

Impatience vs. Incentives*

Marcus M. Opp[†] John Y. Zhu[‡]

September 30, 2014

Abstract

This paper studies the long-run dynamics of contracts in repeated principal-agent relationships with an impatient agent. Despite the absence of exogenous uncertainty, Pareto-optimal dynamic contracts generically oscillate between favoring the principal and favoring the agent.

Econometrica 3rd round revision

*We are grateful for various insightful discussions with Willie Fuchs and Nicolae Gârleanu and helpful comments from three anonymous referees, as well as Matthew Jackson, the co-editor. Moreover, we want to thank Peter DeMarzo, Darrell Duffie, Mike Fishman, Rich Kihlstrom, Natalia Kovrijnykh, Jonathan Levin, Dmitry Livdan, Andrey Malenko, John Morgan, Christine Parlour, Ilya Strebulaev, Tim Worrall and Bilge Yilmaz as well as seminar participants at MIT (Organizational Economics Lunch), Berkeley-Stanford (Joint Finance Seminar), the Econometric Society Summer meeting (USC), the Society of Economic Dynamics, and the Midwest Economics Association Meeting for helpful comments.

[†]University of California, Berkeley (Haas), e-mail: mopp@haas.berkeley.edu.

[‡]University of Pennsylvania, Wharton, e-mail: zhuyiran@wharton.upenn.edu.

1 Introduction

We study optimal contracting in a repeated principal-agent framework where the agent is more impatient than the principal. Differential discounting creates gains from trade across time and tends to push optimal contracts toward favoring the more patient principal in the long-run. This *impatience* force conflicts with the *incentives* force that tends to push optimal contracts toward favoring the agent in the long-run. We show that Pareto-optimal contracts generically resolve this impatience vs. incentives conflict by oscillating between favoring the principal and favoring the agent over time.

In our model, a contract stipulates for each period a transfer to the agent and an action. Our model admits a broad interpretation of “action.” It can be effort by the agent or investment by the principal or a collaborative venture by both. Each action generates a surplus, a portion of which is transferred to the agent according to the contract. If the agent deviates from the contract, he receives a deviation payoff that is a function of the action that was supposed to be taken, but loses a fraction of the stipulated transfer. A contract is incentive-compatible if the agent never wants to deviate and the principal’s interim participation constraint is always satisfied.

In this setting, we first prove the existence of a unique stationary Pareto-optimal contract - the steady state. We then show that all non-stationary Pareto-optimal contracts oscillate around this focal steady state. These contracts may feature oscillating transfers with constant action or co-moving oscillating transfers and actions. Oscillation can be damped, converging to the steady state, or persist in the long run. In the latter case, the amplitude of oscillation can grow over time causing even arbitrarily low participation constraints to bind in the long run, distorting contract dynamics.

There are two features of the model that drive the oscillation phenomenon: First, the agent is more impatient than the principal. Second, the agent loses a fraction of the stipulated transfer when he deviates. The first feature ensures that Pareto-optimal contracts have binding IC constraints. Binding IC constraints plus the second feature imply that any above steady state transfer to the agent must be followed by a below steady state transfer, and any below steady state transfer must be followed by an above steady state one. Oscillation emerges.¹

Our setting is broadly applicable, in the spirit of Ray (2002), and nests environments studied in many influential papers, such as Thomas and Worrall (1988), Thomas and Worrall (1994), and Albuquerque and Hopenhayn (2004). Yet, we show by adding even an infinitesimal amount of relative impatience on the agent side, virtually all Pareto-optimal contracts oscillate around a focal Pareto-optimal steady state. This contrasts with the standard result under equal discounting (e.g., Becker and Stigler (1974), Harris and Holmstrom (1982), and Ray (2002)), when payments can always be backloaded without affecting payoffs, and any Pareto-optimal payoff can be sustained by a contract that favors the agent in the long-run.

The rest of the paper is organized as follows: Section 2 describes our general model and its applicability to various agency problems, and provides a basic intuition for the

¹Thus, cycles require neither exogenous shocks/uncertainty (see Aguiar, Amador, and Gopinath (2009)) nor self-fulfilling changes in beliefs as in Zhu, Wright, and He (2013), Gu, Mattesini, Monnet, and Wright (2013) and Rocheteau and Wright (2013).

oscillation principle. Section 3 presents all formal results.

2 Model

The model is an infinitely-repeated principal agent relationship with perfect information and transferable utility. It consists of four ingredients: discount factors $\delta_P \geq \delta_A$ for the principal and agent, respectively, a deviation parameter $\theta \in [0, 1]$, a mapping (π, d) from an abstract action set \mathcal{A} to \mathbb{R}^2 with compact image, and a principal outside option O_P . Each action a produces surplus $\pi(a)$. Let u_A denote the portion transferred to the agent. A contract is a sequence of transfers and actions $\{(\tilde{u}_{A,t}, \tilde{a}_t)\}_{t=0}^\infty$. There exist public randomization devices, so each $\tilde{u}_{A,t}$ and \tilde{a}_t can be random. For every agent promised value U_A , the Principal's Problem is the following maximization:

$$\max_{\{(\tilde{u}_{A,t}, \tilde{a}_t)\}_{t=0}^\infty} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \delta_P^t (\pi(\tilde{a}_t) - \tilde{u}_{A,t}) \right]$$

s.t.

$$\tilde{u}_{A,t} + \delta_A U_{A,t+1} \geq D(\tilde{u}_{A,t}, \tilde{a}_t) := (1 - \theta)\tilde{u}_{A,t} + d(\tilde{a}_t) \quad \forall t \quad \forall \text{ realizations of } (\tilde{a}_t, \tilde{u}_{A,t}) \quad (1)$$

$$U_{P,t} \geq O_P \quad \forall t \quad (2)$$

$$U_{A,0} \geq U_A \quad (3)$$

where $U_{A,t} := \mathbb{E}_t \left[\sum_{s=t}^{\infty} \delta_A^{s-t} \tilde{u}_{A,s} \right]$ and $U_{P,t} := \mathbb{E}_t \left[\sum_{s=t}^{\infty} \delta_P^{s-t} (\pi(\tilde{a}_s) - \tilde{u}_{A,s}) \right]$ denote the date t continuation payoffs of the agent and principal.

