ is very small, the principal can freeze lending for several rounds and let $U$ move closer to $\underline{U}$. She grants the loan only when $U$ is close enough to $\underline{U}$. The agent will not default since defaulting will wipe out his promised utility.

**Renegotiation-proofness.** The optimal policy, as described in sections 4 and 5, 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-proofness (Farrell and Maskin, 1989) implies a lower boundary on the utility vectors visited (except in the event that $\mu$ 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 (which must serve as a lower reflecting boundary) or the upper absorbing boundary, $\mu$.

**More types.** There is no conceptual difficulty in accommodating more than two 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, for some fixed pair of the low type's and the high type's utilities? It appears that the convex hull of utilities from frontloading and backloading policies traces out the lowest utility that a medium type can obtain for any such pair, but the set of incentive-compatible utilities has full dimension. The maximum utility of the medium type obtains when one of his incentive constraints binds, but according to the binding constraint, there are two possibilities, 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.

# 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] Alderman, H. (2002). "Do local officials know something we don't? Decentralization of targeted transfers in Albania," *Journal of Public Economics*, **83**, 375–404.

[3] Bardhan, P. and D. Mookherjee (2005). "Decentralizing antipoverty program delivery in developing countries," *Journal of Public Economics*, **89**, 675–704.

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

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

[6] Boleslavsky, R. and D. Kelly (2014). "Dynamic Regulation Design Without Payments: The Importance of Timing," *Journal of Public Economics*, **120**, 169–180.

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

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

[9] Clementi, G. and H. Hopenhayn (2006). "A Theory of Financing Constraints and Firm Dynamics," *The Quarterly Journal of Economics*, **121**, 229–265.

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

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

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

[13] Escobar, J. and J. Toikka (2013). "Efficiency in Games with Markovian Private Information," *Econometrica*, **81**, 1887–1934.

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

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

[16] Frankel, A. (2016). "Discounted Quotas," *Journal of Economic Theory*, **166**, 396–444.

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

[18] Garrett, D. and A. Pavan (2015). "Dynamic Managerial Compensation: A Variational Approach," *Journal of Economic Theory*, **159**, 775–818.

[19] Green, E. (1987). "Lending and the Smoothing of Uninsurable Income," in *Contractual Arrangements for Intertemporal Trade* (E. C. Prescott and N. Wallace, Eds.), Minnesota Press, Minnesota, 1987.

[20] Guo, Y. (2016). "Dynamic Delegation of Experimentation," *American Economic Review*, **106**, 1969–2008.

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

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

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

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

[25] 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.

[26] 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.

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

[28] Lipnowski, E. and J. Ramos (2017). "Repeated Delegation," working paper, University of Chicago.

[29] Malenko, A. (2018). "Optimal Dynamic Capital Budgeting," working paper, MIT Sloan.

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

[31] Möbius, M. (2001). "Trading favors," working paper, Harvard University.

[32] Ng, C. (2011). "Online Mechanism and Virtual Currency Design for Distributed Systems," PhD dissertation, Harvard.

[33] Olszewski, M. and M. Safronov (2018a). "Efficient Cooperation by Exchanging Favors," *Theoretical Economics*, *forthcoming*.

[34] Olszewski, M. and M. Safronov (2018b). "Efficient Chip Strategies in Repeated Games," *Theoretical Economics*, *forthcoming*.

[35] Pavan A., I. Segal, and J. Toikka (2014). "Dynamic Mechanism Design: A Myersonian Approach," *Econometrica*, **82**, 601–653.

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

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

[38] Saraswathi, A., Y. Kalaashri, and S. Padmavathi (2015). "Dynamic Resource Allocation Scheme in Cloud Computing," *Procedia Computer Science*, **47**, 30–36.

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

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

[41] 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.

[42] 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.

# A Proofs for section 4

*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 $\mu$ 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 $p_l$.

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

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

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 $I \subset [\overline{U}, \mu]$, for $U \in I$,

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

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 $[\overline{U}, \mu]$, 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 $\mu$). 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 $\mu$). 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 $\varepsilon$ 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.

The rest of the proof shows that $W$ is differentiable. Since $W$ is differentiable on $U \in [0, \underline{U})$, we focus on $U \in [\underline{U}, \mu]$. We let $w^-(U)$ be the left derivative of $W(U)$, and $w^+(U)$ the right derivative. Concavity of $W$ implies that (i) both $w^-$ and $w^+$ are monotone decreasing, (ii) $w^-(U) \geqslant w^+(U), \forall U$, and (iii) $w^+(U) \geqslant w^-(U'), \forall U' > U$. We also know that $w^-(\mu) = \frac{l-c}{l}$, and that $w^-(U), w^+(U) \in \left[\frac{l-c}{l}, \frac{h-c}{h}\right]$.

