RATIONALITY
& COMPETITION
CRC TRR 190

Of Restarts and Shutdowns: Dynamic Contracts with Unequal Discounting

Ilia Krasikov (Penn State University;)
Rohit Lamba (Penn State University)
Thomas Mettral (HU Berlin)

Discussion Paper No. 94

April 24, 2018

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

Of restarts and shutdowns: dynamic contracts with unequal discounting*

Ilia Krasikov Rohit Lamba Thomas Metrall

March 2018

Abstract

A large supplier (principal) contracts with a small firm (agent) to repeatedly provide working capital in return for payments. The total factor productivity of the agent is private and follows a Markov process. Moreover, the agent is *less patient* than the principal. We solve for the optimal contract in this environment. Distortions are pervasive and efficiency unattainable. The optimal contract is characterized by two key properties: restart and shutdown, which capture various aspects of contracts offered in the marketplace. The optimal distortions are completely pinned down by the number of low TFP shocks since the last high shock. Once a high shock arrives, the contract loses memory and repeats the same cycle, we call this endogenous resetting feature *restart*. If ex ante agency frictions are high, the principal commits to not serving the low type, we call this *shutdown*. The principal prefers a patient agent if the interim agency friction, as measured by the persistence of the private information is large, and she prefers

*Krasikov: Penn State University, izk113@psu.edu; Lamba: Penn State University, rlamba@psu.edu; Metrall: Humboldt University of Berlin, thomas.schacherer@hu-berlin.de. We are indebted to Leeat Yariv for conversations that instigated this paper, Vitor Farinha Luz for fruitful initial discussions and Daniel Baron for detailed feedback. We are also grateful to Nageeb Ali, Marco Battaglini, Vijay Krishna, Roland Strausz and participants at the third annual conference on relational contracting at Northwestern University for their comments. Shomak Chakrabarti provided excellent research assistance, thanks are due to him as well. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 is gratefully acknowledged.

an impatient agent if it is small. Finally, when global incentive constraints bind, we (i) provide the complete recursive solution, and (ii) characterize a simpler incentive compatible contract that is approximately optimal.

1 Introduction

Ever so often, as the juggernaut of a literature ferries along, we must stop it in the tracks, to evaluate certain assumptions that we may then consider standard. One such assumption in dynamic models of mechanism design and agency models of dynamic contracting is that all parties have an equal rate of time preference. A significant parametric restriction, it is at times a simplifying device and at other times a modeling habit. Allowing for unequal discounting reveals to the economist the robustness of her results to the wider parametric range, and in the process she may uncover hitherto unexplored dynamic tradeoffs.

This paper studies a dynamic screening model with persistent private information where the principal is more patient than the agent. One may think of a venture capitalist investing in a startup, a government deciding on tax schedules with objective of redistribution amongst a population, or an intermediary supplying a vital input to a firm to produce a final good. We focus on the last interpretation, but urge the reader to think of the framework more broadly, distilling through it two key economic forces: unequal discounting and persistent agency frictions. The interaction of the two produces intertemporal gains from time scripted trade and intertemporal costs of incentive provision.

There are at least three motivations for analyzing the said model. First, in many long-term contractual situations constrained by private information one party is "financially bigger" or more integrated in capital markets than the other; an easy way of capturing this asymmetry is unequal discounting.[1] In fact, the literature is rife with evidence of limited access to finance as a binding constraint in economic transactions.[2] What kind of contracts do we expect to observe in

[1]We have $\delta_P = e^{-r}$ and $\delta_A = e^{-s}$ where r and s are respectively the interest rates faced by the principal and agent in the market with $s \geq r$, and the exponential representation approximates a continuously compounded principal amount.

[2]In a survey of 1050 CFOs across the US, Europe and Asia, Campello et al. (2010) find a considerable impact of credit constraints on real firm behavior in the aftermath of the Great Recession. Deaton (1991) and Carroll (1992) make the theoretical and empirical case respectively of the importance of liquidity constraints in analyzing consumption. In the celebrated

such environments? Second, behaviorally speaking, it is natural for two parties in a contract to have different time horizons, or different assessment of the probability survival of the transaction; both situations can be represented, at least to a reduced form, by unequal discounting.[3] And, third, from a more theoretical perspective, how robust are the predictions in the burgeoning literature on dynamic mechanism design to the violations of the assumption of equal discounting? How do allocative distortions evolve and influence long-term efficiency?[4]

We are not the first ones to study dynamic contracting with unequal discounting, however, to the best of our knowledge, this is the first paper to explore its implications in a dynamic screening or adverse selection model with persistent private information.[5] The word persistent is imperative for it adds a realistic dimension to the underlying agency frictions[6], and as we will see later, it also adds memory to allocative distortions. The realism though comes with a technical challenge- it introduces potentially binding global incentive constraints.

The formal model entails a "small" firm (agent) with a private production technology, its total factor productivity (TFP) changes periodically according to a two state Markov process, and a "large" supplier (principal) of capital that is critical for production. The principal is more patient than the agent. A contract here is a dynamic menu of capital allocations to the agent in return for payments to the principal. We solve for the profit maximizing contract of the principal subject to incentive compatibility and individual rationality constraints for the agent.

In order to relax future incentive constraints and thereby reduce information rents, the large supplier wants to backload payoffs for the small firm as much

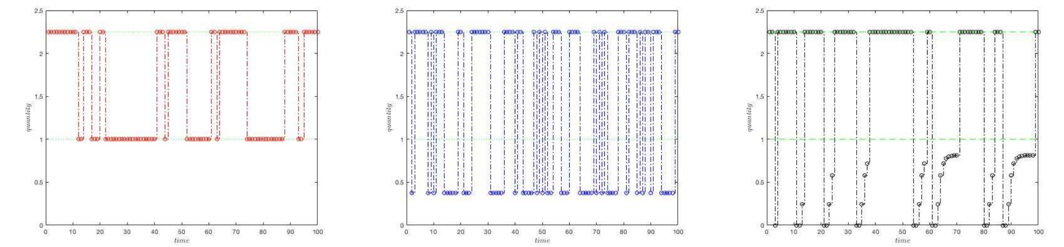
Eaton and Gersovitz (1981), the borrower faces a higher interest rate spread with incomplete markets and defaulting risk.

[3]For example, Edmans et al. (2017) document misaligned intertemporal incentives in corporations between the shareholders and CEOs.

[4]See excellent surveys by Vohra (2012), Krähmer and Strausz (2015), Pavan (2016), and Bergemann and Välimäki (2017) on dynamic mechanism design models where the principal and agent(s) have the same rate of time preference.

[5]The question has been studied in relational contracting by Opp and Zhu (2015), in dynamic moral hazard by DeMarzo and Sannikov (2006) and Biais et al. (2007), and in the public finance literature with risk averse consumers by Farhi and Werning (2007) and Acemoglu et al. (2008). See section 7 for further details.

[6]İmrohoroğlu and Tüzel (2014) find the average persistence in total factor productivity of firms in Compustat data from 1962 to 2009 to be 0.7. Gomes (2001) estimates firm productivity in Compustat data from 1979 to 1998 through an AR(1) process and pegs the autocorrelation coefficient to be at 0.62.



(a) Correlated types equal discounting (b) iid types unequal discounting (c) Correlated types unequal discounting

Figure 1: Sample of allocations across time

as current incentive and individual rationality constraints would permit. On the other hand, unequal discounting ensures that the supplier wants to frontload payoffs of the small firm to arbitrage from the difference in interest rates to the extent future incentive and individual rationality constraints would allow. These two forces work in opposite directions leading to a cyclical pattern in optimal distortions. The efficient amount of capital is supplied for the high TFP shock; however, the low type is distorted and extent of this distortion, viz. its distance from efficiency, is governed by this cyclical property we call *restart*.

Dynamic distortions under the restart property are a function of the number of consecutive low shocks, once a high shock arrives the process repeats again. Figure 1 plots a sample of optimal allocations where the two horizontal lines depict the efficient levels for the high and low TFP shocks, respectively. In each case the first period type is high. With persistent (or Markovian) types and equal discounting the allocation is exclusively efficient. With independent types and unequal discounting, the distortions persist but they do not have any memory. Finally, with persistent types and unequal discounting, distortions have infinite memory along consecutive low shocks, but these are revised every time a high shock arrives.[7]

As can be inferred from Figure 1c, for consecutive low shocks the optimal allocation first falls and then rises to converge to a fixed value. In the figure this convergent value clearly lies above zero. However, if the agency problem is acute, the distortions do not decrease enough for the allocation to converge to a positive number. In such a situation the optimal contract shuts down for the low TFP shock, it gets zero supply across time. Both restarts and shutdowns capture

[7]Note that in all three cases in Figure 1 the optimal contract is restart, but in the first two it is trivially so since distortions along the low sequence of shocks have no memory.

certain salient features of real world contracts.[8] Both features are absent in the equal discounting model.

The nature of dynamic distortions poses a question to the literature on dynamic (Myersonian) mechanism design- a slight perturbation of the standard model of equal discounting renders long-term efficiency unachievable, distortions are pervasive. With equal discounting, Besanko (1985) and Battaglini (2005) show that ex post distortions converge to zero in the long run for the AR(1) and two type Markov models respectively. Garrett et al. (2018) show that distortions converge to zero on average for more general types' processes.[9] Our results make clear that these predictions will not hold for unequal discounting. In the language of financial economics, the Modigliani-Miller theorem does not hold even asymptotically; capital structure is perennially relevant and long-run value of economic surplus follows a non-trivial invariant distribution.[10]

Does the principal prefer a patient or impatient agent? Using ex ante profit as the objective, we show that the answer to this question depends on the extent of interim agency friction as measured by the persistence in the agent's private information. For limited agency problems (when private information is almost independent), the principal prefers a patient agent. However, for large levels of agency frictions (when private information is highly correlated), the principal actually prefers the agent to be myopic. The principal incurs two costs: dynamic information rent and intertemporal cost of incentive provision. For limited agency friction the first component is small, and the latter is a decreasing function of the difference in discounting- so a patient agent decreases the overall cost of incentive provision. However, when agency friction is large, the first component dominates

[8]Restart contracts exposit a natural environment in which an endogenous resetting of the terms is optimal. Debt contracts such as for home loans or insurance contracts often feature such properties (see Fuster and Willen (2017)). Some supply manufacturing contracts allow for revisiting terms as part of the ex ante agreement (see Lyon (1996)). Shutdown exhibits a situation where the big party to the contract commits to not supplying capital to an agent with inferior technology. It represents an endogenous decision of dissolution of a small firm or a business model of sorting by the principal in which she finds it profitable only to contract with a high quality client. There is a fairly large literature on the dynamics of firm growth and survival (see Evans (1987) and Clementi and Hopenhayn (2006)). There is also a rich discussion of screening along the quality dimension in industrial organization (see Tirole (1988)).

[9]See also Bergemann and Strack (2015) for the evolution of dynamic distortions in the continuous time setting.

[10]This is in contrast to Krasikov and Lamba (2017) who show that with hard financial constraints modeled through the limited liability restriction and equal discounting, efficiency is achieved almost surely in the long-run.

the second, and therefore having an impatient agent aids in reducing the overall cost of incentives, even though it increases the second intertemporal part.

Finally we tackle what we regard as an important challenge for dynamic contracting- binding global incentive constraints (see Battaglini and Lamba (2017) and Sannikov (2014)). Unequal discounting leads to the downward and upward incentive constraints binding simultaneously for certain parameters. The optimal contract then looses the restart feature and can have a very complicated form. We do two things. First we completely characterize the optimal recursive contract and exposit the basic intuition through simple pictures. Second we look for the optimal restart contract, that is we restrict our search to a subclass of incentive compatible contracts that have the restart property. When the first-order approach is valid, it coincides with the optimum, and when global incentives bind it provides an approximately optimal alternative that is incentive compatible and relatively easy to characterize. Our theoretical bounds on the performance of optimal restart contracts depend only on the fundamentals, and show a moderate loss of the ex ante objective.

The technical arguments we develop to provide theoretical bounds could have a more general appeal in solving such models. In a nutshell, the value of the objective under the first-order approach, say A, is always (weakly) higher than the value of the global optimum, say B, since the latter is calculated under a strict superset of constraints. The former ignores all the "upward" incentive constraints. The main problem is that when the first-order approach fails, B is endogenous to the set of binding constraints, and generally hard to calculate. Therefore, we restrict attention to restart contracts and calculate the optimal value of the objective, say C. When the first-order approach is valid, $A = B = C$, and when it is not, $A > B > C$, so we can evaluate $A - C$ (or $\frac{A}{C}$) which forms an upper bound on the gap we are interested in, viz. $B - C$ (or $\frac{B}{C}$). This gap $A - C$ is generated by sensitivity analysis: a method of approximating the amount of slack that needs to be added to the "upward" incentive constraints so that the value of the objective in the new auxiliary problem coincides with that in the first-order optimum.

2 Model

2.1 Primitives

A firm (agent) with access to a production technology approaches a supplier (principal) of a key input; the former is a "small player" while the latter is a "big player" in the market.[11] The total factor productivity (TFP) of the firm is its private information. They agree to sign a (dynamic) contract whereby endogenous levels of input are supplied by the principal every period, in return for monetary payments by the agent. Formally, the agent's stage (or per-period) preferences are given by $\theta R(k) - p$ where k is the input supplied by the principal, p is the payment made by the agent, θ is the total factor productivity, and $R(\cdot)$ is a concave production function that satisfies Inada conditions.[12] TFP or technology "shocks" can take values in $\Theta = \{\theta_H, \theta_L\}$, where $\theta_H, \theta_L > 0$ and $\theta_H - \theta_L = \Delta\theta > 0$. We will often refer to it as the agent's type. The first period type is drawn from a prior $\mu = \{\mu_H, \mu_L\}$, and then evolves according to a Markov process: $f(\theta_H|\theta_i) = \alpha_i$, $f(\theta_L|\theta_i) = 1 - \alpha_i$, for $i = H, L$, which satisfies first-order stochastic dominance: $\alpha_H \geq \alpha_L$. The principal does not observe the output, and therein lies the asymmetric information or agency friction. Her stage preference is simply $p - k$.

The contract lasts for T discrete periods, where for the most part we will consider $T = 2$ and $T = \infty$. Both principal and agent discount future utility, but importantly we *do not restrict them to have the same discount factor*; these are denoted by δ_P and δ_A respectively where $\delta_P \geq \delta_A$. The principal can commit to a long-term contract. The set of all parameters of the model is given by $\Gamma = \{R(\cdot), \Theta, \mu, f, \delta_P, \delta_A\}$.

Invoking the revelation principle, a direct mechanism is denoted by $m = \langle \mathbf{k}, \mathbf{p} \rangle = \left(k(\hat{\theta}_t|h^{t-1}),\, p(\hat{\theta}_t|h^{t-1}) \right)_{t=1}^{T}$, where h^{t-1} and $\hat{\theta}_t$ are, respectively, the history of reports up to $t-1$ and current report at time t.[13] The reported history h^t is recursively defined as $h^t = (h^{t-1}, \hat{\theta}_t)$ starting with $h^0 = \emptyset$. The set of all history paths is denoted by H^{t-1}, with $H^0 = \emptyset$. In what follows θ_i^{t-1} stands for the history

[11] Throughout the agent will be referred to as a he and the principal as a she.

[12] Technically: (i) $R'(k) > 0$, $R''(k) < 0$ for all $k \geq 0$, (ii) $R(0) = 0$ and (iii) $\lim_{k \to 0} R'(k) = \infty$, $\lim_{k \to \infty} R'(k) = 0$.

[13] At the cost of minimal confusion, the subscript will be used interchangeably for time and H/L. Also, as is standard, a contract is restricted to lie in l^∞.

with $t-1$ consecutive reports of type θ_i. The principal's objective is to maximize her profit subject to incentive compatibility and participation constraints for the agent. The private history of the agent is given by $h_A^t = (h_A^{t-1}, \hat{\theta}_t, \theta_{t+1})$, starting from $h_A^0 = \theta_1$, where $\hat{\theta}_t$ and θ_t are the reported and actual types, respectively. For a fixed mechanism, the agent faces a dynamic decision problem in which her strategy, $(\sigma_t)_{t=1}^T$, is simply a function that maps his private history into an announcement every period: $h_A^t \mapsto \sigma_t(h_A^t) \in \Theta$.[14]

Finally for any t, partition the set of histories till that time H^t into $\{H_R^t, \theta_L^t\}$, where θ_L^t is the "lowest history" of t consecutive realizations of type θ_L, and H_R^t is the set of all histories where type θ_H is realized at least once. For reasons that will be clear later, we refer to H_R^t as the "restart phase".