(1) is the agent's incentive-compatibility constraint which requires, for every random realization, that the agent's date t continuation payoff $\tilde{u}_{A,t} + \delta_A U_{A,t+1}$ be weakly larger than the total payoff from his *best possible* deviation $D(\tilde{u}_{A,t}, \tilde{a}_t)$. Following Ray (2002), one can think of D as the result of a maximization over a potentially large set of available deviations: θ is the fraction of the current-period transfer lost under deviation and $d(a)$ is the residual deviation payoff which depends on a and includes all subsequent agent payoffs derived from an outside option or under a punishment equilibrium within the game.

(2) is the principal's interim participation constraint.² From now on all contracts are assumed to satisfy (1) and (2). Our goal is to show that when $\delta_P > \delta_A$, the following is generically true:

Theorem 1. *There exists a unique steady state - a Pareto-optimal contract with a constant continuation payoff process $\{(U_{A,t}, U_{P,t})\}_{t=0}^\infty$. The steady state action a^s does not maximize static surplus $\pi(a)$:*

$$a^s = \arg \max_{a \in \mathcal{A}} \pi(a) - \frac{\delta_P - \delta_A}{\theta(\delta_P - \delta_A) + \delta_A} d(a). \quad (4)$$

²Alternatively, one can impose an agent interim participation constraint, or both principal and agent interim participation constraints. The results do not change. Further discussion is provided at the end of the main text, see Remark 2.

If $\theta = 0$ or 1 , every non-steady state Pareto-optimal contract has a continuation payoff process that converges monotonically to the steady state. If $\theta \in (0, 1)$, every non-steady state Pareto-optimal contract has a continuation payoff process that oscillates around the steady state. This oscillation persists in the long run if and only if $\theta \in \left[\frac{\delta_A}{1+\delta_A}, \frac{\delta_A}{\delta_P+\delta_A} \right]$.

Informally, “generically true” means true when corner conditions don’t get in the way. For example, an action solving (4) generically does not maximize surplus but there are clearly exceptions for certain discrete or sufficiently kinked π and d . On the other hand, if one only wants to prove that there are Pareto-optimal payoffs that can only be supported by oscillating contracts, then no further assumptions need to be made except for O_P being sufficiently low.

Our flexible model can speak to a wide spectrum of agency problems:³

Example 1. A government (A) allows a multinational firm (P) to invest I in the country. Investment generates output $Y(I)$ for the multinational and taxes τ . The government can expropriate output up to $Y(I)$, but forfeits tax income forever.

Example 2. An entrepreneur (A) seeks a lender (P) to help finance a product. There are a number of different ways to develop the product. Each option o_i requires some outlay I_{o_i} from the lender and generates some return Y_{o_i} for the entrepreneur. The lender receives a loan repayment R . The entrepreneur can keep Y_{o_i} and strategically default on R , in which case the lender can take the entrepreneur to court. With probability $1 - \theta$ the lender prevails and recoups R . Otherwise he receives nothing.

Example 3. An owner (P) has access to a set of projects and can choose to collaborate with a worker (A) to implement a subset of them. Each subset $\{p_i\}$ requires effort cost $c_{\{p_i\}}$ from the worker, $C_{\{p_i\}}$ from the owner, and produces $\sum_{\{p_i\}} e_{p_i}$ for the owner. In return, the worker receives an up-front wage w . The worker can shirk and keep the wage, in which case the worker is fired but captures unemployment benefits valued at $\delta_A O_A$.

	\mathcal{A}	u_A	π	D	θ	d
Ex. 1	I	τ	$Y(I) - I$	$Y(I)$	1	$Y(I)$
Ex. 2	o_i	$Y_{o_i} - R$	$Y_{o_i} - I_{o_i}$	$Y_{o_i} - (1 - \theta)R$	θ	θY_{o_i}
Ex. 3	$\{p_i\}$	$w - c_{\{p_i\}}$	$\sum_{\{p_i\}} e_{p_i} - c_{\{p_i\}} - C_{\{p_i\}}$	w	0	$c_{\{p_i\}} + \delta_A O_A$

Table 1: Mapping the examples to the model. Notice the action can be taken by the principal (Ex. 1), the agent (Ex. 2), or both (Ex. 3); it can be pecuniary or non-pecuniary, a single object or a set.

The three examples demonstrate how our model can encode various timing and pay conventions. In Example 1, the firm pays taxes *after* observing whether the government has expropriated or not. So the government does not receive any tax revenue in the

³Example 1 is drawn from Thomas and Worrall (1994) and Opp (2012); Example 2 is a modified version of Albuquerque and Hopenhayn (2004) or Clementi and Hopenhayn (2006); Example 3 is inspired by Ray (2002) and Thomas and Worrall (1988).

period it expropriates: $\theta = 1$ and $D = d$. If expropriation occurs, the strongest punishment equilibrium involves the firm withdrawing from the country and so the government receives no future tax revenue either. Therefore, if the government does expropriate its best strategy is to steal the *entire* firm output: $D = Y(I)$. A contract $\{u_{A,t}, I_t\}_{t=0}^{\infty}$ is incentive compatible if and only if $U_{A,t} \geq Y(I_t)$ for all t . In Example 2, the lender has recourse *after* a default and $(1 - \theta)$ represents the probability of contractual enforcement. Alternatively, it could be that the borrower puts up collateral *before* any action, and θ represents the collateral requirement as in Kiyotaki and Moore (1997).⁴ In Example 3, the wage to the worker is pre-paid and so $\theta = 0$ captures the fact that even if the worker shirks, he keeps the wage that has already been paid to him. Of course, if the worker can only abscond with a fraction of the wage due to some inefficiencies, such as the banker in Calomiris and Kahn (1991), then θ can represent the portion of the wage lost during deviation.