We first argue that $W$ is differentiable at $\underline{U}$. Since $W = \overline{W}$ for $U \in [0, \underline{U}]$, $w^-(\underline{U}) = \frac{h-c}{h}$. On the other hand,

$$w^+(\underline{U}) = qw^+(U_h(\underline{U})) + (1-q)w^+(U_l(\underline{U})).$$

Since $U_l(\underline{U}) = \underline{U}$ and $U_h(\underline{U}) < \underline{U}$, we have

$$w^+(\underline{U}) = w^+(U_h(\underline{U})) = \frac{h-c}{h}.$$

This shows that $W$ is differentiable at $\underline{U}$.

If $w^- = w^+$ for all $U$, then $W$ is differentiable. Suppose that $w^-(U') - w^+(U') > 0$ for some $U'$. Then there are only finitely many utility levels at which the difference $(w^- - w^+)$ is weakly larger than the difference at $U'$. We let $K = \max_U (w^-(U) - w^+(U))$, and $A = \{U : w^-(U) - w^+(U) = K\}$. We next argue that the presumption that $K > 0$ and $A \neq \emptyset$ leads to a contradiction.

1. For any $U \in [\underline{U}, (1-\delta)l + \underline{U}]$, we have $p_h = 1, p_l = 0, U_h(U) = \frac{U-(1-\delta)\mu}{\delta}, U_l(U) =$

$\frac{U-(1-\delta)\underline{U}}{\delta}$. Moreover, $U_h(U) \leqslant \underline{U}$. This implies that

$$\begin{aligned} w^-(U) - w^+(U) &= q\left(w^-(U_h(U)) - w^+(U_h(U))\right) + (1-q)\left(w^-(U_l(U)) - w^+(U_l(U))\right) \\ &= (1-q)\left(w^-(U_l(U)) - w^+(U_l(U))\right). \end{aligned}$$

The last step follows from the fact that $w^-(U_h(U)) = w^+(U_h(U))$ for all $U_h(U) \leqslant \underline{U}$. This implies that $A \cap [\underline{U}, (1-\delta)l + \underline{U}] = \emptyset$. Otherwise, if there is some $U \in A \cap [\underline{U}, (1-\delta)l + \underline{U}]$, then the difference $(w^- - w^+)$ at $U_l(U)$ is greater than $K$.

2. For any $U \in [\delta l + \underline{U}, \mu]$, we have $p_h = 1, p_l = \frac{U-\delta l-\underline{U}}{(1-\delta)l}, U_h(U) = \frac{U-(1-\delta)\mu}{\delta}, U_l(U) = \mu$. For $U \in (\delta l + \underline{U}, \mu]$, we have

$$\begin{aligned} w^-(U) &= (1-q)\frac{l-c}{l} + qw^-(U_h(U)), \\ w^+(U) &= (1-q)\frac{l-c}{l} + qw^+(U_h(U)). \end{aligned} \tag{11}$$

This implies that

$$w^-(U) - w^+(U) = q\left(w^-(U_h(U)) - w^+(U_h(U))\right). \tag{12}$$

For $U = \delta l + \underline{U}$, $w^+(U)$ is given by the same formula as in (11). On the other hand,

$$w^-(U) = (1-q)w^-(\mu) + qw^-(U_h(U)) = (1-q)\frac{l-c}{l} + qw^-(U_h(U)).$$

This means that (12) applies to $U = \delta l + \underline{U}$ as well. Based on (12), we conclude that $A \cap [\delta l + \underline{U}, \mu] = \emptyset$. Otherwise, if there is some $U \in A \cap [\delta l + \underline{U}, \mu]$, then the difference $(w^- - w^+)$ at $U_h(U)$ is greater than $K$.

3. For any $U \in ((1-\delta)l + \underline{U}, \delta l + \underline{U})$, we have $p_h = 1, p_l = 0, U_h(U) = \frac{U-(1-\delta)\mu}{\delta}, U_l(U) = \frac{U-(1-\delta)\underline{U}}{\delta}$. This implies that

$$w^-(U) - w^+(U) = q\left(w^-(U_h(U)) - w^+(U_h(U))\right) + (1-q)\left(w^-(U_l(U)) - w^+(U_l(U))\right).$$

Therefore, if $U \in A \cap ((1-\delta)l + \underline{U}, \delta l + \underline{U})$, then $U_h(U), U_l(U)$ must be in $A$ as well. Let $\tilde{U}$ be the maximal utility in $A \cap ((1-\delta)l + \underline{U}, \delta l + \underline{U})$. Then this leads to a contradiction: (i) either $U_l(\tilde{U})$ is also in $A \cap ((1-\delta)l + \underline{U}, \delta l + \underline{U})$, which contradicts the presumption that $\tilde{U}$ is the maximum, (ii) or $U_l(\tilde{U})$ is in $[\delta l + \underline{U}, \mu]$, which contradicts the conclusion that $A \cap [\delta l + \underline{U}, \mu] = \emptyset$ in the previous step.