2.2 Constraints

Define the stage and expected utility of the agent (under truthful reporting) at any history of the contract tree to be

$$u(\theta_t|h^{t-1}) = \theta_t R(k(\theta_t|h^{t-1})) - p(\theta_t|h^{t-1}),$$

$$U(\theta_t|h^{t-1}) = u(\theta_t|h^{t-1}) + \delta_A \mathbb{E}\left[U(\tilde{\theta}_{t+1}|h^{t-1}, \theta_t)|\theta_t\right],$$

It is straightforward to note that a contract can then be expressed as $\langle \mathbf{k}, \mathbf{u} \rangle$ or $\langle \mathbf{k}, \mathbf{U} \rangle$. We shall use the three formulations interchangeably.

A contract is said to be *incentive compatible* if truthful reporting by the agent is always profitable for him. Using the one shot deviation principle, formally, for $i = H, L$ and $\forall h^{t-1}$, $\forall t$

$$IC_i(h^{t-1}) : U(\theta_i|h^{t-1}) \geq \theta_i R(k(\theta_j|h^{t-1})) - p(\theta_j|h^{t-1}) + \delta_A \mathbb{E}\left[U(\tilde{\theta}_{t+1}|h^{t-1}, \theta_j) \mid \theta_i\right],$$
$$\text{with } j \neq i.$$

A contract is said to be *almost incentive compatible* if $IC_i(h^{t-1})$ is required to hold for $i = H, L$ and $\forall h^{t-1} \neq \theta_L^{t-1}$. The difference is that we ignore the agent's incen-

[14] Note that other dynamic screening models can mapped into our framework and all the results in the paper can be analogously stated. For example, we can also consider the regulation model à la Laffont and Tirole (1993) where the principal and agent have preferences $V(k) - p$ and $p - \theta k$ respectively, or the monopolistic screening model à la Mussa and Rosen (1978) where the principal and agent have preferences $p - k^2/2$ and $\theta k - p$, respectively.

tives along the lowest history. $IC_H(h^{t-1})$ will be referred to as the "downward" incentive constraint, and $IC_L(h^{t-1})$ as the "upward" incentive constraint.

A contract is said to be *individually rational* if it offers each type of the agent a non-negative expected utility after every history, formally, for $i = H, L$ and $\forall h^{t-1}$, $\forall t$

$$IR_i(h^{t-1}) : U(\theta_i|h^{t-1}) \geq 0.$$

Individual rationality ensures that the agent is provided with a minimum expected utility at each stage; its normalization to zero is done for simplicity.

2.3 Optimization problem

Define $s(k, \theta) = \theta R(k) - k$ to be the static surplus, written as $s(\theta_t|h^{t-1}) = \theta_t R(k(\theta_t|h^{t-1})) - k(\theta_t|h^{t-1})$ for the direct mechanism. The *efficient input* supply that maximizes the surplus is independent of history and is given by $\theta R'(k^e(\theta)) = 1$. Let $\bar{S} = \sum_{t=1}^{T} \delta_P^{t-1} \mathbb{E}\left[s(\tilde{\theta}_t|\tilde{h}^{t-1})\right]$ be the (ex ante) expected surplus generated by a given contract. Moreover, define

$$\bar{U}_P = \sum_{t=1}^{T} \delta_P^{t-1} \mathbb{E}\left[u(\tilde{\theta}_t|\tilde{h}^{t-1})\right] \quad \text{and} \quad \bar{U}_A = \sum_{t=1}^{T} \delta_A^{t-1} \mathbb{E}\left[u(\tilde{\theta}_t|\tilde{h}^{t-1})\right].$$

For $\delta_P = \delta_A$, we have $\bar{U}_P = \bar{U}_A$. However, in our framework, the principal and agent evaluate the net present value of agent's utility stream differently. This core departure from the standard model will generate novel dynamic tradeoffs. The principal's problem, say $(\star)$, can be stated as

$$(\star) \qquad \max_{m} \quad \bar{S} - \bar{U}_P,$$

subject to $\mathbf{k} \geq 0$, and

$$IC_H(h^{t-1}), IC_L(h^{t-1}), IR_H(h^{t-1}), IR_L(h^{t-1}),\ \forall\ h^{t-1} \in H^{t-1},\ \forall t,$$

where $IC_i(h^{t-1})$ and $IR_i(h^{t-1})$ are the incentive compatibility and individual rationality constraints, respectively, for type θ_i in period t after history h^{t-1}. The first step is to identify the subset of constraints that bind at the optimum. These are then used to substitute $\bar{U}_P$, and express the objective only in terms of

$\mathbf{k}$. Pointwise optimization of allocations along all histories then yields the optimal contract.

3 Sequential first-order approach

3.1 Two period problem

We start with $T = 2$ and invoke the so-called first-order approach, wherein we maximize the objective subject to the "downward" incentive constraints and the individual rationality constraint of the "low" type. It is easy to show that all the incentive and individual rationality constraints in the relaxed problem can be assumed to hold as equalities.

$$\max_{m} \quad \bar{S} - \bar{U}_P,$$

subject to $\mathbf{k} \geq 0$,

$$IC_H(h) \text{ and } IR_L(h), \quad \text{for } h = \emptyset, H, L.$$

The economic force here, different than in the standard model, is that for the same sequence of stage utilities, the agent and the principal evaluate expected utility differently. Thus, in order to employ the Myersonian pointwise maximization of virtual surplus (that is, surplus minus information rents), evaluation of $\bar{U}_A$ will not do. Instead, we need to calculate the vector of stage payoffs $\mathbf{u}$ and then aggregate them to $\bar{U}_P$ using the principal's discount factor.

The second period incentive and individual rationality constraints give

$$u(\theta_H|\theta_i) = \Delta\theta R(k(\theta_L|\theta_i)) \quad \text{and} \quad u(\theta_L|\theta_i) = 0, \quad \text{for } i = H, L.$$

Through binding IC_{HL} and IR_L constraints, we get

$$U(\theta_H) = \Delta\theta R(k(\theta_L)) + \delta_A(\alpha_H - \alpha_L)\Delta\theta R(k(\theta_L|\theta_L)) \quad \text{and} \quad U(\theta_L) = 0.$$

Let $\mathbb{P}(h)$ be the ex ante probability of history h. Parsing out the two types of costs incumbent on the principal, we have $\bar{U}_P = \bar{U}_A + I$, where

$$\begin{aligned}\bar{U}_A = \mu_H U(\theta_H) + \mu_L U(\theta_L) &= \mu_H \Big[\Delta\theta R(k(\theta_L)) + \delta_A(\alpha_H - \alpha_L)\Delta\theta R(k(\theta_L|\theta_L))\Big] \\ &= \frac{\mu_H}{\mu_L}\Delta\theta R(k(\theta_L))\mathbb{P}(\theta_L) + \delta_P \frac{\mu_H}{\mu_L}\Bigg(\underbrace{\frac{\delta_A}{\delta_P}\frac{\alpha_H - \alpha_L}{1-\alpha_L}}_{=:b}\Bigg)\Delta\theta R(k(\theta_L|\theta_L))\mathbb{P}(\theta_L^2), \qquad (1)\end{aligned}$$

$$\begin{aligned}I &= (\delta_P - \delta_A)\sum_{i=H,L}\mu_i\Big[(\alpha_i u(\theta_H|\theta_i) + (1-\alpha_i)u(\theta_L|\theta_i))\Big] \\ &= \delta_P\Bigg[\Bigg(\underbrace{\frac{\delta_P - \delta_A}{\delta_P}\frac{\alpha_H}{1-\alpha_H}}_{=:a_H}\Bigg)\Delta\theta R(k(\theta_L|\theta_H))\mathbb{P}(\theta_H\theta_L) \\ &\quad + \Bigg(\underbrace{\frac{\delta_P - \delta_A}{\delta_P}\frac{\alpha_L}{1-\alpha_L}}_{=:a_L}\Bigg)\Delta\theta R(k(\theta_L|\theta_L))\mathbb{P}(\theta_L^2)\Bigg]. \qquad (2)\end{aligned}$$

Here, $\bar{U}_A$ is the standard (dynamic) information rent that the principal has to provide the agent, and I is the additional intertemporal cost of incentive provision. Since the amount of surplus that principal has to part with is expressible in terms of quantities, we can now calculate the first-order optimal contract. Define $\mathcal{K}_L(x) = (R')^{-1}\left(\frac{1}{\theta_L - x\Delta\theta}\right)$ for $x\Delta\theta < \theta_L$ and zero otherwise.

Proposition 1. The following supply contract characterizes the solution to the relaxed problem:

$$\begin{aligned}&k^{\#}(\theta_H|h) = k^e(\theta_H), \\ &k^{\#}(\theta_L|h) = \mathcal{K}_L(\rho(\theta_L|h)), \qquad \text{for } h = \emptyset, \theta_H, \theta_L,\end{aligned}$$

where $\rho(\theta_L) = \frac{\mu_H}{\mu_L}$, $\rho(\theta_L|\theta_L) = \rho(\theta_L)b + a_L$ and $\rho(\theta_L|\theta_H) = a_H$.

This result precisely pins down dynamic distortions in the two period screening contract with unequal discounting. The high type is always supplied the efficient allocation, the supply to the low one is distorted downwards. Distortions are pervasive in that $k(\theta_L|h) < k^e(\theta_L)$ for all h. To grasp the intuition, consider the following chain of arguments. Assume that rent of the type θ_H after history $h = \emptyset$ is increased by $\Delta\theta\varepsilon$.[15] The expected utility of the agent goes up by $\mathbb{P}(\theta_H)\Delta\theta\varepsilon$,

[15] This is done by increasing the first period allocation of θ_L by an amount x such that

which is the principal's cost for providing the agent with the requisite incentives. Concomitantly, the expected surplus changes by $\mathbb{P}(\theta_L)\Delta S(\varepsilon)$ where $\Delta S(\varepsilon)$ is the associated change in expected surplus. Thus, the net change in marginal cost-marginal benefit ratio is proportional to $\rho(\theta_L) = \frac{\mathbb{P}(\theta_H)}{\mathbb{P}(\theta_L)} = \frac{\mu_H}{\mu_L}$.

Next, assume that a rent of type θ_H after history $h = \theta_H$ is increased by $\Delta\theta\varepsilon$. This increase costs the principal $\delta_P \mathbb{P}(\theta_H^2)\Delta\theta\varepsilon$. Moreover, the ex ante expected utility of the agent increases by $\delta_A \mathbb{P}(\theta_H^2)\Delta\theta\varepsilon$, all of which can then be extracted by the principal. Therefore, the aggregate cost to the principal of this change is given by $(\delta_P - \delta_A)\mathbb{P}(\theta_H^2)\Delta\theta\varepsilon$. As before, the benefit of this change is given by increase in surplus generated by increasing the allocation to type θ_L (after history $h = \theta_H$), which in a slight abuse of notation can be given by $\delta_P \mathbb{P}(\theta_H\theta_L)\Delta S(\varepsilon)$. The net change in marginal cost-marginal benefit ratio is therefore proportional to $\rho(\theta_L|\theta_H) = a_H$.[16]

Finally, assume that the rent of type θ_H after history $h = \theta_L$ increases by $\Delta\theta\varepsilon$. As before, aggregate incentive cost to the principal equals $(\delta_P - \delta_A)\mathbb{P}(\theta_L\theta_H)\Delta\theta\varepsilon$. The change also leads to an increase in the ex ante utility of the low type agent by $\delta_A \mathbb{P}(\theta_L\theta_H)\Delta\theta\varepsilon$ all of which can be extracted upfront by the principal through the binding IR_L constraint. However, in order to maintain the IC_H constraint, she also needs to provide the high type agent with an additional utility worth $\delta_A \mathbb{P}(\theta_H^2)\Delta\theta\varepsilon$. Therefore, the total cost to the principal of this change is given by $\Big[(\delta_P - \delta_A)\mathbb{P}(\theta_L\theta_H) + \delta_A(\mathbb{P}(\theta_H^2) - \mathbb{P}(\theta_L\theta_H))\Big]\Delta\theta\varepsilon$. The benefit of this change is of course $\delta_P \mathbb{P}(\theta_L^2)\Delta S(\varepsilon)$. The expected net change in marginal cost-marginal benefit ratio is proportional to $\rho(\theta_L|\theta_L) = \rho(\theta_L)b + a_L$.

Coefficients b and a in equations (1) and (2) represent the distortions with respect to the intertemporal cost of incentive provision and standard information rent; the former is purely a manifestation of differential discounting. In addition, transfers are uniquely pinned down. This is in striking contrast to the standard quasilinear model of dynamic screening with equal discounting where aggregate utility (that is $U(\theta_H)$ and $U(\theta_L)$) is uniquely pinned down up to a constant, but a continuum of transfers implement the optimum.

Remark 1. Given $\mathbf{k}^{\#}$, the vector of optimal utilities $\mathbf{U}^{\#}$ with a cardinality of six, is uniquely pinned down by the set of six binding constraints.

$R(k(\theta_L) + x) - R(k(\theta_L)) = \varepsilon$.

[16] For context, note that $\mathbb{P}(\theta_H^2) = \mu_H\alpha_H, \mathbb{P}(\theta_H\theta_L) = \mu_H(1-\alpha_H), \mathbb{P}(\theta_L\theta_H) = \mu_L\alpha_L, \mathbb{P}(\theta_L^2) = \mu_L(1-\alpha_L)$.

The intuition for this is fairly straightforward. Even with an arbitrarily small difference in discounting, the principal wants to lend an infinite amount of money in the first period, only to demand it back in the second. He is however restricted in this "arbitrage" by the agent's individual rationality constraint. Therefore, irrespective of the history, the agent's individual rationality, and hence incentive compatibility constraints bind, leading to a system of six equalities. All six prices, which enter linearly in this optimization problem, are thus uniquely determined.

We also note that the first-order approach may not always be valid, that is $(\mathbf{k}^{\#}, \mathbf{U}^{\#})$ may violate the first period "upward" incentive constraint IC_L. In the static model $k^{\#}(\theta_H) \geq k^{\#}(\theta_L)$ is a necessary and sufficient condition for the validity of the first-order approach, and this condition is always satisfied. In the dynamic model, for the first-order optimal contract to satisfy IC_L, a weighted average of allocation that follow the "high" type $(k^{\#}(\theta_H), k^{\#}(\theta_H|\theta_H), k^{\#}(\theta_L|\theta_H))$ must be greater than the corresponding weighted average of the allocations that follow the "low" type $(k^{\#}(\theta_L), k^{\#}(\theta_H|\theta_L), k^{\#}(\theta_L|\theta_L))$, where the weights are determined by the Markov matrix.[17] With equal discounting this three dimensional vector is pointwise greater for the "high" history. However, with unequal discounting, if a_H is very large, that is $k^{\#}(\theta_L|\theta_H)$ is highly distorted and significantly less than $k^{\#}(\theta_L|\theta_L)$, then the desired average notion of monotonicity fails culminating in a failure of the first-order approach. Parametrically speaking, IC_L binds for low levels of ex ante agency friction and high levels of interim agency friction, that is smaller values of $k^e(\theta_H) - k^e(\theta_L)$ and larger values of α_H respectively.

To end the description of the two period model, we provide a simple sufficient condition for the validity of the fist-order approach. Although there are much weaker sufficient conditions, Corollary 1 provides one that is easy to state.

Corollary 1. Suppose $R(k^e(\theta_H)) \geq 2R(k^e(\theta_L))$. Then, the first-order optimal contract solves $(\star)$.

3.2 Infinite horizon problem

We extend the relaxed problem (or first-order) approach adopted in the two period model to the infinite number of periods- here all "upward" incentive constraints are ignored. In the appendix, we show that for all h^{t-1}, $IC_H(h^{t-1})$ and

[17] See for example Corollary 1 in Pavan et al. (2014).