Examples 2 and 3 also highlight an important aspect of the model's flexibility. In many applications it is more natural to think of the agent's utility u_A as a sum $m + h(a)$ where $h(a)$ is the component intrinsic to the action a stipulated by the contract and m is the monetary transfer stipulated by the contract. This is in contrast to the model setup where the entire u_A is thought of as the transfer.

The difference is important because when $u_A = m + h(a)$, one should:

1. Apply the θ parameter only to the m component of u_A instead of to the entire u_A .
2. Think of D as $D(m, a) = (1 - \theta)m + \hat{d}(a)$ instead of $D(u_A, a) = (1 - \theta)u_A + d(a)$.

However, Example 2 shows how the model can easily accommodate this mismatch. Simply decompose $\hat{d}(a)$ into $(1 - \theta)h(a) + (\hat{d}(a) - (1 - \theta)h(a))$, define $d(a) := \hat{d}(a) - (1 - \theta)h(a)$, and now $D(m, a)$ can be written in the correct form $D(u_A, a)$.⁵

Remark 1. *Any transferable utility model where agent utility has the more common form $u_A(m, a) := m + h(a)$ and $D(m, a) = (1 - \theta)m + \hat{d}(a)$ is quasilinear in the monetary transfer with $\theta \in [0, 1]$ can be mapped into the model.*

The Oscillation Principle

The basic intuition for oscillation around the steady state relies on two features of the model: inefficient deviation ($\theta > 0$) and relative impatience of the agent ($\delta_A < \delta_P$). In particular, a non-trivial action set is actually not necessary for oscillation to emerge. Thus, to highlight the basic mechanics of oscillation, we will for now consider the simplest version of the model where the action set is a singleton $\{a^s\}$.⁶ Let u_A^s be the steady state utility transfer and let $U_A^s := u_A^s / (1 - \delta_A)$ be the agent's steady state continuation payoff.

⁴The settings of Geanakoplos (2009) and Brunnermeier and Pedersen (2009) show how the collateralization parameter θ can emerge endogenously.

⁵The quantity $d(a) := \hat{d}(a) - (1 - \theta)h(a)$ lacks the economic significance of $\hat{d}(a)$, but that is of no concern since the model does not require d to satisfy anything beyond having a compact image. d has compact image so long as \hat{d} and h have compact image.

⁶Think of a binary effort/shirking decision where the shirking option is never relevant.

To highlight the significance of assuming inefficient deviation, rewrite the IC constraint $u_{A,t} + \delta_A U_{A,t+1} \geq (1 - \theta)u_{A,t} + d(a^s)$ as follows:

$$\theta u_{A,t} + \delta_A U_{A,t+1} \geq d(a^s) \quad (5)$$

Notice $\theta > 0$ means that transfers today relax the IC constraint today.

Relative impatience of the agent implies that IC constraints should always bind. Otherwise, moving some of tomorrow's transfer to today would lead to a Pareto-improvement and IC constraints would still be respected. In particular, the IC constraint of the steady state must bind:

$$\theta u_A^s + \delta_A U_A^s = d(a^s) \quad (6)$$

Now, to see why oscillation emerges, first consider an agent payoff $U_A > U_A^s$. To deliver U_A , the principal can for example provide the agent with an above steady state initial transfer followed by the steady state continuation payoff. But (6) plus the assumption $\theta > 0$ implies that the IC constraint would be slack. Thus, the principal can do better by further frontloading utility until the IC constraint binds. In the end, the agent receives an above steady state transfer today $u_A > u_A^s$ followed by a *below* steady state continuation payoff tomorrow $U_A^+ < U_A^s$.

Next, consider the opposite case $U_A < U_A^s$. Mirroring the previous case, the principal could try a below average initial transfer followed by the steady state continuation payoff. But now (6) plus the assumption $\theta > 0$ implies that the IC constraint is violated. Thus, the initial transfer must be further diminished and the continuation payoff must be increased. In the end, the agent receives a below steady state transfer today $u_A < u_A^s$ followed by an *above* steady state continuation payoff tomorrow $U_A^+ > U_A^s$.

We have now shown that if today's payoff is above the steady state then tomorrow's should be below and if today's is below then tomorrow's should be above. Oscillation around the steady state results.

Notice how inefficient deviation and relative impatience of the agent interact to generate oscillation. Relative impatience makes binding IC constraints uniquely optimal. Then inefficient deviation ensures that binding IC constraints plus above (below) steady state payoffs imply below (above) steady state continuation payoffs.

If either feature is missing, the argument for oscillation falls apart. If deviation is efficient ($\theta = 0$), then to deliver an above (below) steady state payoff with binding IC, the principal can simply provide an above (below) steady state initial transfer followed by the steady state continuation payoff. As a result, all contracts converge monotonically to the steady state. If the principal and agent are equally patient, then Pareto-optimal contracts no longer need to be maximally frontloaded and IC constraints no longer need to bind. Starting with an oscillating contract, one can always further backload payments in a payoff neutral way until the contract continuation payoff process no longer oscillates and instead, converges monotonically to a steady state.

In this primer on oscillation, we have neglected to discuss how participation constraints can distort oscillation and ultimately, the optimal action sequence if the model possesses a nontrivial action set. Participation constraints matter because the higher the θ , the greater the amplitude of oscillation. We will show that when $\theta > \delta_A/(1 + \delta_A)$ oscillations become explosive if we were to show no regard for participation constraints.

This means that any participation constraint, no matter how low, would eventually be violated. Optimally adjusting the explosive oscillation so as to respect participation constraints leads to nontrivial distortions of the action sequence and the oscillation dynamic itself. We now explore this as part of the formal analysis of the Principal’s Problem.

3 Analysis

We start our formal analysis with a preliminary lemma that reveals the role of public randomization and implies that we can restrict our large, abstract action set to a small set of *efficient actions*.