■

**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 5

*Proof of Lemma 4.* We focus on the case with $1 - \rho_h > \rho_l$. We first make a few observations about $V$:

1. Because $(0,0)$ is delivered by never providing the good and $(\mu_h, \mu_l)$ is delivered by always providing the good, $(0,0)$ and $(\mu_h, \mu_l)$ are the lowest and highest utility pairs that the principal can deliver. It follows that $V \subset [0, \mu_h] \times [0, \mu_l]$. Moreover, for any $(U(h), U(l)) \in V$, $0 < U(h) < \mu_h$ if and only if $0 < U(l) < \mu_l$.

2. Payoffs in $\text{co}\{b^m, f^m : m \geqslant 0\}$ can be implemented without eliciting the agent's information. It follows that $\text{co}\{b^m, f^m : m \geqslant 0\} \subset V$.

3. The set $V$ is convex, because any convex combination of $U, U' \in V$ can be implemented by flipping a coin before round zero.

4. Let $SE(x) := \min_{(x,U(l)) \in V} U(l)$ and $NW(x) := \max_{(x,U(l)) \in V} U(l)$. Then $SE(x)$ and $NW(x)$ are the lowest and the highest payoffs to a current low type, given that a current high type gets $x$. The convexity of $V$ implies that $SE(\cdot)$ is convex and $NW(\cdot)$ is concave.

5. We next argue that $SE(x) = \frac{l}{h}x$ for $x \in [0, (1-\delta)h]$. First, $SE(x) \leqslant \frac{l}{h}x$ for $x \in [0, (1-\delta)h]$ because the principal can provide the good in the current round with probability $\frac{x}{(1-\delta)h}$. Second, we argue that $SE(x)$ cannot be lower than $\frac{l}{h}x$. For any policy that gives a current high type $x$, a current low type can secure $\frac{l}{h}x$ by reporting in the same way

as the current high type does. This is because the current low type's and the current high type's future values converge, and the ratio of their future values is larger than $\frac{l}{h}$. A similar argument shows that $NW(x) = \mu_l + \frac{l}{h}(x - \mu_h)$ for $x \in [\mu_h - (1-\delta)h, \mu_h]$.

6. Throughout the proof, we let $SE'(\cdot)$ be the right-derivative and $NW'(\cdot)$ be the left-derivative. It follows from (4.) and (5.) that $SE'(\cdot) \geqslant \frac{l}{h}$ and $NW'(\cdot) \geqslant \frac{l}{h}$.

For any $x \in (0, \mu_h)$, $SE(x)$ is given by the following program:

$$
\begin{aligned}
\min_{p_h, p_l \in [0,1], U_h, U_l \in V} & U(l) \\
s.t.,\ x &= (1-\delta)p_h h + \delta\left((1-\rho_h)U_h(h) + \rho_h U_h(l)\right) \\
&\geqslant (1-\delta)p_l h + \delta\left((1-\rho_h)U_l(h) + \rho_h U_l(l)\right) \\
U(l) &= (1-\delta)p_l l + \delta\left(\rho_l U_l(h) + (1-\rho_l)U_l(l)\right) \\
&\geqslant (1-\delta)p_h l + \delta\left(\rho_l U_h(h) + (1-\rho_l)U_h(l)\right).
\end{aligned}
$$

The four constraints are $PK_h, IC_h, PK_l, IC_l$, respectively. Because we assume that $x \in (0, \mu_h)$, the minimal $U(l)$ must be strictly positive. It follows that in the solution either $p_l > 0$ or $U_l > (0,0)$. We make three observations regarding the solution of this program:

1. $IC_l$ must bind. If not, we can reduce $p_l$ or $U_l$. This has no impact on $PK_h$ and makes $IC_h$ easier to satisfy. This reduction will lower the objective $U(l)$.

2. $U_h$ must be on the lower boundary $SE(\cdot)$. If not, we can replace $U_h$ with

$$
\left(U_h(h) + \varepsilon, U_h(l) - \varepsilon \frac{1-\rho_h}{\rho_h}\right)
$$

for small $\varepsilon > 0$. This has no impact on the promised utility to the current high type, and thus no impact on $PK_h, IC_h$. This lowers the RHS of $IC_l$, so allows us to lower the objective $U(l)$.

3. Either $p_h = 1$ or $U_h = (0,0)$. If not, we can replace $p_h, U_h$ with

$$
p_h = p_h + \frac{\delta(h(1-\rho_h) + l\rho_h)}{(1-\delta)h^2}\varepsilon, \quad U_h(h) = U_h(h) - \varepsilon, \quad U_h(l) = U_h(l) - \frac{l}{h}\varepsilon.
$$

This change has no impact on $PK_h, IC_h$, but lowers the RHS of $IC_l$. Hence, we can further lower the objective $U(l)$. We can replace $U_h$ with the new utility pair because we have shown that $SE'(\cdot)$ is at least $\frac{l}{h}$.