$IR_L(h^{t-1})$ bind at the optimum, and $IR_H(h^{t-1})$ is trivially satisfied. Using the binding $IR_L(h^{t-1})$ constraints, we have

$$u(\theta_H|h^{t-1}) = U(\theta_H|h^{t-1}) - \delta_A \alpha_H U(\theta_H|h^{t-1}, \theta_H) \quad \text{and}$$
$$u(\theta_L|h^{t-1}) = -\delta_A \alpha_L U(\theta_H|h^{t-1}, \theta_L). \tag{3}$$

In addition, the following identity is generated by the inductive application of binding $IC_H(h^{t-1})$ and $IR_L(h^{t-1})$ constraints:

$$U(\theta_H|h^{t-1}) = \sum_{s=0}^{\infty} (\delta_A(\alpha_H - \alpha_L))^s \Delta\theta R(k(\theta_L|h^{t-1}, \theta_L^s)) \tag{4}$$

Equations (3) and (4) give the expression for $\bar{U}_P$ in terms of the allocation. As before, we can parse it out into two components: $\bar{U}_P = \bar{U}_A + I$, such that

$$\bar{U}_A = \mu_H U(\theta_H) + \mu_L U(\theta_L) = \sum_{t=1}^{\infty} \delta_P^{t-1} \cdot \frac{\mu_H}{\mu_L} b^{t-1} \cdot \Delta\theta R(k(\theta_L|\theta_L^{t-1}))\mathbb{P}(\theta_L^t), \text{ and} \tag{5}$$

$$\begin{aligned} I &= \frac{\delta_P - \delta_A}{\delta_P} \sum_{t=2}^{\infty} \delta_P^{t-1} \mathbb{E}\left[U(\tilde{\theta}_t|\tilde{h}^{t-1})\right] \\ &= \frac{\delta_P - \delta_A}{\delta_P} \sum_{t=2}^{\infty} \delta_P^{t-1} \cdot \left(\hat{\rho}_t - \frac{\mu_H}{\mu_L} b^{t-1}\right) \cdot \Delta\theta R(k(\theta_L|\theta_L^{t-1}))\mathbb{P}(\theta_L^t) \\ &+ \frac{\delta_P - \delta_A}{\delta_P} \sum_{h^{t-1}} \sum_{s=0}^{\infty} \delta_P^{t+s} \cdot \rho_t \cdot \Delta\theta R(k(\theta_L|h^{t-1}, \theta_H, \theta_L^s))\mathbb{P}(h^{t-1}, \theta_H, \theta_L^{s+1}), \end{aligned} \tag{6}$$

where $\hat{\rho}_t$ and ρ_t are functions of $(\alpha_H, \alpha_L, \delta_P, \delta_A, \mu)$. We are now ready to provide the closed form expression for the first-order optimal contract.

Proposition 2.

$$k^{\#}(\theta_H|h^{t-1}) = k^e(\theta_H), \quad \forall h^{t-1},$$
$$k^{\#}(\theta_L|h^{t-1}) = \begin{cases} \mathcal{K}_L(\hat{\rho}_t), & \text{if } h^{t-1} = \theta_L^{t-1}, \\ \mathcal{K}_L(\rho_s), & \text{if } h^{t-1} = (h^{\tau-1}, \theta_H, \theta_L^{s-1}), \text{ s.t. } \tau + s = t, \end{cases}$$

where $\hat{\rho}_t = b\hat{\rho}_{t-1} + a_L$, $\hat{\rho}_1 = \frac{\mu_H}{\mu_L}$ and $\rho_{t+1} = b\rho_t + a_L$, $\rho_1 = a_H$.

Persistence in private information leads to the *propagation* of distortions. Each consecutive low shock produces a sequence of distortions that infinitely propagates along the lowest history from that point on. Thus after any history of types, along a sequence of low shocks new distortions are recursively added at each point. Perhaps surprisingly, their aggregate effect can be exactly pinned down. Proposition 2 points to two immediately observable properties: first the high type is always provided the efficient allocation, and second, the distortion for the low type is a function of the number of consecutive low shocks. These can be formalized through the following definition.

Definition 1. A contract m is **restart** if for all t and h^{t-1}

$$k(\theta_H|h^{t-1}) = k(\theta_H) \quad \text{and} \quad k(\theta_L|h^{t-1}, \theta_H, \theta_L^{s-1}) = k(\theta_L|\theta_H, \theta_L^{s-1}),\ \forall s.$$

Note that allocation in a restart contract can be succinctly expressed by two sequences, one that represents the optimal allocation along the lowest history, and the other that represents it in the restart phase.

Remark 2. Suppose m is a restart contract. Then, $\exists$ two sequences $\{\hat{k}_t\}$ and $\{k_t\}$ such that for all t and h^{t-1}, $k(\theta_L|\theta_L^{t-1}) = \hat{k}_t$ and $k(\theta_L|h^{t-1}, \theta_H, \theta_L^{s-1}) = k_s$.

From Proposition 2 we can conclude that the first-order optimal contract satisfies the restart property with $\hat{k}_t = \mathcal{K}_L(\hat{\rho}_t)$ and $k_t = \mathcal{K}_L(\rho_t)$, where $\hat{\rho}_t$ documents the distortions along the lowest history, and ρ_t documents those in the restart phase. Figure 2 explains the dynamics. The contract starts in the white circle. The first period type draw initializes the contract leading it to one of two gray circles, labeled θ_H and θ_L. From then on, the contract transits among the grey circles depending on whether a high or low type is realized. The allocation (and expected utility) supplied to the agent is printed on each gray circle.

We also show that the distortions in the restart phase are monotonically decreasing, implying $k_{t+1} \geq k_t$ with a strict inequality for $k_t > 0$. If $k_1 = 0$ and there exists a τ such that $k_\tau > 0$, then the contract features temporary shutdwon. It is also possible that $\lim_{t\to\infty} k_t = 0$, then we say that the contract is permanently shutdown for the low type. More generally, we can define shutdown as follows.

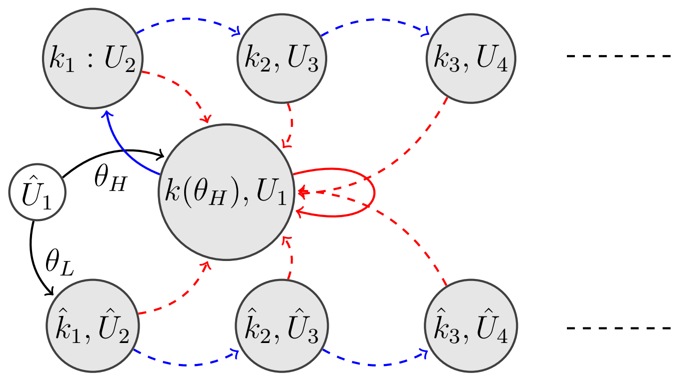


Figure 2: The evolution of allocation and expected utility in a restart contract. A red/blue arrow indicates a transition, because of a high/low draw. A solid/dashed arrow corresponds to the probability of transition $\alpha_j/1-\alpha_j$ where $j = H/L$ if the arrow is solid/dashed.

Definition 2. A contract m is (permanently) **shutdown** if for all t and all $h^{t-1} \neq \theta_L^{t-1}$, $k(\theta_L|h^{t-1}) = 0$. Shutdowns are temporary if $k(\theta_L|h^{t-1}) = 0$ only for some $h^{t-1} \neq \theta_L^{t-1}$.[18]

The following list consolidates the key properties exhibited by the dynamic distortions of the first-order optimal contract.

Corollary 2. The first-order optimal contract satisfies the following properties:

(a) Distortions are monotonically decreasing in the restart phase: $\rho_t > \rho_{t+1}$ $\forall t$.

(b) Distortions are monotone along the lowest history: $\hat{\rho}_t \gtrless \hat{\rho}_{t+1}$ $\forall t$ whenever $\frac{\mu_H}{\mu_L} \gtrless \frac{a_L}{1-b}$.

(c) Distortions are pervasive: $\lim\limits_{t\to\infty} \hat{\rho}_t = \lim\limits_{t\to\infty} \rho_t = \frac{a_L}{1-b} > 0$.

(d) There are shutdowns in the restart phase: $\mathcal{K}_L(\rho_t) = 0$ for some t whenever $\theta_L \leq \rho_1 \Delta\theta$.

(e) Shutdowns are permanent: $\mathcal{K}_L(\rho_t) = 0$ for all t whenever $\theta_L \leq \lim\limits_{t\to\infty} \rho_t \Delta\theta$.

[18]Note that we can extend the defintion to include the lowest history as well. As we will see in Corollary 2, distortions along the lowest history are either monotonically increasing or decreasing. If the distortions converge to a value that keeps the allocation at zero, then the contract feature shutdown at the lowest history too. We ignore the lowest history here for simplicity of exposition.

What about transfers? As we explained in the two period model, the principal's desire to frontload agent's payoff as much as possible leads to all individual rationality and hence incentive compatibility constraints in the relaxed problem to bind. Therefore despite quasi-linearity the set of optimal expected utilities and transfers is unique. Along with allocations made every period, Figure 2 depicts the expected utility promised to the agent on the realization of a high type in the next period.

Remark 3. For all histories $U^{\#}(\theta_L|h^{t-1}) = 0$, and the expected utilities (and transfers) of the high type inherit the restart property- $U^{\#}(\theta_H|\theta_L^{t-1}) = \hat{U}_t$ and $U^{\#}(\theta_L|h^{t-1}, \theta_H, \theta_L^{s-1}) = U^{\#}(\theta_L|\theta_H, \theta_L^{s-1}) = U_s$ for two unique sequences of values $\{\hat{U}_t\}$ and $\{U_t\}$.

Finally, we register some simple results for specific parametric constellations that follow directly from Proposition 2. First, note that the first-order optimal contract is never efficient. Unequal discounting renews the potency of private information periodically so that even far into the future the distortions do not disappear. Second, for the iid model, the first-order approach is valid, and distortions are still pervasive though they do not have any memory. Third, for perfect persistence too the first-order approach is valid, the optimal contract has infinite memory and it converges to the efficient allocation in the long-run. Each of these produce the opposing conclusion for the equal discounting model.

Corollary 3. Optimal distortions in special cases of the Markov process are as follows.

(a) Correlated types ($1 > \alpha_H > \alpha_L$). For $\delta_P > \delta_A$, the first-order optimal contract is never efficient. For $\delta_P = \delta_A$ the first-order optimal contract is optimal and it converges to the efficient allocation along every history: $a_H = a_L = 0$, and $\rho_t = 0, \hat{\rho}_t = \frac{\mu_H}{\mu_L}\left(\frac{\alpha_H - \alpha_L}{1-\alpha_L}\right)^{t-1}$ $\forall\ t$.

(b) iid types ($\alpha_H = \alpha_L < 1$). The first-order optimal contract is optimal. For $\delta_P > \delta_A$, the optimal contract is never efficient but distortions have limited memory: $b = 0$, $\rho_t = \hat{\rho}_t = a_L\ \forall\ t \geq 2$. For $\delta_P = \delta_A$ the optimal contract is efficient starting period 2: $\rho_t = 0\ \forall\ t$, $\hat{\rho}_t = 0\ \forall\ t \geq 2$.

(c) Constant types ($\alpha_H = 1 - \alpha_L = 1$). The first-order optimal contract is optimal. For $\delta_P > \delta_A$, the optimal contract is efficient in the long run:

$\hat{\rho}_t = \frac{\mu_H}{\mu_L}\left(\frac{\delta_A}{\delta_P}\right)^{t-1} \forall\, t$. For $\delta_P = \delta_A$ an optimal contract is the repetition of the static optimum, it is never efficient: $\hat{\rho}_t = \hat{\rho}_1 \;\forall\, t$.

3.3 Connection to primitives

When is the first-order approach valid, and is it a necessary condition for the optimal contract to satisfy the restart property? The parametric space for which the "upward" incentive constraint binds can further by divided into two regions- one where it binds for finite time, and another where it binds perennially. It turns out, as is intuitive, that the optimal contract looses the restart property when the "upward" incentive constraint binds. So, corresponding to the two aforementioned parametric regions, the optimal contract is either eventually restart or never restart.

Definition 3. A contract m is **eventually restart** if there exists a $t^* < \infty$, a constant k_H and a sequence $\{k_t\}$ such that for all $t \geq t^*$ and h^{t-1},

$$k(\theta_H|h^{t-1}) = k_H \quad \text{and} \quad k(\theta_L|h^{t-1}, \theta_H, \theta_L^{s-1}) = k_s,\ \forall s.$$

In contrast, a contract that is not eventually restart is succinctly referred to as **never restart**.

It is easy to see that the first-order optimal contract is immediately restart, $t^* = 1$. Almost incentive compatibility, that is incentive compatible along all histories except potentially the lowest one, precisely characterizes eventually restart contracts.

Proposition 3. Suppose the first-order optimal contract is almost incentive compatible. Then, the optimal contract is eventually restart: there exists $t^* < \infty$ such that for all $t \geq t^*$ and h^{t-1},

$$k^*(\theta_H|h^{t-1}) = k^e(\theta_H) \quad \text{and} \quad k^*(\theta_L|h^{t-1}, \theta_H, \theta_L^{s-1}) = \mathcal{K}_L(\rho_s),\ \forall s,$$

where $\rho_t = b\rho_{t-1} + a_L$, $\rho_1 = a_H$. The converse is also true: if the first-order optimal contract is not almost incentive compatible, then the optimal contract is never restart.

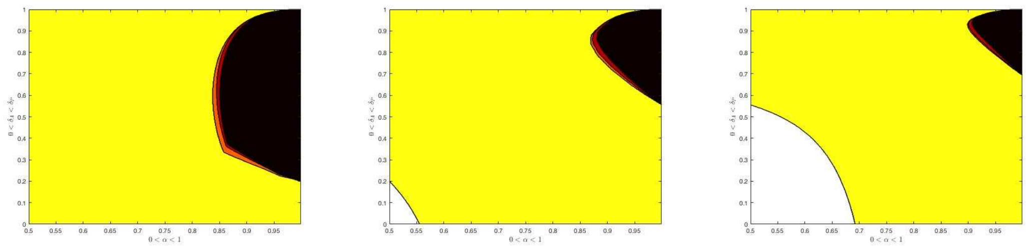
Therefore, if the first-order approach fails, it is either still valid eventually or it is not valid at all. Our next result identifies eventually restart contracts in terms of the primitives.

Corollary 4. Let $C = R(k^e(\theta_H)) + \delta_A(\alpha_H - \alpha_L)U^{\#}(\theta_H|\theta_H)$. The first-order optimal contract is optimal if and only if $\max\left\{U^{\#}(\theta_H), \lim_{t\to\infty} U^{\#}(\theta_H|\theta_L^{t-1})\right\} \leq C$. Moreover, the optimal contract is eventually restart if and only if $\lim_{t\to\infty} U^{\#}(\theta_H|\theta_L^{t-1}) \leq C$.

Since $\mathbf{U}^{\#}$ is uniquely pinned down, Corollary 4 presents a condition on the primitives of the environment. Recollect from Corollary 2(b) that distortions along the lowest history are either decreasing or increasing, therefore, the tightest upward incentive constraint is either the one in the first period or "the one at infinity", hence $\max\left\{U^{\#}(\theta_H), \lim_{t\to\infty} U^{\#}(\theta_H|\theta_L^{t-1})\right\} \leq C$ ensures that the first-order optimal contract is incentive compatible along the lowest history. Next, distortions in the restart phase are monotonically decreasing along consecutive low cost realizations (Corollary 2(a)); moreover, distortions along the lowest history and in the restart phase converge to the same value (Corollary 2(c)). Putting these together we get that $\lim_{t\to\infty} U^{\#}(\theta_H|\theta_L^{t-1}) \leq C$ ensures almost incentive compatibility, that is incentive compatibility in the restart phase. Therefore, if the first-order optimal contract satisfies $\lim_{t\to\infty} U^{\#}(\theta_H|\theta_L^{t-1}) \leq C$, it is almost incentive compatible and hence eventually restart.

Figure 3 partitions the parameter space along the set of binding constraints for a specific example. White and yellow regions represent the validity of the first-order approach where the optimal contract is immediately restart, the dark region is the space where the optimal contract is never restart, and the region in between represents cases where the first-order approach is valid after finite time and the optimal contract is eventually restart. Moreover, the white portion in the southwest corner represents the case of (permanent) shutdown, no capital is supplied to the low type. For larger values of $\Delta\theta$, signifying greater ex ante agency friction, it is easier to separate the two types, and hence the first-order approach is more likely to be satisfied.

Discounting and persistence interact in a non-linear fashion. For $\delta_A = 0$ and δ_P, the first-order approach is valid, the same is true for the iid model ($\alpha_H = 1 - \alpha_L$) and perfect persistence ($\alpha_H = 1 - \alpha_L = 1$). More generally, high levels of discounting and persistence are required for the first-order approach to fail.



(a) $\Delta\theta = 0.25$ (b) $\Delta\theta = 1.25$ (c) $\Delta\theta = 2.25$

Figure 3: Partitioning parameter space into set of binding constraints. White & yellow: first-order approach works and optimal contract is restart. White: low type is shutdown. Black: upward constraint binds ad infinitum. $\alpha_H = 1 - \alpha_L = \alpha$ on the x-axis, δ_A on the y-axis. $\delta_P \approx 1$, $R(k) = 2\sqrt{k}$, $\theta_L = 1$.