Lemma 1. *Fix a model $(\theta, \delta_A, \delta_P, \mathcal{A}, d, \pi)$. Any alternate model $(\theta, \delta_A, \delta_P, \hat{\mathcal{A}}, \hat{d}, \hat{\pi})$ where $Im(\hat{\mathcal{A}}) = Conv(Im(\mathcal{A}))$ generates the same Pareto-frontier with the same Pareto-optimal continuation payoff processes.*

Certainly the alternate model $\hat{\mathcal{A}}$ can achieve any payoff the original model can achieve. To prove the converse, suppose there was a contract in the alternate model that called for action \hat{a}_t , transfer $\hat{u}_{A,t}$ and continuation payoff $\hat{U}_{A,t+1}$. First, for any action $\hat{a} \in \hat{\mathcal{A}}$, there exists a random action $\tilde{a} \in \mathcal{A}$ satisfying $\mathbb{E}\pi(\tilde{a}) = \pi(\hat{a})$ and $\mathbb{E}d(\tilde{a}) = d(\hat{a})$. Note, however, the IC constraint (1) must now be satisfied for any random realization $d(\tilde{a})$, and not just for the average realization $d(\hat{a})$. This can be achieved by *only* fine-tuning transfers $\tilde{u}_{A,t} = \hat{u}_{A,t} + \frac{d(\tilde{a}_t) - d(\hat{a}_t)}{\theta}$ and leaving the continuation payoff fixed at $\hat{U}_{A,t+1}$. By construction, payoffs are unaffected since $\mathbb{E}\tilde{u}_{A,t} = \hat{u}_{A,t}$, and the continuation payoff process is identical. This proves Lemma 1.

When a model’s $Im(\mathcal{A})$ is convex, public randomization provides no benefits. In particular, all Pareto-optimal payoffs can be delivered by contracts with deterministic actions, transfers, and continuation payoff processes. And since Lemma 1 implies that it is without loss of generality to focus on models where $Im(\mathcal{A})$ is convex, we have now proved:

Corollary 1. *Any Pareto-optimal payoff of any model can be delivered by a contract with a deterministic continuation payoff process.*

In particular, this is true even if we are in a model where any such contract must involve random actions and transfers. This implication is important for the interpretation of our results as it allows us to highlight that oscillation of continuation payoffs is not an artifact of randomization.

Our analysis can be further simplified by noting that any Pareto-optimal contract must only use *efficient* actions. An action a is efficient if for any other action a' , $\pi(a') < \pi(a)$ or $d(a') > d(a)$ or $(\pi(a'), d(a')) = (\pi(a), d(a))$. Let $\hat{\mathcal{A}}^*$ be the set of efficient actions, then it is without loss of generality to focus on models with action space of the form $\hat{\mathcal{A}}^*$. Figure 1 shows a representative $Im(\mathcal{A})$, its convex hull $Im(\hat{\mathcal{A}})$, and the efficient frontier $Im(\hat{\mathcal{A}}^*)$. By construction, $Im(\hat{\mathcal{A}}^*)$ is a concave, strictly increasing function over $[d_{\min}, d_{\max}]$; π is an implicit function of d ; the action space can be identified with the interval $[d_{\min}, d_{\max}]$; and $Im(\hat{\mathcal{A}}^*)$ is just the graph of $\pi(d)$. From now on, for the sake

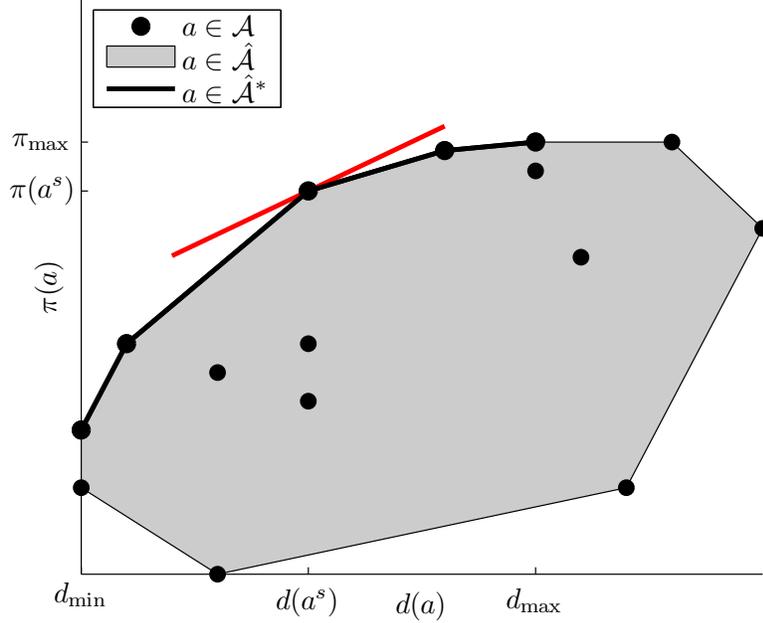


Figure 1: The graph plots the images of the original action set \mathcal{A} , $\hat{\mathcal{A}}$, and the set of efficient actions $\hat{\mathcal{A}}^*$. The latter set defines π as a function of d . The red line has the Pareto-optimal steady state slope $\frac{\delta_P - \delta_A}{\theta(\delta_P - \delta_A) + \delta_A}$ (see Lemma 3) and identifies the steady state action.

of simplicity, we will refer to actions as d , surpluses as $\pi(d)$, and we will, without loss of generality, disallow public randomization.

The set of contract payoffs is compact. Any continuation contract of a Pareto-optimal contract must be Pareto-optimal. Thus, when dealing with Pareto-optimal contracts, we may write $(U_{A,t}, V(U_{A,t}))$ for $(U_{A,t}, U_{P,t})$. From now on, we will refer to the Pareto frontier as V and Pareto-optimal contracts as V -contracts.

Lemma 2. $V(U_A)$ is a concave, strictly decreasing function over its domain $[U_A^{\min}, U_A^{\max}]$ and satisfies $V(U_A^{\max}) = O_P$.