These observations show that the utility pair $(x, SE(x))$ is delivered by a policy in which (i) the current low type is indifferent between reporting low and high, (ii) the continuation utility pair after high report is still on the lower boundary $SE(\cdot)$, (iii) $p_h$ is as high as possible. Combining (i) and (ii), we conclude that the current low type's payoff can be calculated as if he always reports high. According to (iii), the good is provided as long as possible after high reports. This shows that the lower boundary $SE(\cdot)$ is characterized by the frontloading policies.

Using a similar argument, we next show that $NW(\cdot)$ is given by the backloading policies. For any $x \in (0, \mu_l)$, $NW^{-1}(x)$ is given by the following program:

$$
\begin{aligned}
\min_{p_h, p_l \in [0,1], U_h, U_l \in V} & U(h) \\
s.t., \; U(h) &= (1-\delta)p_h h + \delta\left((1-\rho_h)U_h(h) + \rho_h U_h(l)\right) \\
&\geqslant (1-\delta)p_l h + \delta\left((1-\rho_h)U_l(h) + \rho_h U_l(l)\right) \\
x &= (1-\delta)p_l l + \delta\left(\rho_l U_l(h) + (1-\rho_l)U_l(l)\right) \\
&\geqslant (1-\delta)p_h l + \delta\left(\rho_l U_h(h) + (1-\rho_l)U_h(l)\right).
\end{aligned}
$$

Because we assume that $x \in (0, \mu_l)$, the minimal $U(h)$ must be strictly positive. It follows that in the solution either $p_h > 0$ or $U_h > (0, 0)$. We again make three observations:

1. $IC_h$ must bind. If not, we can reduce $p_h$ or $U_h$. This has no impact on $PK_l$ and makes $IC_l$ easier to satisfy. This reduction will lower the objective $U(h)$.

2. $U_l$ must be on the upper boundary $NW(\cdot)$. If not, we can replace $U_l$ with

$$\left(U_l(h) - \varepsilon, U_l(l) + \varepsilon \frac{\rho_l}{1-\rho_l}\right)$$

   for small $\varepsilon > 0$. This has no impact on the promised utility to the current low type, and thus no impact on $PK_l, IC_l$. This lowers the RHS of $IC_h$, so allows us to lower the objective $U(h)$.

3. Either $p_l = 0$ or $U_l = (\mu_h, \mu_l)$. If not, we can replace $p_l, U_l$ with

$$p_l = p_l - \frac{\delta(h\rho_l + l(1-\rho_l))}{(1-\delta)hl}\varepsilon, \;\; U_l(h) = U_l(h) + \varepsilon, \;\; U_l(l) = U_l(l) + \frac{l}{h}\varepsilon.$$

   This change has no impact on $PK_l, IC_l$, but lowers the RHS of $IC_h$. Hence, we can further lower the objective $U(h)$. We can replace $U_l$ with the new utility pair because we have shown that $NW'(\cdot)$ is at least $\frac{l}{h}$.

These observations show that the utility pair $(NW^{-1}(x), x)$ is delivered by a policy in which (i) the current high type is indifferent between reporting low and high, (ii) the continuation utility pair after low report is still on the upper boundary $NW(\cdot)$, (iii) $p_l$ is as low as possible. This shows that the upper boundary $SE(\cdot)$ is characterized by the backloading policies. ■

## 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 5.1 for the entire $V$.

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

$$\underline{v}_h^m = \mathbf{E}\left[\delta^m \underline{U}(v_m) \mid v_0 = h\right], \quad \underline{v}_l^m = \mathbf{E}\left[\delta^m \underline{U}(v_m) \mid v_0 = l\right]. \tag{13}$$

Intuitively, this is the payoff from waiting for $m$ rounds from initial value $h$ or $l$, and then getting $\underline{U}$. (See Footnote 24 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 $\{\underline{v}^m\}_{m\geqslant 0}$. Figure 6 illustrates this construction. Note that $\underline{V}$ has a nonempty interior if and only if $\rho_l$ 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 $\phi$,*

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

*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 $m \geqslant 1$, the utility $\underline{v}^m$ can be delivered by not supplying the unit in the current round and setting the future utility to be $\underline{v}^{m-1}$, regardless of the report. The agent's promised utility becomes $\underline{U}(= \underline{v}^0)$ after $m$ 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 $\hat{u}^m := (\hat{u}^m_h, \hat{u}^m_l)$, $m \geqslant 0$ as follows:

$$\hat{u}^m_h = \mathbf{E}\left[(1-\delta)\sum_{n=1}^{m} \delta^n v_n \mid v_0 = h\right], \quad \hat{u}^m_l = \mathbf{E}\left[(1-\delta)\sum_{n=1}^{m} \delta^n v_n \mid v_0 = l\right]. \tag{14}$$