4 Recursive approach: a full characterization

4.1 A restatement of the problem

In order to fully characterize the optimal contract, even when it is never restart, we turn to the recursive approach. It is well known that in order *recursify* a dynamic contracting sequence problem with an N-state Markov chain of types, the state variable of promised utility is required to be N-dimensional (Fernandes and Phelan (2000)). In our model, it is easy to show that $IC_L(h^{t-1})$ will always bind for the optimal contract, hence, $U^*(\theta_L|h^{t-1}) = 0 \;\forall\; h^{t-1}$. Thus, even though agent's type follows a two state Markov process, a one dimensional state variable, viz. $U(\theta_H|h^{t-1}) = w \in \mathbb{R}_+$, will suffice to encode all the required history dependence.

In the appendix, we show that the following recursive formulation is equivalent to the sequence problem described in $(\star)$. From the second period onwards, for a promised expected utility of w to the high type and last period type j, define the objective as follows:

$$(\mathcal{RP}) \qquad S_j(w) = \max_{(\mathbf{z},\mathbf{k})\in\mathbb{R}^4_+} \alpha_j\big[s(k_H,\theta_H) - (\delta_P - \delta_A)\alpha_H z_H + \delta_P S_H(z_H)\big] +$$
$$+ (1-\alpha_j)\big[s(k_L,\theta_L) - (\delta_P - \delta_A)\alpha_L z_L + \delta_P S_L(z_L)\big] \text{ s.t.}$$
$$w \geq \Delta\theta R(k_L) + \delta_A(\alpha_H - \alpha_L) z_L$$
$$w \leq \Delta\theta R(k_H) + \delta_A(\alpha_H - \alpha_L) z_H$$

The objective is to maximize the surplus when expected utility promised to the agent is fixed at $(w, 0)$ or $\alpha_j w + (1-\alpha_j)0$ in expectation. There are four choice variables: working capital advances $\mathbf{k} = (k_H, k_L)$ and expected utilities $\mathbf{z} = (z_H, z_L)$; note that z_i represents the utility promised to the high TFP type next period if the current type is θ_i. The term $(\delta_P - \delta_A)\alpha_i z_i$ captures the intertemporal cost of incentive provision incurred by the principal in providing a continuation value of z_i. The two constraints are the "downward" and "upward" incentive constraints, respectively. The participation constraint of θ_H type is subsumed in the recursive domain.

At date $t = 1$, the problem is different for two reasons: the belief equals the prior and contract has not yet been initialized. To initialize the contract, $w = U(\theta_H) - U(\theta_L)$ must be chosen. The problem reads as follows:

$$(\diamond) \qquad \Pi^* = \max_{(w,\mathbf{z},\mathbf{k})\in\mathbb{R}^5_+} -\mu_H w + \mu_H\big[s(k_H,\theta_H) - (\delta_P - \delta_A)\alpha_H z_H + \delta_P S_H(z_H)\big] +$$
$$+ \mu_L\big[s(k_L,\theta_L) - (\delta_P - \delta_A)\alpha_L z_L + \delta_P S_L(z_L)\big] \text{ s.t.}$$
$$w \geq \Delta\theta R(k_L) + \delta_A(\alpha_H - \alpha_L) z_L$$
$$w \leq \Delta\theta R(k_H) + \delta_A(\alpha_H - \alpha_L) z_H$$

We show that the value functions in $(\star)$ and $(\diamond)$ coincide, justifying our focus on the recursive problem. In what follows the recursive contract is informally characterized, formal details can be found in the appendix.

4.2 Optimal recursive contract

In this subsection, we exposit the properties of the optimal recursive contract, $\langle w^*, \mathbf{k}(\cdot), \mathbf{z}(\cdot)\rangle$, where $(\mathbf{k}(w), \mathbf{z}(w))$ solves $(\mathcal{RP})$ for each $w \geq 0$, and $(\diamond)$ is solved by $(w^*, \mathbf{k}(w^*), \mathbf{z}(w^*))$.[19] We start with registering the monotonicity of allocation with respect to expected utility given to the high type.

For the optimal recursive contract, allocations for the high and low TFP shocks are increasing in the state variable, w. Intuitively speaking, the downward incentive constraint binds only for low values of w. In this case, the allocation and promised expected utility upon announcing the low type (that is, k_L and $\alpha_L z_L$) must be distorted downwards to prevent the high type from misreporting. Indeed, there exists a critical value w_L^* so that the downward incentive constraint binds only for $w \leq w_L^*$. The incentive problem is more severe for low values of w, there exists another threshold w_k^o below which the contract does not supply θ_L.

By the similar reasoning, the allocation and promised expected utility upon announcing the high type (that is, k_H and $\alpha_H z_H$) must be distorted upwards if the upward incentive constraint binds. And, there exists a critical value w_H^* such that this constraint binds if and only if $w \geq w_H^*$. Figure 4a plots the optimal allocation as the function of agent's expected utility. We have the following simple result.

Proposition 4. Allocation in the optimal recursive contract satisfies the following:

(a) $\exists w_H^*$ such that $k_H(w) = k^e(\theta_H)$ if and only if $w \leq w_H^*$, $k_H(\cdot)$ is strictly increasing on $[w_H^*, \infty)$.

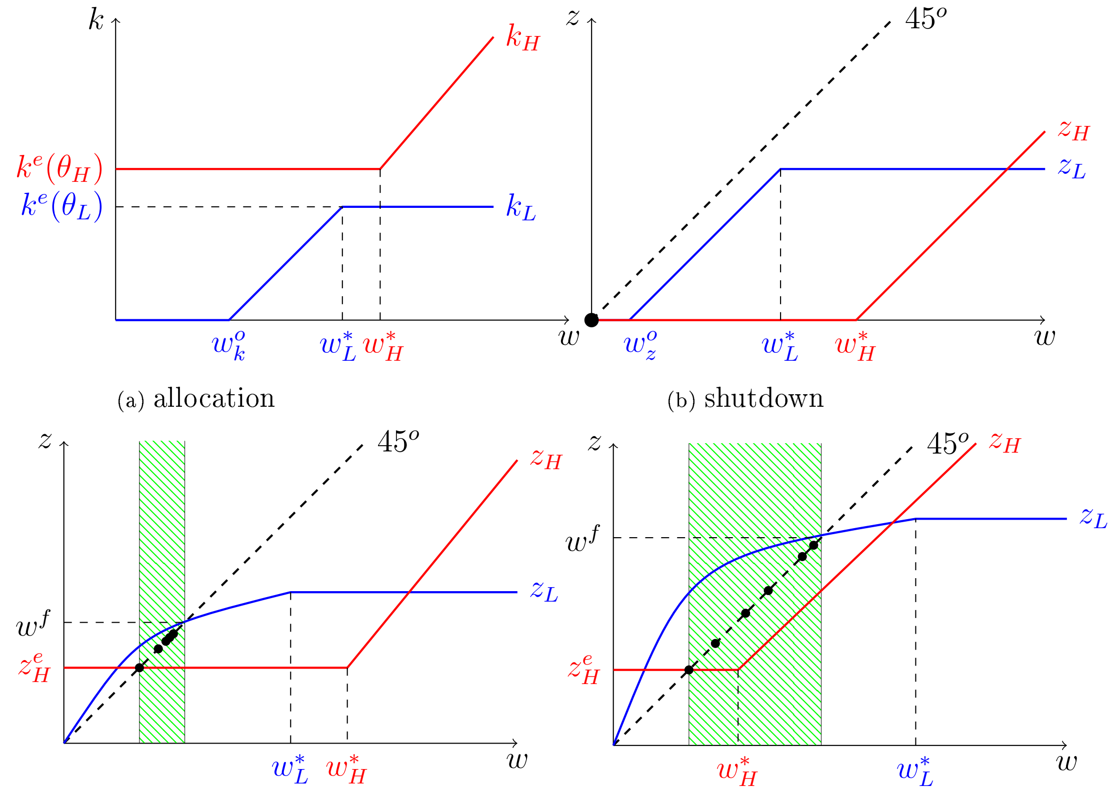
(b) $\exists w_k^o, w_L^*$ such that $k_L(w) = 0$ if and only if $w \leq w_k^o$, $k_L(w) = k^e(\theta_L)$ if and only if $w \geq w_L^*$, $k_L(\cdot)$ is strictly increasing on $[w_k^o, w_L^*]$.

The dynamics of promised expected utility are described in Figure 4. In each case z_H and z_L are plotted as functions of w. The 45° line partitions the quadrant into regions where expected utility increases or decreases in the next period. w_H^* and w_L^* are the thresholds as defined above. And the bold dots represent some points in the support of the invariant distribution of the optimal contract. [20]

[19] As in the sequential first-order optimal contract, the allocation and transfers are uniquely pinned down. To be precise, we formally show in the appendix that only z_H could fail to be unique at a single point. The details are provided in the appendix (Claim 4).

[20] The optimal contract induces a Markov process on the recursive domain. Formally, the Markov process is defined by $F_{i|j}^*(A|w) = \mathbb{1}(z_i(w) \in A) f(\theta_i|\theta_j)$ that is the probability that the expected utility promised to the agent in the next period lies in some Borel measurable set $A \subseteq \mathbb{R}_+$ when the type realized is θ_i, given that the current expected utility and last period's

For example, in all the figures the point z_H^e at which $z_H(\cdot)$ intersects the 45° line constitutes a bold dot. Each time a high shock arrives it is possible for the optimal contract to stay at the same expected utility, and it surely does so if the upward constraint is not binding.



(a) allocation

(b) shutdown

(c) eventually restart

(d) never restart

Figure 4: Optimal recursive contract

Consider first the situation depicted in Figure 4b. Here $z_H^e = 0$. Since both curves lie below the 45° line, the recursive contract continually shrinks in expected value. It quickly converges, most often immediately, to the bold point at zero which implies an expected utility of zero and a complete shutdown of the low TFP type. In Figures 4c and 4d, we portray the optimal restart contract which does not feature shutdowns. The realization of a high shock pushes the expected utility towards z_H^e. On the realization of a low shock, promised expected utility above w^f

shock are given by w and θ_j, respectively. By the standard mixing argument, the Markov process can be shown to have the unique invariant distribution, see Theorem 12.12 of Stokey et al. (1989).

contracts, and below w^f it expands. The key condition that characterizes Figure 4c is $w^f \leq w_H^*$. It implies that the upward incentive constraint does not bind in the interval $[z_H^e, w^f]$, and the invariant distribution of the promised expected utility rests therein.[21] In contrast, Figure 4d exposits the case with perennial binding of the "upward" incentive constraint which is captured by the condition $w^f > w_H^*$.

Finally, the only missing piece is initialization- where does the optimal recursive contract start? We show that the recursive contract is initialized at a unique point $w^* \in [0, w_L^*]$. Therefore, at the inception the downward incentive constraint always binds, while the upward constraint may or may not bind. The next proposition summarizes the evolution of expected utility in the optimal recursive contract.

Proposition 5. Expected utility of the agent in the optimal recursive contract satisfies the following:

(a) $\exists w_z^o, z_L^e$ such that $z_L(w) = 0$ if and only if $w \leq w_z^o$, $z_L(w) = z_L^e$ if and only if $w \geq w_L^*$, $z_L(\cdot)$ is strictly increasing on $[w_z^o, w_L^*]$.

(b) $\exists z_H^e$ such that $z_H(w) = z_H^e$ if and only if $w \leq w_H^*$, $z_H(\cdot)$ is strictly increasing on $[w_H^*, \infty)$.

(c) $z_L(\cdot)$ has a unique globally stable fixed point $w^f \in [z_H^e, z_L^*]$, and z_H has a unique fixed point z_H^e which is positive if and only if $\theta_L > \frac{a_L}{1-b}\Delta\theta$.

(d) The thresholds satisfy $z_H^e \leq w^f \leq z_L^e < w_L^*$, $z_H^e < w_H^*$, and $z_L^e \neq z_H^e$ if and only if $z_L^e > 0$.

(e) $\exists w^* \in [0, w_L^*]$ such that the optimal contract starts at this point, and it always stays within $[0, w_L^*]$.

Propositions 4 and 5 precisely characterize the optimal contract. Starting at w^*, each subsequent realization of the agent's type determines the optimal allocation according to Proposition 4 and the optimal expected utility for the next period, the state variable, according to Proposition 5.

[21] To find the support, we repeatedly apply $z_L(\cdot)$ to z_H^e, the bold points in Figure 4c depict this set.

There is of course a one-to-one relationship between the optimal recursive contract, and the sequential optimum. First of all, the “downward” incentive constraints always bind, and the low type always gets the promised utility of zero. The high type allocation can be distorted only upwards, whereas the low type allocation is always distorted downwards.

Moreover, the realization of each θ_H decreases the promised utility offered to the high type in the next period which reduces distortion for the high type allocation, but increases a distortion in the low type. It takes an endogenous number of consecutive θ_H for the “upward” incentive constraint to stop binding. After a finite number of periods, θ_L always increases the promised utility offered to the high type in the next period which tightens the distortion for the high type allocation, but relaxes distortions for the low type allocation. It takes an endogenous number of consecutive θ_L for the “upward” incentive constraint to start binding.

5 Optimal restart contract

When the upward incentive constraint binds forever, the optimal contract is never restart, and it is quite hard to exactly pin down in terms of the sequential formulation. In this section we construct an approximately optimal sequential contract by restricting our search to the class of all incentive compatible restart contracts. There are two reasons for this restriction: (i) it is a fairly intuitive criterion and simple to describe, and (ii) the first-order optimal contract falls within this class, and so if it is indeed globally optimal there is no loss. Our approach is similar in spirit to Chassang (2013) in that it emphasizes the search for approximately optimal contracts by constraining the instruments available to the principal, but it is also different in that we do still demand incentive compatibility.

The “downward” incentive constraint always bind for the optimal restart contract.[22] This immediately implies (see Figure 2) that the optimal restart contract takes the following form: there exist sequences (k_t, U_t) and $(\hat{k}_t, \hat{U}_t)$ and a number $k(\theta_H)$ such that $\forall t$, $k(\theta_L|\theta_L^{t-1}) = \hat{k}_t$, $U(\theta_H|\theta_L^{t-1}) = \hat{U}_t$ and $\forall h^{t-1}$,

[22]This could be shown by the argument similar to Lemma 2 in the appendix.

$k(\theta_H|h^{t-1}) = k(\theta_H)$ and $U(\theta_L|h^{t-1}) = 0$,

$$k(\theta_L|h^{t-1},\theta_H,\theta_L^{s-1}) = k_s \quad \text{and} \quad U(\theta_H|h^{t-1},\theta_H,\theta_L^{s-1}) = U_s,\ \forall s.$$

The optimal restart contract can now be characterized.

Proposition 6. The following supply contract characterizes the restart optimum:

$$k^R(\theta_H|h^{t-1}) = k^R(\theta_H) \geq k^e(\theta_H),$$
$$k^R(\theta_L|h^{t-1}) = \begin{cases} \mathcal{K}_L(\hat{\rho}_t), & \text{if } h^{t-1} = \theta_L^{t-1}, \\ \mathcal{K}_L(\rho_s), & \text{if } h^{t-1} = (h^{\tau-1},\theta_H,\theta_L^s)\,,\ \text{s.t. } \tau + s = t-1, \end{cases}$$

where $\hat{\rho}_t = \max\{b\hat{\rho}_{t-1} + a_L, \gamma\}$, $\hat{\rho}_1 \geq \frac{\mu_H}{\mu_L}$ and $\rho_t = \max\{b\rho_{t-1} + a_L, \gamma\}$, $\gamma < \rho_1 \leq a_H$ for some $\gamma \in \left[\frac{a_L}{1-b}, a_H\right]$.[23]

The optimal restart contract resembles the first-order optimal one (see Proposition 2), but there are three noticeable differences: (i) the high type allocation is (potentially) distorted upwards, (ii) the initial distortion at the lowest history is higher and that in the restart phase is lower, and (iii) there is a floor on distortions, so the contract has a finite memory along consecutive low TFP shocks.[24] Closed form expressions of the distortions and the floor are determined by analyzing the complementary slackness of "upward" incentive constraints.