Proof. Concavity follows from concavity of $\pi(d)$. Suppose $V(U_A^{\max}) > O_P$. Take the V -contract that delivers U_A^{\max} and increase the initial transfer by $V(U_A^{\max}) - O_P$. This contract still satisfies (1) and (2) and delivers $> U_A^{\max}$ payoff to the agent. Contradiction. \square

Assumption 1. O_P is sufficiently low.⁷

Paired with the lemma below, this assumption highlights the fact that even though the surplus maximizing action d^{\max} is sustainable, dynamic trading gains may cause the principal and agent to prefer a steady state with a lower static surplus.

⁷A sufficient upper bound is $O_P \leq \left(\pi(d_{\max}) - \frac{1 - \delta_A}{\theta(1 - \delta_A) + \delta_A} d_{\max} \right) / (1 - \delta_P)$. The right hand side is the principal's payoff under the unique stationary contract that sustains the surplus maximizing action with the smallest possible stationary transfer to the agent.

Lemma 3. *A stationary contract $\{(u_{A,t} = u_A^s, d_t = d^s)\}_{t=0}^\infty$ is a V -contract if and only if*

$$d^s \in \arg \max_{d \in \mathcal{A}} \pi(d) - \frac{\delta_P - \delta_A}{\theta(\delta_P - \delta_A) + \delta_A} d. \quad (7)$$

$$u_A^s = \frac{d^s}{\theta + \frac{\delta_A}{1 - \delta_A}}. \quad (8)$$

In particular, d^s is generically smaller than d_{\max} .

Lemma 3 characterizes the steady state of V -contracts. The reason d^s is generically smaller than d_{\max} is due to an important tradeoff between dynamic trading gains and static surplus: When the agent is more impatient, shifting the steady state payoff allocation in favor of the principal allows for larger initial transfers to the agent. This frontloading realizes potential gains from trading across time. The tradeoff is that shifting the steady state in favor of the principal tightens IC constraints. Since IC constraints were already binding to begin with, a concomitant decrease in the steady state action d^s , which loosens IC constraints, is required. This leads to a smaller static surplus.⁸ The optimal degree of the tradeoff is parameterized by the coefficient $\frac{\delta_P - \delta_A}{\theta(\delta_P - \delta_A) + \delta_A}$ in (7), which is proportional to the relative impatience of the agent. When the agent is as patient as the principal, $\delta_A = \delta_P$ and d^s simply maximizes surplus. As δ_A decreases, the relative value of dynamic trading gains increases and the steady state action and surplus decrease relative to the static optimum.

Proof. The IC constraint requires $\delta_A u_A^s / (1 - \delta_A) \geq -\theta u_A^s + d^s$ or equivalently $u^s \geq d^s / (\theta + \delta_A / (1 - \delta_A))$. If (8) did not hold, then the IC constraint would be slack each date. One can then easily achieve a Pareto-improvement by increasing $u_{A,0}$ slightly and decreasing $u_{A,1}$ slightly. Contradiction. This proves (8). To prove (7), fix a generic V -contract $\{(u_{A,t}, d_t)\}_{t=0}^\infty$ and consider two perturbations. First, for a small real ϵ , let $u_{A,0} \rightarrow \hat{u}_{A,0} := u_{A,0} + \epsilon$ and starting at date 1, enact the V -contract with agent payoff $\hat{U}_{A,1} := U_{A,1} - \theta\epsilon/\delta_A$. This perturbation is incentive-compatible. The agent's payoff is $U_{A,0} + (1 - \theta)\epsilon$. By definition, the principal's payoff must be weakly smaller than his payoff under the V -contract with the same agent payoff as the perturbation contract:

$$\pi(d_0) - u_{A,0} - \epsilon + \delta_P V\left(U_{A,1} - \frac{\theta\epsilon}{\delta_A}\right) \leq V(U_{A,0} + (1 - \theta)\epsilon)$$

Letting ϵ be infinitesimally positive and negative, we derive the two fundamental differential conditions linking the payoff U_A and continuation payoff U_A^+ of any V -contract:

$$(1 - \theta)V^+(U_A) \geq -1 + \frac{\delta_P}{\delta_A} \cdot \theta \cdot V^-(U_A^+) \quad (9)$$

$$(1 - \theta)V^-(U_A) \geq -1 + \frac{\delta_P}{\delta_A} \cdot \theta \cdot V^+(U_A^+) \quad (10)$$

⁸See Acemoglu, Golosov, and Tsyvinski (2008), Aguiar, Amador, and Gopinath (2009), and Opp (2012) for examples on investment distortions with heterogeneous discounting. We contribute relative to these papers by highlighting the efficiency of such distortions and being able to characterize the solution for arbitrary action sets.

(9) and (10) provide a useful necessary condition for when a Pareto-optimal payoff $(U_A, V(U_A))$ can be achieved by a stationary contract:

$$-V^-(U_A) \leq \frac{\delta_A}{\theta\delta_P + (1-\theta)\delta_A} \leq -V^+(U_A) \quad (11)$$

In the second perturbation, for a small real ϵ , let $d_0 \rightarrow \hat{d}_0 := d_0 + \epsilon$ and let $u_{A,0} \rightarrow \hat{u}_{A,0} := u_{A,0} + \epsilon/\theta$. Using arguments similar to before, we can establish:

$$1 - \theta\pi^-(d_0) \leq -V^-(U_A) \leq -V^+(U_A) \leq 1 - \theta\pi^+(d_0) \quad (12)$$

(11) and (12) together imply that if (u^s, d^s) is a steady state then

$$\pi^+(d_s) \leq \frac{\delta_P - \delta_A}{\theta\delta_P + (1-\theta)\delta_A} \leq \pi^-(d_s) \quad (13)$$

This proves the only if direction of (7). If $\pi(d)$ is strictly convex then we're done. Otherwise, there may be multiple solutions to (7). Let $\{(u^s, d^s)\}_{t=0}^\infty$ and $\{(\hat{u}^s, \hat{d}^s)\}_{t=0}^\infty$ be two steady states satisfying (7) and (8). Then the slope between the two steady states is $\delta_A/(\theta\delta_P + (1-\theta)\delta_A)$. (11) now implies that both are V -contracts. To complete the proof, it suffices to show that any stationary contract $\{(u_{A,t} = u_A^s, d_t = d^s)\}_{t=0}^\infty$ satisfying (7) and (8) must satisfy the principal's interim participation constraint (2). This is true by Assumption 1. \square