This is the utility vector if the principal supplies the unit from round 1 to round $m$, independently of the reports. Note that the unit is not supplied in round zero. We note that $\hat{u}^0 = (0,0)$ and $\hat{u}^m$ is an increasing sequence (in both coordinates) contained in $V$, where $\lim_{m\to\infty} \hat{u}^m = b^1$. The sequence $\{\hat{u}^m\}_{m\geqslant 0}$ 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 $\{f^m\}_{m\geqslant 0}$ and $\{b^m\}_{m\geqslant 0}$ 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*

$$p_h = \min\left\{1, \frac{U(h)}{(1-\delta)h}\right\}, \quad p_l = \max\left\{0, 1 - \frac{\mu_l - U(l)}{(1-\delta)l}\right\}, \tag{15}$$

*and $U_h \in P_f$. If $U \in V_f$, then $U_l \in P_f$. If $U \in V_b$, $U_l$ is chosen such that $IC_h$ binds. (This implies that $U_l \in P_b$ for $U \in P_b$.)*

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 will 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 $(\hat{u}^m_h, \hat{u}^m_l) \in P$, the high type is promised the expected utility from forgoing the unit in round zero and consuming the unit from round 1 to $m$. 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

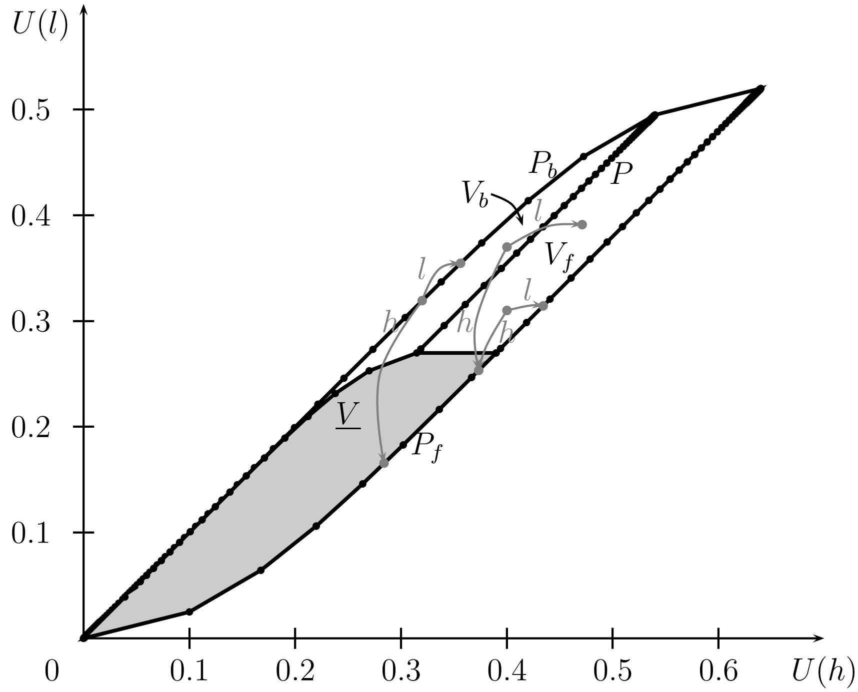


Figure 6: Incentive-compatible set $V$ and the optimal policy for $(\delta, \rho_h, \rho_l, l, h) = (9/10, 1/3, 1/4, 1/4, 1)$.

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 (see also Footnote 18). 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 $U \in V \setminus \underline{V}$, the vector $(p_h, p_l) = (1, 0)$ is used (unless $U$ is too close to $(0, 0)$ or $(\mu_h, \mu_l)$), 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 to 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 $(\mu_h, \mu_l)$ 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

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

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 $m$ conditional on the current type into (13), we obtain the analytic expression of $\{\underline{v}^m\}_{m \geqslant 0}$:

$$\underline{v}_h^m = \delta^m \left((1-q)\underline{U}(l) + q\underline{U}(h) + (1-q)(1-\rho_h-\rho_l)^m(\underline{U}(h) - \underline{U}(l))\right),$$
$$\underline{v}_l^m = \delta^m \left((1-q)\underline{U}(l) + q\underline{U}(h) - q(1-\rho_h-\rho_l)^m(\underline{U}(h) - \underline{U}(l))\right).$$

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

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

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 $m \geqslant 1$, $\underline{v}^m$ can be supported by setting $p_h = p_l = 0$ and $U_h = U_l = \underline{v}^{m-1}$. Therefore, $\underline{v}^m$ 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.