How well does the optimal restart contract perform? By definition, the principal's profit from the optimal restart contract is lower than the optimal contract, $\Pi^R \leq \Pi^*$. Unfortunately, the gap between the two is very hard to theoretically compute when the first-order approach is not valid. However, we can still bound the loss by using the expression for the first-order optimum, $\Pi^\#$, which is calculable in closed form. Since $\Pi^* \leq \Pi^\#$, we must have $\Pi^* - \Pi^R \leq \Pi^\# - \Pi^R$.[25]

We use sensitivity analysis to assess the gap. Attach a Lagrange multiplier to each "upward" incentive constraint and evaluate the multipliers at the restart

[23]In fact, it is easy to show that $\frac{a_L}{1-b} \leq a_H$ holds for any parameter constellations of δ_A, δ_P, α_H and α_L. Hence, the interval is never empty.

[24]However, it must be noted that the optimal restart contract has positive memory in that it is not the same as the static optimum, it does strictly better than the repetition of the static optimum.

[25]Calculating the first-order optimum involves the maximization of the same objective in ($\star$) but with a strict subset of constraints, so even if the first-order approach is not valid it gives an upper bound on the optimal value of the objective, Π^*.

optimum. Quantify how much slack needs to be added to these constraints so that the solution then coincides with the first-order optimum. [26,27] The general estimate can then be written as

$$\Pi^{\#} - \Pi^{R} \leq \text{Lagrange multipliers} \cdot \text{Slack}.$$

Proposition 7. There exist two bounds, B_a and B_r, function of primitives Γ, such that $\Pi^* - \Pi^R \leq B_a(\Gamma)$ and $1 - \frac{\Pi^R}{\Pi^*} \leq B_r(\Gamma)$.

One is an additive bound, and the other is a bound on the ratio. In the appendix we provide closed form expressions in terms of fundamentals. Figure 5 depicts the loss from using the optimal restart contract for a specific example- as before we set $\theta_L = 1$, $\delta_P = 0.8$ and $R(k) = 2\sqrt{k}$. The unshaded region represents the validity of the first-order approach so the optimal restart contract coincides with the first-order optimum. When the first-order approach is not valid the analytical bound never exceeds 3.5 percent and the actual loss is never more than 2 percent.[28]

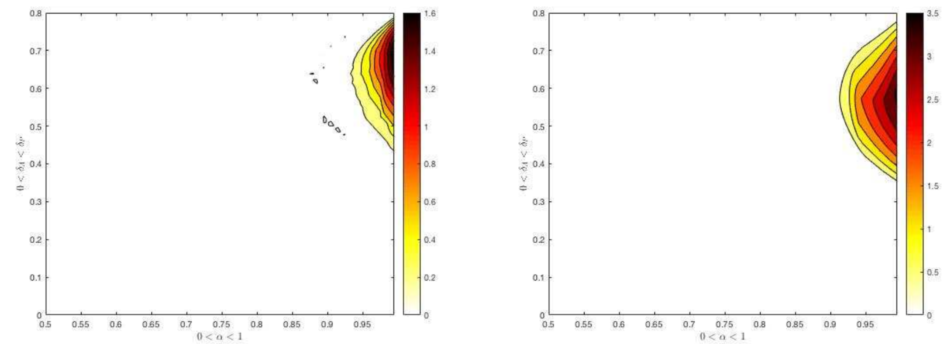
6 Comparative Statics

Does the principal favor the impatient agent or the patient agent and what determines the ranking if there exists any? Broadly speaking, if the Markov process

[26]Formally, we look at the problem of maximizing principal's profit Π over the class of restart contracts subject to two sets of incentive constrains, namely "downward" (IC_H) and "upward" (IC_L): $\max_{m: m \text{ is restart}} \Pi(m)$ subject $IC_H(m) \geq 0$ and $IC_L(m) \geq 0$. Here, we use the notation $IC_i(m) \geq 0$ to indicate that agent's utility if truth-telling minus his utility if deviating is nonnegative. Our goal is to quantify principal's profit at the solution, say m^R. To do this, consider the relaxed problem when IC_L was not present. In this case, the solution is the so-called first-order optimal contract $m^{\#}$. Next, consider the auxiliary problem: $\max_{m: m \text{ is restart}} \Pi(m)$ subject to $IC_H(m) \geq 0$ and $IC_L(m) \geq -\varepsilon$, and denote its solution by $m^A(\varepsilon)$ with the corresponding Lagrange multiplier $\lambda(\varepsilon)$. Clearly, $m^A(0) = m^R$ and $m^A[IC_L(m^{\#})] = m^{\#}$, that is $\varepsilon = IC_L(m^{\#})$ is the "minimal" slack needed for IC_L not to bite. It turns out that $\Pi[m^A(\varepsilon)]$ viewed as a function of ε is convex, therefore by the envelope argument: $\Pi(m^{\#}) - \Pi(m^R) \leq \lambda(0) \cdot IC_L(m^{\#})$.

[27]Our approach of slacking upward incentive constraints and quantifying the loss associated from the exercise has a flavor of Madarász and Prat (2017) where a robust approach to multidimensional screening entails the principal giving up profits in order to relax global incentive constraints.

[28]By actual loss, we mean the exact numerical value of the loss associated with using the optimal restart contract as opposed to the optimal contract, and by analytical loss we mean the value of the theoretical bound, B_r, for which no optimization is required, it is simply a function of the fundamentals of the model.



(a) actual loss (b) analytical loss

Figure 5: Percentage loss, $\left(1 - \frac{\Pi^R}{\Pi^*}\right) * 100$. $\alpha_H = 1 - \alpha_L = \alpha$ on the x-axis and δ_A on the y-axis.

is not too persistent (in the neighborhood of iid), then the principal prefers the patient agent, and if it is very persistent (in the neighborhood of constant types), the principal prefers the impatient agent. While a global comparative static is elusive, we can find a theoretical result for the limit cases and provide clear numerical arguments for the intermediate ones.

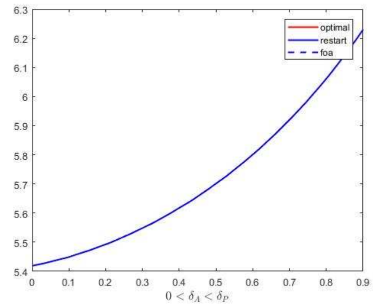
Proposition 8. Let $\alpha_H = 1 - \alpha_L = \alpha$. Principal's ex ante payoff in the first-order optimal, optimal and optimal restart contracts varies with δ_A as follows:

(a) principal prefers patient agent ($\delta_A = \delta_P$) for α sufficiently close to $\frac{1}{2}$.

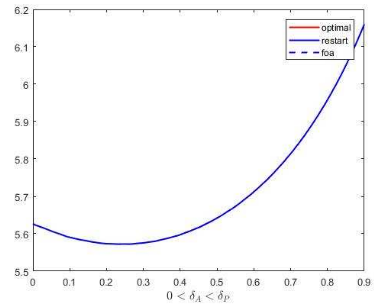
(b) principal prefers myopic agent ($\delta_A = 0$) for α sufficiently close to 1.

Figure 6 plots principal's profit in the first-order optimal contract (dotted blue), optimal contract (red) and the optimal restart contract (blue) as a function of δ_A for the different persistence levels of symmetric Markov chain, $\alpha \in \{0.7, 0.8, 0.9, 0.95, 0.99\}$, all other parameters are the same as before. Conceptually, the principal has to internalize two types of costs – standard information rent and intertemporal cost of incentive provision, and two types of benefits – standard surplus generated by the transaction and the gain from differential discounting.

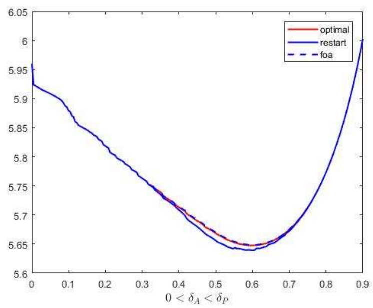
At very low levels of persistence the standard information rent the principal has to pay is quite low, she extracts a large part of the surplus as profit, and does not find it worthwhile to pay the extra intertemporal cost of incentive provision to benefit from differential interest rates. As persistence increases the traditional information rent goes up and the intertemporal cost of incentive provision goes



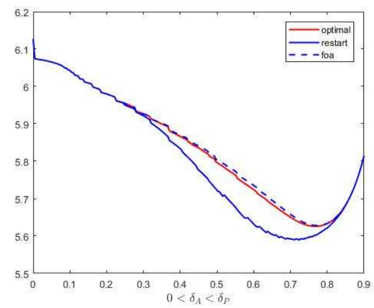
(a) $\alpha = 0.7$



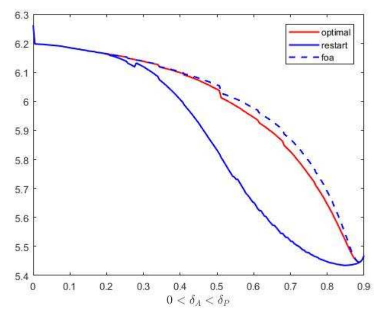
(b) $\alpha = 0.8$



(c) $\alpha = 0.9$



(d) $\alpha = 0.95$



(e) $\alpha = 0.99$

Figure 6: Principal's profit

down. Therefore, the principal's preference for the forward-looking aptitude of the agent is proportional to the strength of the agent as measured by the extent of his private information. Interestingly, for intermediate levels of persistence, say $\alpha = 0.9$, the princpal prefer either a completely myopic agent ($\delta_A = 0$) or completely forward looking one ($\delta_A = \delta_P$), but not goldilocks, see Figure 6c. The non-monotonicity is generated by the rate of change of benefits- standard economic surplus and gains from differential discounting- as a function of discounting and persistence.

7 Concluding remarks

We analyzed a dynamic principal-agent model with persistent private information and unequal discounting. Unequal discounting creates intertemporal gains from trade, and its interaction with Markovian private information produces intertemporal costs of incentive provision. The optimal contract is completely characterized; two key properties underlying the dynamics are restart and shutdown. When the first-order approach does not work, we also characterize the optimal restart contract which provides a simpler and approximately optimal alternative.

Unequal discounting has been explored to varying degrees in dynamic games and contracts. It is well known that in repeated interactions with differential rate of time preference, payoffs for the impatient player can be frontloaded and the set of equilibria expands favoring the patient player (see the classic Lehrer and Pauzner (1999)). In a very elegant paper, Opp and Zhu (2015) analyze the general relational contracting model of Ray (2002) with unequal discounting. They, however, do not deal with agency frictions, all actions and information are publicly observable. Incentive constraints therein are the equivalent of punishment phase in repeated games, a resort to autarky on deviation from the prescribed plan. The threat of autarky generates backloading of payments and unequal discounting does the frontloading, leading to a cyclical pattern similar to our paper.

Biais et al. (2007) explored the implications of unequal discounting in a dynamic model of moral hazard with limited liability and the possibility of liquidation. There exists a reflective boundary below the efficient level that pushes the optimal contract back towards the liquidation region, and the contract is liquidated almost surely in the long-run. The propagation of distortions is sustained in our model through persistence in agency frictions whereas the same is done in their framework by limited liability and the threat of liquidation.[29]

Our paper is also related to the political economy and public finance literature that uses unequal discounting as a motivation for long-run distortions. Acemoglu et al. (2008) show that when politicians are less patient than the citizens, positive aggregate labor and capital taxes are charged forever to correct for political econ-

[29]Biais et al. (2007) also invoke unequal discounting for a technical reason- the continuous time limit of their discreet time model is not well defined for equal discounting. No such problem exists in our framework.

omy distortions. Farhi and Werning (2007) find that with risk averse agents in an overlapping generations model when the social discount factor is higher than the private one, consumption exhibits mean reversion with no immiseration.[30] While the former contains the long-run inefficiency flavor of our results, the latter shows dynamics similar to the optimality of restarts.

One can also ask the question – what if the agent is more patient than the principal? Though most of our applications fit the patient principal model, this is an interesting theoretical question in its own right. It turns the model as stated is then not "compact"; the lack of an upper bound on transfers that the principal can pay means that the agent will lend the principal an unbounded amount of money in a hope to claw it back in the future. Imposing an upper bound rectifies the problem – the optimal allocation rule in the equal discounting case continues to be the optimum for the model with $\delta_A > \delta_P$, and transfers are uniquely pinned down through the upper bound.

Going forward, we believe it will be useful to study the dynamics generated by the interaction of persistent private and unequal discounting under the presence of one or some combination of the following economics forces: privately known discounting, hidden savings, risk aversion and limited liability.

8 Appendix

8.1 Sequential approach

First, we establish the set of binding constraints in Lemmata 1 and 2. The former says that the individual rationality constraints of the low type bind in ($\star$). The latter claims that the "downward" incentive compatibility constraints bind in the relaxed problem.

Lemma 1. Let m be any incentive compatible and individually rational contract with $U(\theta_L|h^{t-1}) > 0$ for some h^{t-1}. There exists another incentive compatible and individually rational contract m' with $U'(\theta_L|h^{t-1}) > 0$, and it delivers higher ex-ante profit to the principal.

[30] A similar mechanism is generated through the interaction of aggregate shocks and unequal discounting in Aguiar et al. (2009) with an application to foreign direct investment and sovereign debt.

Proof. Suppose $h^{t-1} \neq \emptyset$. Alter m by decreasing $U(\theta_L|h^{t-1})$ by small $\varepsilon > 0$, but keeping $U(\theta_H|h^{t-1}) - U(\theta_L|h^{t-1})$ fixed. The new contract is still incentive-feasible and the net change of objective is proportional to $\delta_P^{t-2}(\delta_P - \delta_A)\mathbb{P}(h^{t-1}) > 0$. The similar argument applies to $h^{t-1} = \emptyset$. □

The bottom line of Lemma 1 is that there is no loss of generality to set $U(\theta_L|h^{t-1}) = 0$ for any h^{t-1}, which we implicitly impose slightly abusing our notations.

Lemma 2. Take an individually rational contract satisfying the "downward" incentive constraints with $IC_H(h^{t-1})$ being slack for some h^{t-1}. There exists another incentive compatible and individually rational contract with binding $IC_H(h^{t-1})$ that delivers higher ex-ante profit to the principal.

Proof. Suppose $h^{t-1} \neq \emptyset$. Decrease $U(\theta_H|h^{t-1})$ by small $\varepsilon > 0$ so that $IC_H(h^{t-1})$ still holds. The new contract is individually rational, and it satisfies the "downward" incentive constraints. Moreover, the net change of principal's revenue is proportional to $\delta_P^{t-2}(\delta_P - \delta_A)\mathbb{P}\big(h^{t-1}\big) > 0$. The case of $h^{t-1} = \emptyset$ is obvious. □

Proof of Proposition 1. By Lemmata 1 and 2, the solution to the relaxed problem satisfies the downward incentive constraint and individually rationality of the low type as equalities. Using the binding constraints, the objective could be expressed only in terms of allocation as the surplus minus the information rent and intertemporal cost of incentive provision. The precise expressions are derived in the main text in Equations 1 and 2. Clearly, the objective is strictly concave, thus the first-order conditions are sufficient to characterize the first-order optimum. □

Proof of Corollary 1. For the first-order optimal contract, the second period "upward" incentive constraints are trivially satisfied as $k(\theta_H|\theta_i) > k(\theta_L|\theta_i)$ for $i = H, L$. The first period "upward" incentive constraint is implied by $k^{\#}(\theta_H) = k^e(\theta_H)$ and $2\Delta\theta R(k^e(\theta_L)) > U^{\#}(\theta_H)$. □

Proof of Proposition 2. The case of $T = \infty$ is essentially similarly to the two period model, although calculations are heavier. Recall that the cost of implementing an allocation is $\bar{U}_P = \sum_{t=1}^{T} \delta_P^{t-1}\mathbb{E}\left[u(\tilde{\theta}_t|\tilde{h}^{t-1})\right]$, and it could be parsed into the information rent and intertemporal cost of incentive provision:

$$\bar{U}_P = \sum_{t=1}^{T} \delta_P^{t-1} \mathbb{E}\left[u(\tilde{\theta}_t|\tilde{h}^{t-1})\right]$$
$$= \underbrace{\mathbb{E}\left[U(\tilde{\theta}_1)\right]}_{\bar{U}_A\text{: agent's ex ante utility}} + \underbrace{\frac{\delta_P - \delta_A}{\delta_P} \sum_{t=2}^{\infty} \mathbb{E}\left[U(\tilde{\theta}_t|\tilde{h}^{t-1})\right]}_{I\text{: intertemporal cost of incentive provision}}$$

The key is to invoke the binding constraint to obtain the expression for $U(\theta_L|h^{t-1}) = 0$ and $U(\theta_H|h^{t-1})$ as a function of $k(\theta_L|h^{t-1}, \theta_L^s)$ with $s \geq 0$ as given in Equation 4, Equation 5 directly follows from

$$\bar{U}_A = \mu_H U(\theta_H) = \mu_H \sum_{s=0}^{\infty} (\delta_A(\alpha_H - \alpha_L))^s \Delta\theta R(k(\theta_L|\theta_L^s)).$$