Lemma 3 establishes the first part of Theorem 1. It implies that the steady state is unique if there is a unique maximizer of (7), which is true outside of the knife-edge case when $\pi(d)$ has an entire edge with slope exactly equal to $(\delta_P - \delta_A)/(\theta\delta_P + (1-\theta)\delta_A)$. In the latter half of the analysis, we will make an assumption that eliminates the knife-edge case. So from now on we will refer to a unique steady state. While Lemma 3 itself characterizes the steady state, the proof of Lemma 3 contains all the technical ingredients needed to show when and how non-steady state V -contracts oscillate around the steady state. This will establish the second half of Theorem 1.

We begin the analysis by first supposing that the only available action is the steady state action d^s . This is exactly the premise of our earlier primer on oscillation in Section 2. In that analysis, we argued that since IC constraints must bind, the continuation payoff $U_A^+(U_A)$ of an above (below) steady state payoff U_A must be weakly below (above) the steady state, resulting in oscillation. The precise relation between the two values is:

$$U_A^+(U_A) - U_A^s = -(1+r)(U_A - U_A^s) \quad (14)$$

where U_A^s is the steady state agent payoff and $r := (\theta\frac{1+\delta_A}{\delta_A} - 1)/(1-\theta)$ is the growth rate of oscillation for the continuation payoff process. Per-period transfers u_A oscillate around the steady state value u_A^s analogously.

Definition 1. *For the rest of the paper, we will call these contracts described in the primer the benchmark contracts. The benchmark contracts keep the action fixed at d^s and set transfers to maximally exploit dynamic trading gains by keeping the IC constraint binding at all times.*

We now show that when $\theta \leq \frac{\delta_A}{1+\delta_A}$ and the growth rate r is nonpositive, V -contracts are essentially the benchmark contracts. But when $\theta > \frac{\delta_A}{1+\delta_A}$ and the growth rate r is positive, the benchmark contracts violate participation constraints and we explain how the true V -contracts become distorted. In particular, the action sequence also oscillates (around d^s).

Case $\theta \in [0, \frac{\delta_A}{1+\delta_A}]$ Let U_A be any payoff $\in I := [(U_A^+)^{-1}(U_A^{\max}), U_A^{\max}]$. Because the growth rate r is nonpositive, the benchmark contract delivering payoff (U_A, U_P) does not violate participation constraints and is therefore feasible. The resulting payoff frontier is linear and goes through the steady state payoff point:

$$\frac{U_P - U_P^s}{U_A - U_A^s} = -\frac{\delta_A}{\theta\delta_P + (1-\theta)\delta_A} \quad \forall U_A \in I \quad (15)$$

Since V is weakly concave, a linear frontier is unimprovable and therefore each benchmark contract is a V -contract: $V(U_A) = U_P$ for $U_A \in I$.

In the special case $\theta = 0$, the growth rate is -1 and $U_A^+(U_A) = U_A^s$ for all U_A . Thus, I is the entire domain $(-\infty, U_A^{\max}]$ and every V -contract is a benchmark contract: To deliver U_A , the V -contract always calls for the steady state action d^s . The agent receives an initial transfer $u_A = u_A^s + U_A - U_A^s$ and then receives the steady state transfer u_A^s forever.

If $\theta > 0$, then it is possible that $U_A^{\min} < (U_A^+)^{-1}(U_A^{\max})$. For a $U_A \in [U_A^{\min}, (U_A^+)^{-1}(U_A^{\max})]$, the benchmark contract would violate the participation constraint of the principal in the next period. Since the agent continuation payoff is capped by U_A^{\max} , the date 0 action must be distorted downward to $d < d^s$, just enough to maintain incentive compatibility. From the next period onwards, transfers resume oscillating according to (14). Thus, when $r \leq 0$, participation constraints induce only *mild* distortions of V -contracts: by date 1 at the latest, a V -contract becomes a benchmark contract.

When $\theta \in (0, \frac{\delta_A}{1+\delta_A})$, $r < 0$ and the oscillations dampen. Every V -contract converges to the steady state. When $\theta = \frac{\delta_A}{1+\delta_A}$, $r = 0$ and the oscillations persist in the long-run. This establishes Theorem 1 up through $\theta = \frac{\delta_A}{1+\delta_A}$. Next we consider:

Case $\theta > \frac{\delta_A}{1+\delta_A}$ To see how things change, let us start at the agent's maximum payoff $U_{A,0} = U_A^{\max}$. Suppose the principal still tries to use the benchmark contract. Since oscillation is growing, the principal realizes that this contract will violate his own participation constraint the day after tomorrow. Thus, the agent's date 2 payoff must be adjusted downwards. But this violates the IC constraint at date 1. To restore incentive-compatibility at date 1, the principal can either decrease the action $d_1 < d^s$ or increase the agent's date 1 continuation payoff, so that

$$U_{A,1} = U_A^+(U_A^{\max}) > U_A^s - \frac{1}{1+r}(U_A^{\max} - U_A^s)$$

Intuitively, it is optimal to do a little bit of both. But now the date 0 IC constraint is slack and so the principal can increase the date 0 action $d_0 > d^s$ and reap the extra surplus.

We have now shown $d_0 < d^s < d_1$. The degree to which actions are distorted depends on the degree to which participation constraints negatively impact the principal's payoff. Formally, the relationship is captured by (12) which relates the action to the slope of V . To ease the exposition and give contracts the ability to fine-tune action distortions, we assume the following differentiability condition:

Assumption 2. $\pi(d)$ is a strictly concave, continuously differentiable function tracing out slopes $\pi'(d) \in [0, \frac{1}{\theta}]$.