We next show that $V_0 \subset \underline{V}$. Consider a sequence of programs by setting $Z^0 = V$ and defining recursively the supremum score in direction $(\lambda_1, \lambda_2)$ given $Z^{m-1}$ as $K((\lambda_1, \lambda_2), Z^{m-1}) = \sup_{p_h, p_l, U_h, U_l} \lambda_1 U(h) + \lambda_2 U(l)$, subject to $(PK_h)$–$(IC_l)$, $p_l = 0$, $p_h \in [0,1]$, $U_h, U_l \in Z^{m-1}$, $\lambda \cdot U_h \leqslant \lambda \cdot U$ and $\lambda \cdot U_l \leqslant \lambda \cdot U$. Let the set $\mathcal{B}(Z^{m-1})$ be given by

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

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

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

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 $(\mu_h, \mu_l)$. 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 $\delta$ 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 $P_f$

$$\tilde{w}_h'(X(U)) \geqslant \tilde{w}_l'(Y(U)). \tag{16}$$

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 (15), *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

$$(U_{h1}(x), U_{l1}(x)) = (U_{h2}(y), U_{l2}(y)) = (U(h), U(l)). \tag{17}$$

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

$$\begin{aligned} \hat{\hat{w}}_h'(x) &= (1-\rho_h)U_{h1}'(x)\hat{w}_h'\left(\frac{U_{h1}(x)-(1-\delta)h}{\delta}\right) + \rho_h U_{l1}'(x)\hat{w}_l'\left(\frac{U_{l1}(x)}{\delta}\right), \\ \hat{\hat{w}}_l'(y) &= \rho_l U_{h2}'(y)\hat{w}_h'\left(\frac{U_{h2}(y)-(1-\delta)h}{\delta}\right) + (1-\rho_l)U_{l2}'(y)\hat{w}_l'\left(\frac{U_{l2}(y)}{\delta}\right). \end{aligned} \tag{18}$$

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

$$\hat{w}_h'\left(\frac{U(h)-(1-\delta)h}{\delta}\right) \geqslant \hat{w}_l'\left(\frac{U(l)}{\delta}\right). \tag{19}$$

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

$$(1-\rho_h)U_{h1}'(x) + \rho_h U_{l1}'(x) = \rho_l U_{h2}'(y) + (1-\rho_l)U_{l2}'(y) = 1, \tag{20}$$

$$(1-\rho_h)U_{h1}'(x) - \rho_l U_{h2}'(y) \geqslant 0. \tag{21}$$

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 (20) and (21) hold.

Given (18), (20), 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 (17) to (21), 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 (18):

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

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 $(\mu_h, \mu_l)$ 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:

$$\hat{\hat{W}}(U(h)+\varepsilon, U(l), \phi) - \hat{\hat{W}}(U, \phi) \geqslant \hat{\hat{W}}\left(U(h), U(l) + \frac{\phi}{1-\phi}\varepsilon, \phi\right) - \hat{\hat{W}}(U, \phi), \text{ where } \varepsilon > 0. \tag{22}$$

The left-hand side of (22) equals

$$\delta\phi\left(\hat{W}\left(\tilde{U}_h(h), \tilde{U}_h(l), 1-\rho_h\right) - \hat{W}(U_h(h), U_h(l), 1-\rho_h)\right), \tag{23}$$

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 (22) equals

$$\delta(1-\phi)\left(\hat{W}\left(\tilde{U}_l(h), \tilde{U}_l(l), \rho_l\right) - \hat{W}(U_l(h), U_l(l), \rho_l)\right), \tag{24}$$

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 (23) is greater than (24). 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 $\varepsilon$ goes to 0, we obtain that (22) is equivalent to

$$\hat{w}_h'\left(\frac{U(h)-(1-\delta)h}{\delta}\right) \geqslant \hat{w}_l'\left(\frac{U(l)}{\delta}\right), \quad \forall U \in V \setminus (\underline{V} \cup V_h \cup V_l). \tag{25}$$

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 (25) 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 (25) holds given that (i) $\hat{w}_h, \hat{w}_l$ are concave; (ii) inequality (16) 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 (22) is given by

$\phi\varepsilon(1-c/l)$. Inequality (22) 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 subproblems: (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 subproblem. 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 subproblem.

2. For the second subproblem, 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 $\gamma_h, \gamma_l$ 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:

$$\gamma_h \frac{\partial W(U, \rho_l)}{\partial U(h)} + \gamma_l \frac{\partial W(U, \rho_l)}{\partial U(l)} \leqslant \frac{(1-\delta)(c-l)}{\delta}, \tag{26}$$

where the left-hand side is the directional derivative of $W(U,\rho_l)$ along the vector $(\gamma_h,\gamma_l)$. We first show that (26) 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 (26) in $\tilde{w}_h'$ and $\tilde{w}_l'$, it is sufficient to show that (26) 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 (26) equals $-(1-\delta)l\tilde{w}_h'/\delta$, so (26) holds for any $\tilde{w}_h' \in [1-c/l, 1-c/h]$. In case (ii), the left-hand side of (26) 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 (26) holds. (Using similar arguments, we can show that (26) holds for all $U \in P_f \cap V_h$ or $U \in P_f \cap V_l$.)