To obtain Equation 6, notice that

$$\sum_{t=2}^{\infty} \mathbb{E}\left[U(\tilde{\theta}_t|\tilde{h}^{t-1})\right] = \sum_{h^{t-1}} \delta_P^{t-1} \mathbb{P}(h^{t-1}, \theta_H) \sum_{s=0}^{\infty} (\delta_A(\alpha_H - \alpha_L))^s \Delta\theta R(k(\theta_L|h^{t-1}, \theta_L^s)).$$

We will simplify this expression by fixing the position of the last θ_H. In particular, for the lowest history, exchange the order of summation to get

$$\sum_{t=2}^{\infty} \delta_P^{t-1} \mathbb{P}(\theta_L^{t-1}, \theta_H) \sum_{s=0}^{\infty} (\delta_A(\alpha_H - \alpha_L))^s \Delta\theta R(k(\theta_L|\theta_L^{t+s-1})) =$$
$$= \sum_{t=2}^{\infty} \delta_P^{t-1} \mathbb{P}^{t-1}(\theta_L^t) \Delta\theta R(k(\theta_L|\theta_L^{t-1})) \frac{\alpha_L}{1-\alpha_L} \sum_{s=0}^{t-2} \left(\frac{\delta_A}{\delta_P} \frac{\alpha_H - \alpha_L}{1-\alpha_L}\right)^s,$$

which is exactly the first term in Equation 6 defining I. The second term is derived similarly by first summing over the histories $\{(h^{t-1}, \theta_H, \theta_L^s)\}_{s \geq 0}$ for fixed h^{t-1}, and then over h^{t-1}. □

Proof of Corollary 2. Consider $f(x) = bx + a_L$ with $b = \frac{\delta_A}{\delta_P} \frac{\alpha_H - \alpha_L}{1-\alpha_L}$ and $a_i = \frac{\delta_P - \delta_A}{\delta_P} \frac{\alpha_i}{1-\alpha_i}$ for $i = H, L$. So, the distortions satisfy $\rho_{t+1} = f(\rho_t)$ and $\hat{\rho}_{t+1} = f(\hat{\rho}_t)$ for all t. It is easy to see that f has only one non-zero fixed point, namely $c = \frac{a_L}{1-b}$, and it is globally stable. So, *(a), (b)* and *(c)* are established. To see *(d)* and *(e)*, recall the definition of $\mathcal{K}_L(x) = (R')^{-1}\left(\frac{1}{\theta_L - x\Delta\theta}\right)$ for $x\Delta\theta < \theta_L$ and zero otherwise. □

8.2 Recursive approach

In this subsection, we simplify ($\star$) and formulate its recursive analogue mentioned in the main text. We introduce an auxiliary sequential problem to derive ($\mathcal{RP}$). Let $\Pi(\theta_t|h^{t-1})$ be the expected lifetime profit at the end of date t, assuming truthful reporting at date t and further

$$\Pi(\theta_t|h^{t-1}) = s\big(k(\theta_t|h^{t-1}), \theta_t\big) - u(\theta_t|h^{t-1}) + \delta_P \mathbb{E}\left[\Pi(\tilde{\theta}_{t+1}|h^{t-1}, \theta_t)|\theta_t\right].$$

Suppose that the agent truthfully reported (h^{t-1}, θ_j) before date $t \geq 2$. In addition, the principal committed to deliver exactly w to the high type at this date. Then, if possible, define $S_j(w)$ by

$$(\mathcal{SP}) \quad S_j(w) = \max_{\langle \mathbf{U}, \mathbf{k} \rangle} \alpha_j \left[\Pi(\theta_H|h^{t-1}, \theta_j) + w\right] + (1-\alpha_j)\Pi(\theta_L|h^{t-1}, \theta_j),$$

$$\text{s.t. } U(\theta_H|h^{t-1}, \theta_j) = w, \text{ and } IC_i(h^{t+s}),\ IR_H(h^{t+s}), \quad \forall h^{t+s} \in H^{t+s}\Big|_{(h^{t-1}, \theta_j)}, \forall s.$$

Notice that the optimal value is independent of h^{t-1}, thus we simply write $S_j(w)$.

Let W be the largest set of w such that the constraints set in ($\mathcal{SP}$) is non-empty. W is the familiar recursive domain described in Spear and Srivastava (1987) and it has a very simple structure.

Claim 1 (Recursive domain)**.** $W = \mathbb{R}_+$.

Proof. First of all, $w \geq 0$ by $IR_H(h^{t-1}, \theta_j)$. To see that the program is feasible for $w \geq 0$, take $k(\theta_H|h^{t-1}, \theta_j) = R^{-1}\left(\frac{w}{\Delta\theta}\right)$ and $k(\theta_H|h^{t+s}) = k(\theta_L|h^{t+s}) = U(\theta_H|h^{t+s}) = 0$ for any $h^{t+s} \in H^{t+s}\Big|_{(h^{t-1}, \theta_j)}$ $\forall s \neq 0$. □

It is easy to see that ($\mathcal{SP}$) could be restated as ($\mathcal{RP}$), and the problem at the initial date is equivalent to ($\diamond$). To formally show equivalence of the sequential and recursive formulations, we need to introduce auxiliary definitions. The policy correspondence is a correspondence which maps w into $(\mathbf{Z}(w), \mathbf{K}(w))$ that is the set of optimal choices in ($\mathcal{RP}$). We say that a contract is generated from the policy correspondence if $k(\theta_i|\theta_j, h^{t-1}) \in \mathbf{K}_i(U(\theta_H|\theta_j, h^{t-1}))$ and $U(\theta_H|\theta_j, h^{t-1}, \theta_i) \in \mathbf{Z}_i(U(\theta_H|\theta_j, h^{t-1}))$ for $i, j = H, L$ and $\forall h^{t-1}$, $\forall t$.

Claim 2.

(a) There exists a unique continuous bounded function satisfying the Bellman equation in $(\mathcal{RP})$.

(b) The policy correspondence is non-empty, compact-valued and upper hemi-continuous.

(c) A contract is generated from the policy correspondence if and only if it solves $(\mathcal{RP})$ with $w = U(\theta_H|\theta_j)$ for $j = H, L$.

(d) Value functions in $(\mathcal{SP})$ and $(\mathcal{RP})$, as well as in $(\star)$ and $(\diamond)$ coincide.

Proof. The result follows from Exercises 9.4, 9.5 in Stokey et al. (1989). □

In the rest of the subsection, we outline several standard properties of the value function (Claim 3), establish uniqueness of transfers (Claim 4) and prove Propositions 4, 5.

Claim 3 (Properties of the value function)**.**

(a) Each S_j is concave.

(b) Each S_j is continuously differentiable on $\mathbb{R}_{++}$.

(c) Each S_j is locally strictly concave at w satisfying $S_j'(w) > 0$.

Proof.

Part (a). The argument is standard, we need to show that the Bellman operator, implicitly defined in $(\mathcal{RP})$, preserves concavity. Indeed, the constraints set is convex and $s(\cdot, \theta)$ is concave. So, concavity is preserved by the Bellman operator. Since the set of concave functions is closed in the space of continuous bounded functions, the result follows from Theorem 3.1 and its Corollary 1 in Stokey et al. (1989).

Part (b). We established concavity of the value function using the standard argument. As for differentiability, the standard argument of Benveniste and Scheinkman (1979) is not applicable in our context, because it might not to be possible to change $\mathbf{k}$ keeping $\mathbf{z}$ constant. We give a different argument that is close to Rincón-Zapatero and Santos (2009) in its spirit. We shall use the fact S_j is concave, thus it is subdifferentiable. Take m^* which solves $(\mathcal{SP})$ with

$U^*(\theta_H|\theta_j) = w$. Using the generalized first-order and envelope conditions for $(\mathcal{RP})$, we argue that there exists some finite time s such that the value function is differentiable at $U^*(\theta_H|\theta_j, \theta_L^{s-1})$. Then, the value function turns out to be differentiable at the original point, w.

Before we show differentiability, we shall validate that the first-order conditions are sufficient to characterize a solution. In particular, we show that Slater's condition holds which is sufficient to guarantee that the first-order approach with Lagrange multipliers in l^1 is valid in $(\mathcal{SP})$, because of concavity and boundedness of these problems (see Morand and Reffett (2015)).

We claim that, for any $w > 0$, there exists a feasible point such that the constraint map is uniformly bounded away from 0. The argument is constructive. Since $w > 0$, there exists $k_H > k_L > 0$ satisfying:

$$\frac{\Delta\theta}{1 - \delta_A(\alpha_H - \alpha_L)} R(k_L) < w < \frac{\Delta\theta}{1 - \delta_A(\alpha_H - \alpha_L)} R(k_H)$$

Take $k(\theta_H|\theta_j, h^{t-1}) = k_H$, $k(\theta_L|\theta_j, h^{t-1}) = k_L$ and $U(\theta_H|\theta_j, h^{t-1}) = w \ \forall h^{t-1} \ \forall t$.

Now, we are in a position to show that S_j is continuously differentiable. Let m^* be a solution to $(\mathcal{SP})$ at $t = 2$. It is clear that the capital supplied to θ_H can be distorted only upwards, thus $k^*(\theta_H|\theta_j, h^{t-1}) > 0$ is uniquely defined $\forall h^{t-1}$ by strict concavity of the objective. In addition, if $k^*(\theta_L|\theta_j, h^{t-1}) > 0$, then it is unique by strict concavity of the objective.

Next, consider $(\mathcal{RP})$, its solution exists and coincides with one found in $(\mathcal{SP})$ by the previous claim. Since S_j is concave, its subdifferential at $w > 0$ is well-defined and it equals to $\partial S_j(w) = [S_j^+(w), S_j^-(w)]$, and at $w = 0$ it is $S_j^+(0)$ where a plus/minus denotes a right/left subderivative.

Let $\alpha_j \rho_H$ and $(1-\alpha_j)\rho_L$ be Lagrange multipliers for the "upward" and "downward" incentive constraints, respectively. And, $\rho_j(w)$ be some Lagrange multiplier supporting a solution, whereas $\rho_j^-(w)/\rho_L^+(w)$ be the highest/smallest such Lagrange multiplier. Finally, denote by $(\mathbf{z}(w), \mathbf{k}(w))$ some point in the optimal correspondence.

The first-order conditions with respect to $\mathbf{k}$ are $k_i(w) = \mathcal{K}_i(\rho_i(w))$ for $i = H, L$ where $\mathcal{K}_H(x) = (R')^{-1}\left(\frac{1}{\theta_H + x\Delta\theta}\right)$ and $\mathcal{K}_L(\cdot)$ is defined as before. By the above argument, $\mathbf{K}_H(w)$ is a singleton and $\rho_H^+(w) = \rho_H^-(w) = \rho_H(w)$ for any w. In addition, if $k_L(w) > 0$, then $\mathbf{K}_L(w)$ is a singleton and $\rho_L^+(w) = \rho_L^-(w) = \rho_L(w)$.

So, for $w > 0$, the Lagrange multipliers might be not unique only if there exists some $\rho_L(w) \geq \theta_L/\Delta\theta > 0$. Given this $\rho_L(w) > 0$, the "downward" incentive constraint binds and we have that $z_L(w) = \frac{w}{\delta_A(\alpha_H - \alpha_L)} > w > 0$ is uniquely defined.

Then, the envelope conditions give $S_j^-(w) - S_j^+(w) = (1-\alpha_j)(\rho_L^-(w) - \rho_L^+(w))$. It is immediate that S_j is differentiable at w if and only if $\rho_L(w)$ is unique. The first-order condition with respect to z_L when $z_L(w) > 0$ reads as follows:

$$\delta_P S_L^-(z_L(w)) \geq \alpha_L(\delta_P - \delta_A) + (\alpha_H - \alpha_L)\delta_A \rho_L(w) \geq \delta_P S_L^+(z_L(w))$$

If $\rho_L(z_L(w))$ is unique, then $\rho_L(w)$ is so and S_j is differentiable at w. Now, define recursively $z_L^s = z_L(z_L^{s-1})$ with $z_L^0 = w > 0$ for some selection from $\bar{z}_L$. There are two potential cases, namely $\rho_L(z_L^s)$ is unique for some s or it is not for all s. In the former case, S_j is differentiable at w by our previous argument. In the latter case, $z_L^s = \frac{w}{\delta_A^s(\alpha_H - \alpha_L)^s} \to \infty$ as $s \to \infty$ which is impossible, because any solution must be in l^∞.

Finally, continuous differentiability of S_j is implied by differentiability and concavity.

Part (c). Suppose that $S_j'(w) = S_j'(w+\varepsilon) > 0$ for some $w, \varepsilon > 0$. Consider m^* and m^ε solving $(\mathcal{SP})$ at w and $w+\varepsilon$, respectively. Since $s(\cdot, \theta)$ is strictly concave, it must be that $\mathbf{k}^* = \mathbf{k}^\varepsilon$. Otherwise, it would be the case that $S_j'(w) < S_j'(w+\varepsilon)$.

Now, since $S_j'(w) = S_j'(w+\varepsilon) > 0$, the envelope theorem implies that the "downward" incentive constraint binds in each case. By the first-order and envelope conditions, see Equations 7, 8 and 9, it will continue to bind along the sequence of θ_L's, thus

$$w = \Delta\theta \sum_{s=0}^{\infty} (\delta_A(\alpha_H - \alpha_L))^s R\big(k^*(\theta_L | h^{t-2}, \theta_j, \theta_L^s)\big) = w + \varepsilon.$$

The last assertion is a clear contradiction. The similar argument establishes that $S_j'(w - \varepsilon) > S_j'(w)$. □

Now, we derive the optimality conditions which are useful for our characterization of the optimal contract. Let $(1-\alpha_j)\rho_H$ and $\alpha_j \rho_L$ be Lagrange multipliers on the constraints in $(\mathcal{RP})$. And, let $\mu_H \rho_H$ and $\mu_L \rho_L$ be Lagrange multipliers on the constraints in $(\diamond)$. We denote by $(\mathbf{z}(w), \mathbf{k}(w))$ some selection from the optimal correspondence and by $\rho(w)$ some corresponding Lagrange multipliers.

So, the first-order conditions are $k_i(w) = \mathcal{K}_i(\rho_i(w))$ for $i = H, L$ and

$$S'_H\big(z_H(w)\big) - \alpha_H \frac{\delta_P - \delta_A}{\delta_P} + (\alpha_H - \alpha_L)\frac{\delta_A}{\delta_P}\rho_H(w) \begin{cases} = 0, & \text{if } z_H(w) > 0, \\ \leq 0, & \text{if } z_H(w) = 0, \end{cases} \tag{7}$$

$$S'_L\big(z_L(w)\big) - \alpha_L \frac{\delta_P - \delta_A}{\delta_P} - (\alpha_H - \alpha_L)\frac{\delta_A}{\delta_P}\rho_L(w) \begin{cases} = 0, & \text{if } z_L(w) > 0, \\ \leq 0, & \text{if } z_L(w) = 0. \end{cases} \tag{8}$$

In addition, the Envelope theorem gives

$$S'_j(w) = (1 - \alpha_j)\rho_L(w) - \alpha_j \rho_H(w), \text{ for } j = H, L. \tag{9}$$

We proceed by characterizing properties of the recursive optimum. Although, S_j might be not globally strictly concave, we are able to identify next period promised utilities when the incentive constraints do not bind. To be specific, $z_L(w) = z^e_L$ if the "downward" constraint is slack and $z_H(w) = z^e_H$ if the "upward" constraint is slack. By part (c) of Claim 2, there exists unique z^e_j satisfying $z^e_j > 0$ and $S'_j(z^e_j) = \alpha_j \frac{\delta_P - \delta_A}{\delta_P}$ or $z^e_j = 0$ and $S'_j(0) \leq \alpha_j \frac{\delta_P - \delta_A}{\delta_P}$. Then, define two thresholds $w^*_j = \Delta\theta R(k^e(\theta_j)) + \delta_A(\alpha_H - \alpha_L)z^e_j > 0$.

We also argue that the Lagrange multipliers are unique. Let m^* be a solution to $(\mathcal{SP})$ at $t = 2$. It is clear that the capital supplied to θ_H can be distorted only upwards, thus $k^*(\theta_H|h^{t-2}, \theta_j) > 0$ is uniquely defined by strict concavity of the objective. It follows from Claim 2 that $\rho_H(w) = \mathcal{K}_H^{-1}(k^*(\theta_H|h^{t-2}, \theta_j))$, and $\rho(\cdot)$ is continuous in w, because m^* changes continuously with w. It remains to select $\rho_L(w)$ to satisfy the envelope condition.