Assumption 2 simplifies the exposition substantially by ensuring differentiability of the value function with $V'(U_A)$ ranging from 0 down to -1 . Also, the cumbersome directional derivative inequalities in the proof of Lemma 3 become simple derivative equations. In particular, (12) simplifies to

$$\pi'(d) = \frac{1 + V'(U_A)}{\theta}$$

Since V is concave, (12) implies that action d and surplus $\pi(d)$ are weakly increasing in U_A , formalizing the intuition that distortions from the benchmark are optimally stronger the further the distance from the steady state (with $V'(U_A^s) = -\frac{\delta_A}{\theta\delta_P + (1-\theta)\delta_A}$). Combined with (10) and (9), (12) also reveals how an action distortion today is optimally balanced against an opposing distortion tomorrow:⁹

$$\delta_P [\pi'(d^+) - \pi'(d^s)] = -\frac{1}{1+r} [\pi'(d) - \pi'(d^s)] \quad (16)$$

The impact of tomorrow's distortion is naturally discounted by the principal's time-preference while the impact of today's distortion is discounted by the growth rate. Intuitively, the higher the growth rate r of oscillation the more severe the distortions imposed by participation constraints in the future. As a result, the optimal action adjusts more today.

As long as $d < d_{\max}$, the associated optimal transfer sequence can be obtained from the action sequence and binding IC. The resulting implications of the action adjustments are separately analyzed in two subcases.

When $\theta \in \left(\frac{\delta_A}{1+\delta_A}, \frac{\delta_A}{\delta_A+\delta_P}\right)$, action adjustments today are still relatively small compared to tomorrow's adjustments. As a result, transfers and action distortions explosively oscillate according to (16). The participation constraint of the principal will be reached in a finite number of periods. From then on, d and u_A perpetually oscillate between two distinct points, exhibiting long-run fluctuations.

When $\theta > \frac{\delta_A}{\delta_A+\delta_P}$, so that $\delta_P > \frac{1}{1+r}$, today's adjustment become so strong compared to tomorrow that action distortions and transfers damped oscillate and hence converge to the steady state. The economic environment exhibits no long run fluctuations. In the limit, as $\theta \rightarrow 1$ ($r \rightarrow \infty$), the damped oscillation of V -contracts becomes trivial and there is monotonic convergence to the steady state. This establishes the second half of Theorem 1.

⁹This first-order condition holds if the participation constraint does not bind in the subsequent period.

A Final Remark With the proof of Theorem 1 completed, we have now shown that the impatience versus incentives conflict causes V -contracts to oscillate around the steady state. The one requirement is that neither the agent nor the principal’s participation constraint binds at the steady state - otherwise, the steady state would be at a corner of V and there would be no room to oscillate. We made sure that this requirement was satisfied in the analysis by assuming that there was no agent participation constraint and that the principal’s participation constraint was sufficiently low (Assumption 1). In general, the analysis goes through if both participation constraints are present but don’t bind at the steady state. But what if one of them binds? In this case, it is easy to show that when the agent’s (principal’s) participation constraint binds, all V -contracts monotonically converge to the steady state which is at the left (right) corner of V .

Fix a model where the steady state is not at a corner of V , and consider an alternate version that shifts the d function up or down by a constant: $\hat{d}(a) := d(a) + x$ where $x \in \mathbb{R}$. As x increases, the deviation payoffs increase and one can interpret the incentive force as *getting stronger*. Similarly, as x decreases, the incentive force is *getting weaker*. Lemma 3 implies that a non-corner steady state shifts proportionally with x , which means there exists a bound $\bar{x} > 0$ ($\underline{x} < 0$) such that for all $x \geq \bar{x}$ ($x \leq \underline{x}$), the principal’s (agent’s) participation constraint will bind at the steady state. Combining this observation with the monotone convergence result of the previous paragraph, we now have a complete picture of the impatience versus incentives conflict:

Remark 2. *When the impatience force dominates the incentives force ($x \leq \underline{x}$), the steady state is the leftmost V -contract and all other V -contracts are frontloaded, monotonically converging leftwards to the steady state. When the incentives force dominates the impatience force ($x \geq \bar{x}$), the steady state is the rightmost V -contract and all other V -contracts are backloaded, monotonically converging rightwards to the steady state. When neither force dominates ($\underline{x} < x < \bar{x}$), the impatience vs incentives conflict is nontrivial, and oscillation around the steady state is a generic feature of V -contracts.*

Remark 2 helps put into context the opposing predictions of Ray (2002) and Lehrer and Pauzner (1999). As Ray (2002) points out, when “the agent is more impatient than the principal... the Lehrer-Pauzner findings and the results of [Ray (2002)] tug in different directions. It may be worth exploring if one of the two factors always dominates.” Our results not only show what happens when one factor dominates but also reveal that oftentimes neither factor dominates and oscillation is the natural outcome.

References

- ACEMOGLU, D., M. GOLOSOV, AND A. TSYVINSKI (2008): “Political economy of mechanisms,” *Econometrica*, 76(3), 619–641.
- AGUIAR, M., M. AMADOR, AND G. GOPINATH (2009): “Investment cycles and sovereign debt overhang,” *Review of Economic Studies*, 76(1), 1–31.
- ALBUQUERQUE, R., AND H. A. HOPENHAYN (2004): “Optimal lending contracts and firm dynamics,” *Review of Economic Studies*, 71(2), 285–315.