Lastly, we want to show that (26) holds for $U \in V_b$. Note that $W(U,\rho_l)$ is concave on $V$. Therefore, its directional derivative along the vector $(\gamma_h,\gamma_l)$ 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 $(\gamma_h,\gamma_l)$ points towards the interior of $V$, (26) 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 $\phi_0$ 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 $\phi_0$. 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 (16), 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 $\phi_0'$ must be higher than $U^*$, the optimal initial promise given $\phi_0$. ■

# C Transfers with Limited Liability

Here, we consider the case in which 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 $\mu$, 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 $\varepsilon$ 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 $\mu$, 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 $c - l \leqslant l$, 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. We argue that for $U \in [0, \mu]$, the principal's value function is the same as before. This trivially holds for $U = 0, \mu$. We first show that the principal never makes transfers if $U_l, U_h < \mu$. With abuse of notation, let $t_v$ denote the current-round transfer after $v$ report. Suppose $U_v < \mu$ and $t_v > 0$. We can increase $U_v$ ($v = l$ or $h$) by $\varepsilon$ and reduce $t_v$ 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 $U < \mu$. Suppose $U_l = \mu$, $p_l < 1$ 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 $\mu$. The optimal scheme is to provide the unit forever. ■

# Online appendix

**"Dynamic Allocation without Money"**
**by Yingni Guo and Johannes Hörner**

# D Omitted Proofs

## D.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.

# E 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. Here, we consider the continuous-time limit of the problem, by scaling the transition probabilities according to the usual Poisson limit. Based on Lemmas 4 and 6 and Theorems 2 and 3, we can extend the results in the discrete-time model to the continuous-time model. We start with an overview of this model and the results, relegating some of the details and proofs to section E.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 $\Delta$ 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-compatible set $V$ is characterized by the frontloading and backloading policies. By taking the limit (as $\Delta \to 0$) of the formulas for $\{f^m, b^m\}_{m \in \mathbf{N}}$, we obtain the boundaries of $V$. In particular, the frontloading boundary is given in parametric form by[34]

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

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."[35]

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

[34]See appendix E.1 for the analytic expressions.

[35]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

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

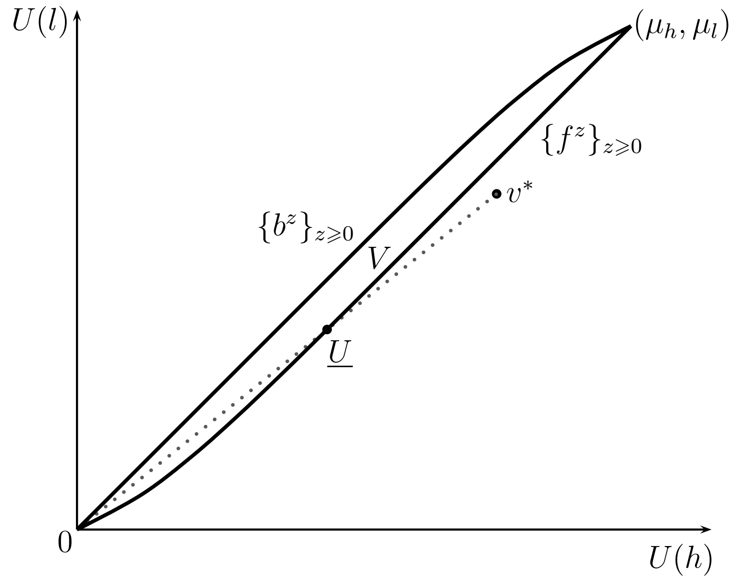


Figure 7: Incentive-compatible set $V$ for $(\lambda_l, \lambda_h, r, l, h) = (1, 1, 1/4, 1, 5/2)$.

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.*, $\lambda_l$ 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.

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 forgoing the unit, he is compensated by a term that equals the rate of

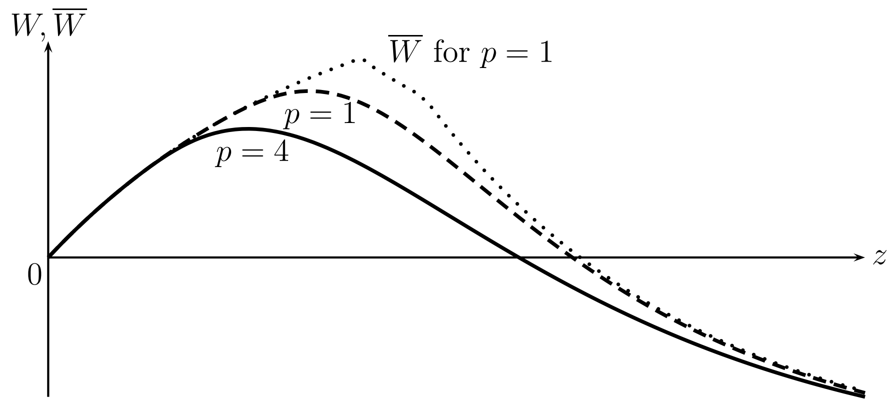


Figure 8: Value function and complete-information payoff as a function of $z$, for $(\lambda_l, \lambda_h, r, l, h, c) = (1/p, 1/p, 1/4, 1, 5/2, 2)$ and $p = 1, 4$.