Proof of Proposition 4. We shall characterize ρ, because its properties translate into $\mathbf{k}$ by the first-order condition $k_i(w) = \mathcal{K}_i(\rho(w))$ for $i = H, L$.

Part (b). If there is no "upward" incentive constraint, then $k_H(w) = k^e(\theta_H)$ and $z_H = z^e_H$ by the first-order conditions and definition of z^e_H. Since this choice is feasible if and only if $w \geq w^*_H$, the result for ρ_H follows. To see monotonicity of $\rho_H(\cdot)$, take $w' > w \geq w^*_H$ and suppose that $\rho_H(w) \geq \rho_H(w')$. Concavity and the first-order conditions imply that $z_H(w) \geq z_H(w')$ which contradicts to $\Delta\theta(R \circ \mathcal{K}_H)(\rho_H(w)) + \delta_A(\alpha_H - \alpha_L)z_H(w) = w < w' = \Delta\theta(R \circ \mathcal{K}_H)(\rho_H(w')) +$

$\delta_A(\alpha_H - \alpha_L)z_H(w')$.

Part (a). By the similar argument to part (b), $\rho_L(\cdot)$ is strictly decreasing on $[0, w_L^*]$, and it is zero afterwards. Finally, since the only feasible choice at $w = 0$ is $k_L(0) = 0$, $w_k^o = \sup\{w \in W : k_L(w) = 0\}$ is well-defined.

$\square$

Now, we turn our attention to $\mathbf{z}$ and start by pointing out uniqueness of transfers.

Claim 4 (Uniqueness of transfers)**.** $\mathbf{Z}_L$ is single-valued, and $\exists$ unique $\bar{w}$ such that $\mathbf{Z}_H$ is single-valued whenever $w_L^* \geq w_H^*$ or $w \neq \bar{w}$. $\bar{w}$ solves $(\alpha_H - \alpha_L)\delta_A\rho_H(\bar{w}) = \alpha_H(\delta_P - \delta_A)$.

Proof. z_L is unique which follows from the last part of Claim 3, whereas z_H might fail to be unique. Intuitively, z_H could be not unique only when there are multiple z_H with $\rho_L\big(z_H(w)\big) = \rho_H\big(z_H(w)\big) = 0$. Such values of z_H are elements of the correspondence $\mathbf{Z}_H$.

Define $\bar{w}$ by $(\alpha_H - \alpha_L)\delta_A\rho_H(\bar{w}) = \alpha_H(\delta_P - \delta_A)$. Clearly, it exists and it is unique, because of monotonicity of ρ_H as shown in the proof of Proposition 4.

Suppose that $w_L^* \geq w_H^*$, then $S_j'(w) = (1 - \alpha_j)\rho_L(w) - \alpha_j\rho_H(w)$ is strictly decreasing on $\mathbb{R}_+$. So, z_H is single-valued by strict concavity of S_j.

If $w_L^* < w_H^*$, then the envelope conditions (Equation 9) imply that $S_j'(w) > 0$ on $[0, w_L^*]$, $S_j'(w) < 0$ on $[w_H^*, +\infty)$ and $S_j'(w) = 0$ for any $w \in [w_L^*, w_H^*]$. Therefore, $\mathbf{Z}_H$ is single-valued on $[0, \bar{w})$ by the last part of Claim 3, and $\mathbf{Z}_H(\bar{w}) = [w_L^*, w_H^*]$ by construction. To see that $\mathbf{Z}_H$ is single-valued on $(\bar{w}, +\infty)$, notice that $w = \Delta\theta(R \circ \mathcal{K}_H)(\rho_H(w)) + \delta_A(\alpha_H - \alpha_L)z_H(w)$ whenever $\rho_H(w) > 0$. Since $\rho_H(w) > 0$ for any $w > \bar{w}$, $z_H(w)$ could be uniquely identified from the "upward" incentive constraint. $\square$

To sum up, $z_H(w)$ is not unique only when $w_L^* < w_H^*$ and $w = \bar{w}$. In what follows, by $z_H(\cdot)$ we mean an arbitrary selection from $\mathbf{Z}_H(\cdot)$.

Proof of Proposition 5.

Part (d). Equation 9 says that $S_H'(w)/\alpha_H - S_L'(w)/\alpha_L = \frac{\alpha_L - \alpha_H}{\alpha_H\alpha_L}\rho_L(w) \leq 0$. Therefore, $z_H^e \leq z_L^e$ with $z_L^e \neq z_H^e$ if and only if $S_L'(0) > \alpha_L\frac{\delta_P - \delta_A}{\delta_P}$ by their definitions and part (c) of Claim 3. For $z_L^e = 0$, $w_L^* > z_L^e$ is trivially satisfied. Suppose that $z_L^e > 0$, then $S_j'(w_L^*) = -\alpha_j\rho_H(w_L^*) \leq 0 < S_j'(z_L^e)$, thus $w_L^* > z_L^e$.

Moreover, notice that $w_H^* = \Delta\theta R(k^e(\theta_H)) + \delta_A(\alpha_H - \alpha_L)z_H^e \leq z_H^e$ if and only if $z_H^e \geq \frac{\Delta\theta}{1-\delta_A(\alpha_H-\alpha_L)} R(k^e(\theta_H))$. On the other hand, $z_H^e < \frac{\Delta\theta}{1-\delta_A(\alpha_H-\alpha_L)} R(k^e(\theta_L))$, because of $z_H^e \leq z_L^e < w_L^*$. So, we can not have $z_H^e \geq w_H^*$.

It remains to establish that $z_H^e \leq w^f$. Of course, it is vacuously true whenever $z_H^e = 0$. So, suppose that $z_H^e > 0$. In this case, $z_H^e \leq w^f$ whenever $\frac{a_L}{1-b} \leq a_H$. To see this, notice that $\rho_L(w^f) \geq \frac{a_L}{1-b}$ with an equality if and only if $\rho_H(w^f) = 0$, as shown in part (c). Suppose that $z_H^e < w^f$, which is equivalent to $\rho_L(w^f) > \rho_L(z_H^e)$ by monotonicity of $\rho_L(\cdot)$. Since $z_H^e < w_H^*$, $\rho_H(w^f) = \frac{a_L}{1-b}$, which contradicts to $\rho_L(w^f) > \rho_L(z_H^e) > 0$.

Recall that $\frac{a_L}{1-b} \leq a_H$ if and only if $\frac{\alpha_L}{1-\alpha_L} \leq \frac{\alpha_H}{1-\alpha_H}\left(1 - \frac{\delta_A}{\delta_P}\frac{\alpha_H-\alpha_L}{1-\alpha_L}\right)$ which is always satisfied.

Parts (a) and (b). We established above that $z_j^e \in [0, w_j^*]$ for $j = H, L$. Monotonicity of $\rho(\cdot)$ as shown in Proposition 4 combined with Equations 7 and 8 yields the result of parts (a) and (b).

Part (c). First, we study fixed points of $Z_H(\cdot)$. In the previous part, we showed that $z_H^e < w_H^*$ which implies that z_H^e is a fixed point of $Z_H(\cdot)$. Suppose that there exists $w \neq z_H^e > 0$ with $w \in \mathbf{Z}_H(w)$. By definition, it must be the case that $\rho_H(w) > 0$.

Consider the equation $w = \frac{\Delta\theta}{1-\delta_A(\alpha_H-\alpha_L)}\big(R\circ\mathcal{K}_H\big)\big(\rho_H(w)\big) > \frac{\Delta\theta}{1-\delta_A(\alpha_H-\alpha_L)} R(k_H^e)$ which is necessary for $w \in \mathbf{Z}_H(w) > 0$ with $\rho_H(w) > 0$. Equation 7 and 9 imply that $(1-\alpha_H)\delta_P\rho_L(w) = \alpha_H(\delta_P - \delta_A) + \big(\alpha_H\delta_P - (\alpha_H - \alpha_L)\delta_A\big)\rho_H(w) > 0$.

Since $\rho_L(w) > 0$, the "downward" constraint binds this period and it will keep binding along the sequence of θ_L's. Formally, let $z_L^s(w)$ be defined by $z_L^s(w) = z_L\big(z_L^{s-1}(w)\big)$ with $z_L^0(w) = w$. By Equation 8, $\rho(z_L^s(w)) > 0$ for any s. Then, iterating along this sequence, we arrive at the following contradiction by using monotonicity of R:

$$w = \Delta\theta \sum_{\tau=0}^{+\infty} (\delta_A(\alpha_H - \alpha_L))^\tau \big(R\circ\mathcal{K}\big)\big(\rho_L(z_L^\tau(w))\big) < \frac{\Delta\theta}{1-\delta_A(\alpha_H-\alpha_L)} R(k_L^e)$$

So, z_H^e is the unique fixed point of $\mathbf{Z}_H$.

Now, we turn our attention to fixed points of z_L. Of course, 0 is always a fixed point, and our goal is to identify a positive fixed point. Suppose there exists $0 < w = z_L(w)$. First of all, $z_L(w) = z_L^e < w_L^* \leq w$ whenever $\rho_L(w) = 0$, therefore it must be the case that $w < z_L^e$ and $\rho_L(w) > 0$.

Consider the equation $w = \frac{\Delta\theta}{1-\delta_A(\alpha_H-\alpha_L)}(R \circ \mathcal{K}_L)(\rho_L(w))$ which is necessary when $w = z_L(w) > 0$ with $\rho_L(w) > 0$. One more necessary condition, due to the Equations 8 and 9, is that $\big((1-\alpha_L)\delta_P - \delta_A(\alpha_H - \alpha_L)\big)\rho_L(w) = \alpha_L(\delta_P - \delta_A) + \alpha_L\delta_P\rho_H(w) > 0$. By monotonicity of ρ (shown in Proposition 4), these two equations have a root if and only if $\theta_L > \frac{a_L}{1-b}\Delta\theta$. And, if such a root exists, then it is unique.

Let w^f be the root of the aforementioned equations for $\theta_L > \frac{a_L}{1-b}\Delta\theta$, and $w^f = 0$, otherwise. For $\theta_L > \frac{a_L}{1-b}\Delta\theta$, global stability follows from $z_L(\cdot)$ crossing the 45-degree line only once and from above, because $w^f < z_L^e$. For $\theta_L/\Delta\theta \leq c$, global stability is trivial, because 0 is the unique fixed point.

Part (e). At the initial date, the first-order conditions with respect to $\mathbf{z}$ coincide with Equations 7 and 8. The extra first condition is $\mu_L\rho_L(w) - \mu_H\rho_H(w) = (\leq)\mu_H$ whenever $w > (=)0$. Existence and uniqueness directly follows from monotonicity of ρ, see proof of Proposition 4. To see that the contract always stays within $[0, w_L^*]$, notice that $S_L'(z_L(w)) > 0$, due to Equation 8, implying that $\rho_L(w) > 0$. For $w \leq w_L^*$, $|z_H(w) - z_H^e| \leq |w - z_H^e|$ yields $z_H(w) \leq w_L^*$, because $z_H^e < w_L^*$ as shown before. □

8.3 Connection to primitives

Proof of Proposition 3 and Corollary 4. First, we show that the first-order optimal contract is optimal if and only if $\max\left\{U^{\#}(\theta_H), \lim_{t\to\infty} U^{\#}(\theta_H|\theta_L^{t-1})\right\} \leq C = \Delta\theta R(k^e(\theta_H)) + \delta_A(\alpha_H - \alpha_L)U^{\#}(\theta_H|\theta_H)$. Given history h^{t-1}, $U^{\#}(\theta_H|h^{t-1}) - C$ is the expected utility which θ_L could obtain by misreporting his type once, and the "upward" incentive constraint requires this object to be non-positive.

By Corollary 2, $U^{\#}(\theta_H|h^{s-1}, \theta_H, \theta_L^{t-1})$ is increasing in t with $\lim_{t\to\infty} U^{\#}(\theta_H|\theta_L^{t-1}) = \lim_{t\to\infty} U^{\#}(\theta_H|h^{s-1}, \theta_H, \theta_L^{t-1})$ for all h^{s-1} and s. In addition, $U^{\#}(\theta_H|\theta_L^{t-1})$ is either globally decreasing or increasing in t depending on the primitives. Obtain the result by combining these two observations.

Next, we establish Proposition 3 and the second part of Corollary 4. Let w^* be the point chosen at initialization. Clearly, z_H^e is attained in finite time, say t^*, along the sequence of θ_H's starting from any w^*. Since $z_H^e \leq w_L^f$, Proposition 5 yields that the optimal contract never leaves the interval $[z_H^e, w^f]$.

Suppose that $w^f \leq w_H^*$, then the "upward" incentive constraints do not have a bite after t^* periods with probability one. Equations 7, 8 and 9 yield that for any $w \in [z_H^e, w^f]$, $\rho_L(z_H(w)) = a_H$ and $\rho_L(z_L(w)) = a_L + b\rho_L(w)$. In other words, the optimal contract will follow the first-order optimal contract described in Proposition 2, and $w^f = \lim_{t\to\infty} U^{\#}(\theta_H|\theta_L^{t-1})$, $w_H^* = C$.

Conversely, suppose $w^f > w_H^*$ and that the optimal contract is eventually restart with some t^*. By Proposition 5, there exists $t > t^*$ such that $w_H^* < z_L^t(z_H^e) < w^f$ implying that $z_H(z_L^t(z_H^e)) \neq z_H^e$ where $z_L^t(\cdot)$ is a product of t consecutive applications of $z_L(\cdot)$ to w. This is a clear contradiction. □

8.4 Optimal restart contract

In this subsection, we characterize the optimal restart contract and assess its performance. Extending Lemma 2, one can show that not only agent's allocation, but his expected utility also follows a restart pattern for the optimal restart contract. Therefore, we represent a restart contract by a pair of sequences $\{U_t, k_t\}$ and $\{\hat{U}_t, \hat{k}_t\}$ as in Remark 3.

Proof of Proposition 6. First, we adjust our previous definitions to respect a structure of restart contracts. A restart contract satisfies the "downward" incentive constraints if for all t,

$$U_t \geq \Delta\theta R(k_t) + \delta_A(\alpha_H - \alpha_L)U_{t+1},$$
$$\hat{U}_t \geq \Delta\theta R(\hat{k}_t) + \delta_A(\alpha_H - \alpha_L)\hat{U}_{t+1}.$$

A restart contract satisfies the "upward" incentive constraints if for all t,

$$U_t \leq \Delta\theta R(k(\theta_H)) + \delta_A(\alpha_H - \alpha_L)U_1,$$
$$\hat{U}_t \leq \Delta\theta R(k(\theta_H)) + \delta_A(\alpha_H - \alpha_L)U_1.$$

Now, we derive principal's expected revenue of a restart contract. Let S_t be the surplus in the restart phase given that θ_L was drawn $t-1$ times since since the last θ_H.

$$\begin{aligned} S_1 &= \alpha_H\Big[s(k(\theta_H),\theta_H) - \alpha_H(\delta_P-\delta_A)U_1 + \delta_P S_1\Big] \\ &\quad + (1-\alpha_H)\Big[s(k_1,\theta_L) - \alpha_L(\delta_P-\delta_A)U_2 + \delta_P S_2\Big], \\ S_t &= \alpha_L\Big[s(k(\theta_H),\theta_H) - \alpha_H(\delta_P-\delta_A)U_1 + \delta_P S_1\Big] \\ &\quad + (1-\alpha_L)\Big[s(k_t,\theta_L) - \alpha_L(\delta_P-\delta_A)U_{t+1} + \delta_P S_{t+1}\Big]. \end{aligned}$$

Next, we solve for principal's expected revenue:

$$\begin{aligned} \Pi = &- \mu_H\hat{U}_1 + \mu_L\sum_{t=1}^{\infty}(\delta_P(1-\alpha_L))^{t-1}\big(s(\hat{k}_t,\theta_L) - \alpha_L(\delta_P-\delta_A)\hat{U}_{t+1}\big)+ \\ &+ \zeta\Big[s(k(\theta_H),\theta_H) - \alpha_H(\delta_P-\delta_A)U_1 \\ &\quad + \delta_P(1-\alpha_H)\sum_{t=1}^{\infty}(\delta_P(1-\alpha_L))^{t-1}\big(s(k_t,\theta_L) - \alpha_L(\delta_P-\delta_A)U_{t+1}\big)\Big] \end{aligned}$$

where $\zeta = \frac{\alpha_L\delta_P+\mu_H(1-\delta_P)}{(1-\delta_P)(1-\delta_P(\alpha_H-\alpha_L))}$. The first term is agent's expected utility, the second term is expected surplus along the lowest history and the third is expected surplus of the restart phase.