- BECKER, G. S., AND G. J. STIGLER (1974): “Law enforcement, malfeasance, and compensation of enforcers,” *The Journal of Legal Studies*, 3(1), 1–18.
- BRUNNERMEIER, M. K., AND L. H. PEDERSEN (2009): “Market Liquidity and Funding Liquidity,” *Review of Financial Studies*, 22(6), 2201–2238.
- CALOMIRIS, C. W., AND C. M. KAHN (1991): “The Role of Demandable Debt in Structuring Optimal Banking Arrangements,” *The American Economic Review*, 81(3), pp. 497–513.
- CLEMENTI, G. L., AND H. A. HOPENHAYN (2006): “A Theory of Financing Constraints and Firm Dynamics,” *The Quarterly Journal of Economics*, 121(1), 229–265.
- GEANAKOPOLOS, J. (2009): “The Leverage Cycle,” Cowles Foundation Discussion Papers 1715, Cowles Foundation for Research in Economics, Yale University.
- GU, C., F. MATTESINI, C. MONNET, AND R. WRIGHT (2013): “Endogenous Credit Cycles,” *Journal of Political Economy*, 121(5), 940 – 965.
- HARRIS, M., AND B. HOLMSTROM (1982): “A theory of wage dynamics,” *Review of Economic Studies*, 49(3), 315–333.
- KIYOTAKI, N., AND J. MOORE (1997): “Credit Cycles,” *Journal of Political Economy*, 105(2), 211–48.
- LEHRER, E., AND A. PAUZNER (1999): “Repeated games with differential time preferences,” *Econometrica*, 67(2), 393–412.
- OPP, M. M. (2012): “Expropriation risk and technology,” *Journal of Financial Economics*, 103, 113–129.
- RAY, D. (2002): “The time structure of self-enforcing agreements,” *Econometrica*, 70(2), 547–582.
- ROCHETEAU, G., AND R. WRIGHT (2013): “Liquidity and asset-market dynamics,” *Journal of Monetary Economics*, 60(2), 275–294.
- THOMAS, J., AND T. WORRALL (1988): “Self-enforcing wage contracts,” *Review of Economic Studies*, 55(4), 541–54.
- (1994): “Foreign direct investment and the risk of expropriation,” *Review of Economic Studies*, 61(1), 81–108.
- ZHU, Y., R. WRIGHT, AND C. HE (2013): “Housing and Liquidity,” Discussion paper, University of Wisconsin, Madison.

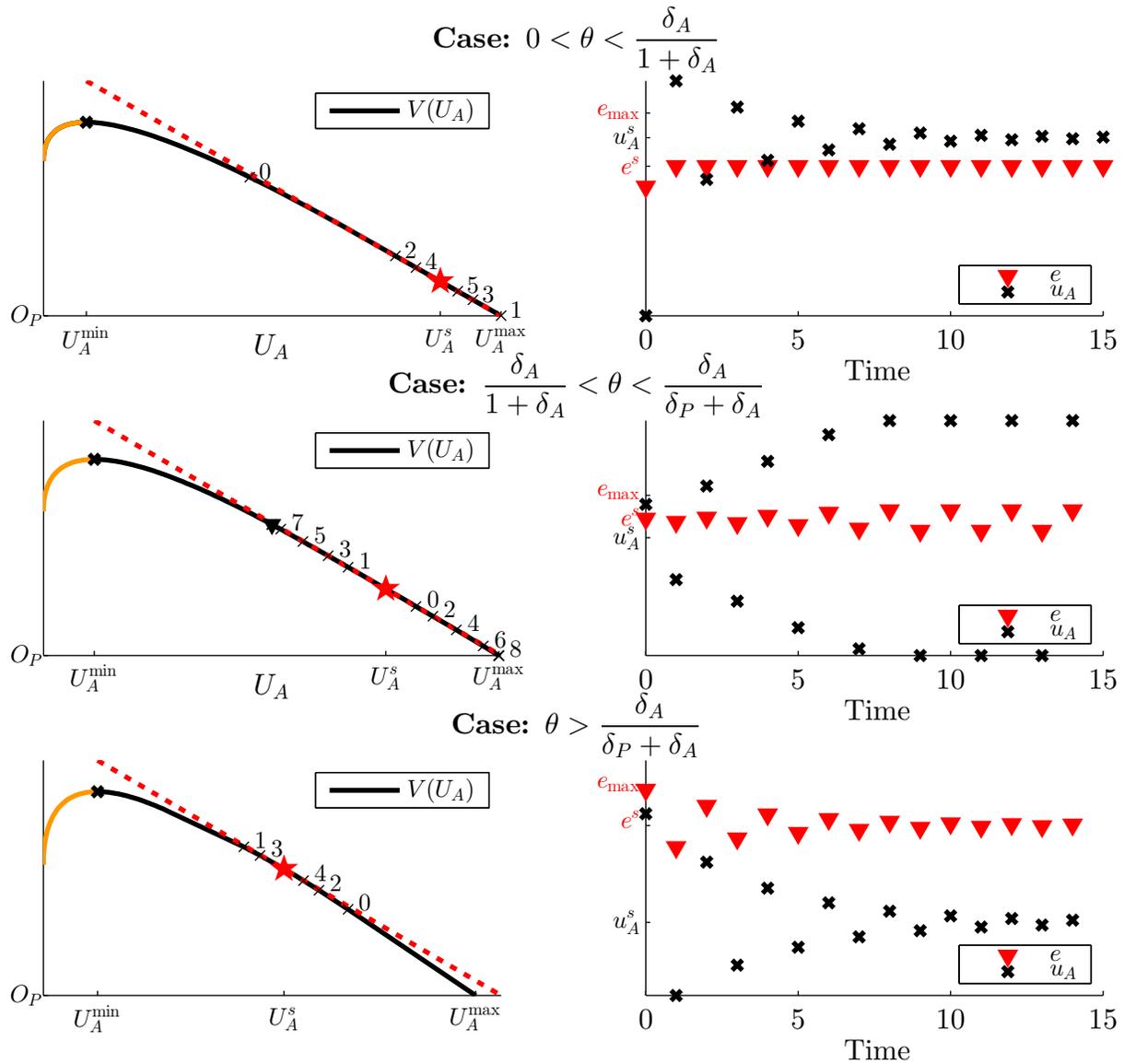


Figure A.1: The graph plots the respective Pareto frontier (left panel) as well as an optimal sequence of transfers/actions (right panel) for the three relevant parameter regions of θ (see Theorem 1). The computation is based on Example 3. The action only represents the effort of the worker, $e \in [0, 1]$, which generates surplus $\pi(e) = 2e - e^2$. We vary the fraction of the wage, θ , that the worker loses upon shirking.