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. 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 $\phi_t$ and the promise to the agent. Continuous time allows to explicitly solve for the principal's value function, although the expression is quite complicated. 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 E.1 for a derivation of the value function. 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 $p = 1$.[36]

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

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

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 quota mechanisms as in Jackson and Sonnenschein (2007).

**Lemma 8.** *It holds that*

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

Lastly, we examine the impact of varying $h$ or $l$ on the value function $W(z)$ numerically, for any fixed $\lambda_h, \lambda_l, r$ and $c$. Since $q$ equals $\lambda_l/(\lambda_h + \lambda_l)$, it is fixed as well. 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.

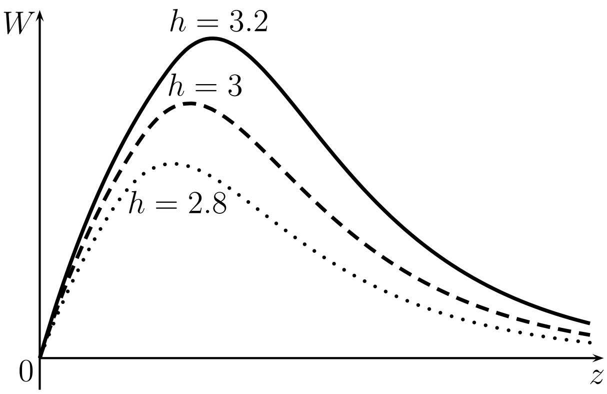


Figure 9: Value function for $(\lambda_l, \lambda_h, r, c) = (1, 1, 1/2, 2)$ and $\mu = 2$.

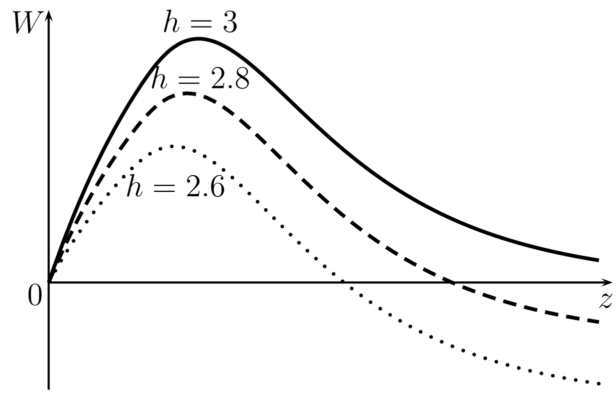


Figure 10: Value function for $(\lambda_l, \lambda_h, r, c) = (1, 1, 1/2, 2)$ and $l = 1$.

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

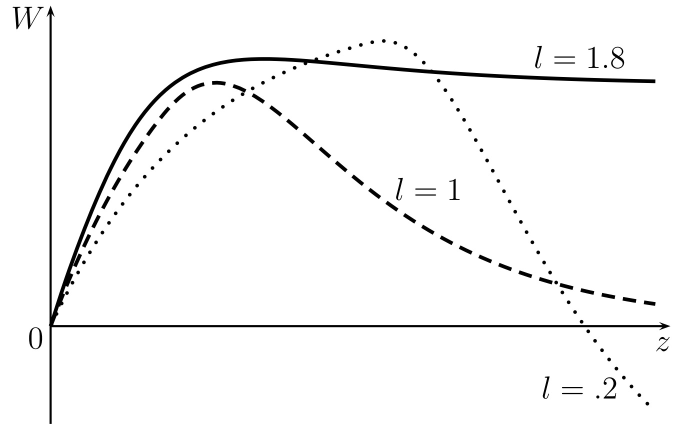


Figure 11: Value function for $(\lambda_l, \lambda_h, r, c) = (1, 1, 1/2, 2)$ and $h = 3$.

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.

## E.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 $\Delta$ to 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 $(\mu_h, \mu_l)$ 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[37]

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

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

$$\begin{aligned} &g(z)((r+\lambda_h)W'(z) + W''(z)) \\ =&(h-l)q\lambda_h e^{-(\lambda_h+\lambda_l)z}W'(z) + \mu(r(\lambda_h+\lambda_l)W(z) + \lambda_l W'(z) - r\lambda_l(h-c)). \end{aligned} \tag{27}$$

[37]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.

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 $\overline{W}_1(z)$:

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

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 $\mu - c$ 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:[38]

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

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

*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

[38]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 $\overline{W}_1(\hat{z})$.

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 $f_h^z, f_l^z$ 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 D.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 $\eta$. 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. ■