First, we ignore the "upward" incentive constraints and maximize Π in the set of restart contracts respecting the "downward" incentive constraints. By Proposition 2, the unique solution is the first-order optimal contract which has $U_t^{\#} := \Delta\theta R(k_t^{\#}) + \delta_A(\alpha_H-\alpha_L)U_{t+1}^{\#}$ and $\hat{U}_t^{\#} := \Delta\theta R(\hat{k}_t^{\#}) + \delta_A(\alpha_H-\alpha_L)\hat{U}_{t+1}^{\#}$ with $k_t^{\#} = \mathcal{K}\big(\frac{a_L}{1-b}(1-b^{t-1}) + a_H b^{t-1}\big)$ and $\hat{k}_t^{\#} = \mathcal{K}_L\big(\frac{a_L}{1-b}(1-b^{t-1}) + \frac{\mu_H}{\mu_L}b^{t-1}\big)$. Let $\Pi^{\#}$ be principal's expected revenue for this contract, then $\Pi^{\#} \geq \Pi^*$ with equality if and only if this contract satisfies the "upward" incentive constraints.

Now, we impose the "upward" incentive constraints and maximize Π in the set of restart contracts respecting IC. Let Π^R be principal's expected revenue for this contract, then $\Pi^R \leq \Pi^*$ with equality if and only if this contract satisfies the "upward" incentive constraints. Define the following Lagrange multipliers for each t:

1. $\zeta\delta_P(1-\alpha_H)(\delta_P(1-\alpha_L))^{t-1}\rho_t$ is the multiplier on $U_t \geq \Delta R(k_t)$ $+ \delta_A(\alpha_H-\alpha_L)U_{t+1}$
2. $\zeta\delta_P(1-\alpha_H)(\delta_P(1-\alpha_L))^{t-1}\eta_t$ is the multiplier on $U_t \leq \Delta R(k(\theta_H))$

$+ \delta_A(\alpha_H - \alpha_L)U_1$

3. $\mu_L(\delta_P(1-\alpha_L))^{t-1}\hat{\rho}_t$ is the multiplier on $\hat{U}_t \geq \Delta R(k_t) + \delta_A(\alpha_H - \alpha_L)\hat{U}_{t+1}$

4. $\mu_L(\delta_P(1-\alpha_L))^{t-1}\hat{\eta}_t$ is the multiplier on $\hat{U}_t \leq \Delta R(k(\theta_H)) + \delta_A(\alpha_H - \alpha_L)U_1$

The first-order conditions are given by $k(\theta_H) = \mathcal{K}_H(\xi)$ and $\forall t$

$$k_t = \mathcal{K}_L(\rho_t), \quad \text{for} \quad \rho_{t+1} = a_L + b\rho_t + \eta_{t+1} \quad \text{with } \rho_1 = a_H - \frac{\delta_A}{\delta_P}\frac{\alpha_H - \alpha_L}{1-\alpha_H}\xi + \eta_1,$$
$$\hat{k}_t = \mathcal{K}_L(\hat{\rho}_t), \quad \text{for} \quad \hat{\rho}_{t+1} = a_L + b\hat{\rho}_t + \hat{\eta}_{t+1} \quad \text{with } \hat{\rho}_1 = \frac{\mu_H}{\mu_L} + \hat{\eta}_1,$$

where $\xi = \sum_{t=1}^{\infty}(\delta_P(1-\alpha_L))^{t-1}\left(\frac{\mu_L}{\zeta}\hat{\eta}_t + \zeta\delta(1-\alpha_H)\eta_t\right)$.

If the "upward" incentive constraints do not bind, then $\xi = 0$ and the first-order contract is optimal. This contract has an infinite memory along the sequence of θ_L's.

So, consider the case that some "upward" incentive constraints bind that is $\xi > 0$. Using complementary slackness, it is easy to see that an optimal restart contract is such that

$$\rho_{t+1} = \max\{\gamma, a_L + b\rho_t\},$$
$$\hat{\rho}_{t+1} = \max\{\gamma, a_L + b\hat{\rho}_t\},$$

for some $\frac{a_L}{1-b} \leq \gamma \leq \rho_1 \leq a_H$. The constant γ is a floor on the distortions along the sequence of θ_L's. In addition, $\eta_1 = 0$ meaning that the optimal restart contract always has some memory, and it is not a static one. To see it, suppose that $\eta_1 > 0$, then $U_1 = \Delta\theta R(k(\theta_H)) + \delta_A(\alpha_H - \alpha_L)U_1$. Since $\frac{a_L}{1-b} \leq \gamma \leq \rho_1$, ρ_t is a non-increasing, therefore U_t is a non-decreasing sequence. Then, the "upward" incentive constraints always bind in the restart phase, and $U_t = U_1$, $k_t = \mathcal{K}_L(\gamma)$. Combing both incentince constraints obtain that $R(k(\theta_H)) = (R \circ \mathcal{K}_L)(\gamma)$ which is a contradiction, because $k(\theta_H) \geq k^e(\theta_H)$ and $\mathcal{K}_L(\gamma) \leq k^e(\theta_L)$. □

Proof of Proposition 7. First, we shall bound $\Pi^{\#} - \Pi^R > 0$. Define the slack variables for the upward incentive constraints by $\varepsilon_t = \left(\hat{U}_t^{\#} - \Delta R(k^e(\theta_H)) - \delta_A(\alpha_H - \alpha_L)U_1^{\#}\right)^+$ and $\hat{\varepsilon}_t = \left(\hat{U}_t^{\#} - \Delta R(k^e(\theta_H)) - \delta_A(\alpha_H - \alpha_L)U_1^{\#}\right)^+$. By the standard perturbation argument,

$$\Pi^{\#} - \Pi^R \leq \sum_{t=1}^{\infty} (\delta_P(1-\alpha_L))^{t-1} \Big[\mu_L \hat{\eta}_t \hat{\varepsilon}_t + \zeta \delta (1-\alpha_H) \eta_t \varepsilon_t \Big],$$

because the "downward" incentive constraints always bind.

Our first bound takes $\varepsilon^{max} = \max\Big\{ \hat{\varepsilon}_0, \Big(\frac{\Delta}{1-\delta_A(\alpha_H-\alpha_L)} \mathcal{K}\left(\frac{a_L}{1-b}\right) - \Delta R(k^e(\theta_H)) - \delta_A(\alpha_H-\alpha_L) U_1^{\#} \Big)^+ \Big\} \geq \varepsilon_t, \hat{\varepsilon}_t$ for all t. Using the first-order condition for U_1 and $\frac{a_L}{1-b} \leq \rho_1$:

$$\Pi^{\#} - \Pi^R \leq \zeta \xi \varepsilon^{max} \leq \frac{\delta_P(1-\alpha_H)}{\delta_A(\alpha_H - \alpha_L)} (a_H - c) \zeta \varepsilon^{max} =: B_a^1$$

Our second bound limits η_t and $\hat{\eta}_t$. Notice that $\rho_{t+1} - a_L - b\rho_t = \eta_{t+1} \leq \gamma(1-b) - a_L \leq (1-b)\big(a_H - \frac{a_L}{1-b}\big)$ and the same is true for $\hat{\eta}_{t+1}$ for any $t \geq 2$. For $t = 1$, $\hat{\eta}_1 \leq \big(a_H - \frac{\mu_H}{\mu_L}\big)^+$,

$$\Pi^{\#} - \Pi^R \leq \mu_L \big(a_H - \frac{\mu_H}{\mu_L}\big)^+ \hat{\varepsilon}_1 + \mu_L (1-b) \big(a_H - \frac{a_L}{1-b}\big) \cdot \sum_{t=2}^{\infty} (\delta_P(1-\alpha_L))^{t-1} \Big[\mu_L \hat{\varepsilon}_t + \zeta \delta (1-\alpha_H) \varepsilon_t \Big] =: B_a^2.$$

Our last bound relies on the optimal static contract. A static contract is such that $k_t = \hat{k}_t = k(\theta_L)$ and $U_t = \hat{U}_t = U(\theta_H)$ for all t. It is easy to show that the optimal static contract has $k^S(\theta_H) = k^e(\theta_H)$ and $k^S(\theta_L) = \mathcal{K}_L(\rho)$ where

$$\rho = \frac{(\mu_H + \zeta(\delta_P - \delta_A)\alpha_H)(1-\delta_P(1-\alpha_L))}{(\mu_L + \zeta \delta_P (1-\alpha_H))(1-\delta_A(\alpha_H - \alpha_L))} + \frac{\alpha_L(\delta_P - \delta_A)}{1-\delta_A(\alpha_H-\alpha_L)}.$$

The profit of the optimal static contract can be found in the closed form, Π^S, using the binding "downward" incentive constraints. And, we have $\Pi^{\#} - \Pi^R \leq \Pi^{\#} - \Pi^S$. Then,

$$\Pi^* - \Pi^R \leq \min\{B_a^1, B_a^2, \Pi^{\#} - \Pi^S\} =: B_a \quad \text{and}$$

$$1 - \frac{\Pi^R}{\Pi^*} \leq \frac{B_a/\Pi^{\#}}{\max\{1 - B_a/\Pi^{\#}, \Pi^S\}} =: B_r.$$

□

8.5 Comparative statics

Proof of Proposition 8. We start by looking at the first-order optimal contract. By Corollary 2, then the first-order optimal contract is essentially static for $\alpha = \frac{1}{2}$. Formally, $\rho_t = \frac{\delta_P - \delta_A}{\delta_P}$ for any t, $\hat{\rho}_t = \frac{\delta_P - \delta_A}{\delta_P}$ for $t \geq 2$, and $\hat{\rho}_1 = \frac{\mu_H}{\mu_L}$. Importantly, $\bar{U}_A$ is independent of δ_A, so $\delta_A = \delta_P$ uniquely maximizes the surplus and minimizes the cost of incentive provision at the same time. Since the profit in the first-order optimal contract is continuous with respect α and $\delta_A = \delta_P$ is a strict maximizer for $\alpha = 1$, it is still a maximizer for $\alpha \approx \frac{1}{2}$.

If $\alpha \to 1$, then $\hat{\rho}_t \to \frac{\mu_H}{\mu_L} \left(\frac{\delta_A}{\delta_P}\right)^{t-1}$ $\forall t$, the intertemporal cost of incentive provision goes to zero. Therefore, $\lim_{\alpha \to 1} \bar{U}_P = \lim_{\alpha \to 1} \bar{U}_A$, and the limit is strictly increasing in δ_A. By continuity, $\delta_A = 0$ is a maximizer for $\alpha \approx 1$.

Finally, by Corollary 2 the first-order optimal contract is incentive compatible for either iid or constant types. Therefore, the proposition is true for the optimal and optimal restart contracts as well. □

References

Acemoglu, D., Golosov, M., and Tsyvinski, A. (2008). Political economy of mechanisms. *Econometrica*, 76(3):619–641.

Aguiar, M., Amador, M., and Gopinath, G. (2009). Investment cycles and sovereign debt overhang. *Review of Economic Studies*, 76(1):1–31.

Battaglini, M. (2005). Long-term contracting with markovian consumers. *American Economic Review*, 95(3):637–658.

Battaglini, M. and Lamba, R. (2017). Optimal dynamic contracting: the first-order approach and beyond. Cornell University and Pennsylvania State University.

Benveniste, L. M. and Scheinkman, J. A. (1979). On the differentiability of the value function in dynamic models of economics. *Econometrica*, 47:727–732.

Bergemann, D. and Strack, P. (2015). Dynamic revenue maximization: a continuous time approach. *Journal of Economic Theory*, 159:819–853.

Bergemann, D. and Välimäki, J. (2017). Dynamic mechanism design: an introduction. Yale University and Aalto University.

Besanko, D. (1985). Multi-period contracts between principal and agent with adverse selection. *Economics Letters*, 17(1-2):33–37.

Biais, B., Mariotti, T., Plantin, G., and Rochet, J.-C. (2007). Dynamic security design: convergence to continuous time and asset pricing implications. *Review of Economic Studies*, 74(2):345–390.

Campello, M., Graham, J. R., and Harvey, C. R. (2010). The real effects of financial constraints: evidence from a financial crisis. *Journal of Financial Economics*, 97(3):470–487.

Carroll, C. D. (1992). The buffer-stock theory of saving: some macroeconomic evidence. *Brookings Papers on Economic Activity*, 1992(2):61–156.

Chassang, S. (2013). Calibrated incentive contracts. *Econometrica*, 81(5):1935–1971.

Clementi, G. L. and Hopenhayn, H. A. (2006). A theory of financing constraints and firm dynamics. *Quarterly Journal of Economics*, 121:229–265.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica*, 59(5):1221–1248.

DeMarzo, P. M. and Sannikov, Y. (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance*, 61(6):2681–2724.

Eaton, J. and Gersovitz, M. (1981). Debt with potential repudiation: theoretical and empirical analysis. *Review of Economic Studies*, 48(2):289–309.

Edmans, A., Fang, V. W., and Lewellen, K. A. (2017). Equity vesting and investment. *Review of Financial Studies*, 30(7):2229–2271.

Evans, D. S. (1987). The relationship between firm growth, size, and age: estimates for 100 manufacturing industries. *Journal of Industrial Economics*, 35(4):567–581.

Farhi, E. and Werning, I. (2007). Inequality and social discounting. *Journal of Political Economy*, 115(3):365–402.

Fernandes, A. and Phelan, C. (2000). A recursive formulation for repeated agency with history dependence. *Journal of Economic Theory*, 91(2):223–247.

Fuster, A. and Willen, P. S. (2017). Payment size, negative equity, and mortgage default. *American Economic Journal: Economic Policy*, forthcoming.

Garrett, D., Pavan, A., and Toikka, J. (2018). Robust predictions of dynamic optimal contracts. Toulouse School of Economics, Northwestern University and MIT.

Gomes, J. F. (2001). Financing investment. *American Economic Review*, 91(5):1263–1285.

İmrohoroğlu, A. and Tüzel, Ş. (2014). Firm-level productivity, risk, and return. *Management Science*, 60(8):2073–2090.

Krähmer, D. and Strausz, R. (2015). Dynamic mechanism design. In *An Introduction to the Theory of Mechanism Design by Tilman Börgers*, chapter 11, pages 204–234. Oxford University Press.

Krasikov, I. and Lamba, R. (2017). A theory of dynamic contracting with financial constraints. Pennsylvania State University.

Laffont, J.-J. and Tirole, J. (1993). *A theory of incentives in procurement and regulation*. MIT press.

Lehrer, E. and Pauzner, A. (1999). Repeated games with differential time preferences. *Econometrica*, 67(2):393–412.

Lyon, B. R. (1996). Empirical relevance of efficient contract theory: inter-firm contracts. *Oxford Review of Economic Policy*, 12(4):27–52.

Madarász, K. and Prat, A. (2017). Sellers with misspecified models. *Review of Economic Studies*, 84(2):790–815.

Morand, O. and Reffett, K. (2015). Lagrange multipliers in convex programs with applications to classical and nonoptimal stochastic one-sector growth models. University of Connecticut and Arizona State University.

Mussa, M. and Rosen, S. (1978). Monopoly and product quality. *Journal of Economic theory*, 18(2):301–317.

Opp, M. M. and Zhu, J. Y. (2015). Impatience versus incentives. *Econometrica*, 83(4):1601–1617.

Pavan, A. (2016). Dynamic mechanism design: robustness and endogenous types. In *2015 World Congress of the Econometric Society*.

Pavan, A., Segal, I., and Toikka, J. (2014). Dynamic mechanism design: a myersonian approach. *Econometrica*, 82(2):601–653.

Ray, D. (2002). The time structure of self-enforcing agreements. *Econometrica*, 70(2):547–582.

Rincón-Zapatero, J. P. and Santos, M. S. (2009). Differentiability of the value function without interiority assumptions. *Journal of Economic Theory*, 144(5):1948–1964.

Sannikov, Y. (2014). Moral hazard and long-run incentives. Stanford University.

Spear, S. E. and Srivastava, S. (1987). On repeated moral hazard with discounting. *Review of Economic Studies*, 54(4):599–617.

Stokey, N. L., Lucas Jr, R. E., and Prescott, E. (1989). *Recursive methods in economic dynamics*. Harvard University Press.

Tirole, J. (1988). *The theory of industrial organization*. MIT press.

Vohra, R. V. (2012). Dynamic mechanism design. *Surveys in Operations Research and Management Science*, 17(1):60–68.