Renegotiation and Dynamic Inconsistency: Contracting with Non-Exponential Discounting*

Doruk Cetemen† Felix Zhiyu Feng‡ Can Urgun§

September 2022

Abstract

This paper studies a continuous-time, finite-horizon contracting problem with renegotiation and dynamic inconsistency arising from non-exponential discounting. The problem is formulated as a dynamic game played among the agent, the principal and their respective future "selves", each with their own discount function. We identify the principal optimal renegotiation-proof contract as a Markov perfect equilibrium (MPE) of the game, prove that such an MPE exists, and characterize the optimal contract via an extended Hamilton-Jacobi-Bellman system. We solve the optimal contract in closed-form when discounting is a function of the time-difference only and demonstrate the applicability of the results in several different settings.

Keywords: continuous-time contracting, dynamic inconsistency, renegotiation, extended HJB system, non-atomic games

JEL code: D82, D86, D91

*Previously titled: *Contracting with Non-Exponential Discounting: Moral Hazard and Dynamic Inconsistency*. We are grateful to Marzena Rostek (the editor), an anonymous associate editor, and three anonymous referees for their insightful suggestions. We thank Paulo Barelli, Jess Benhabib, Alessandro Bonatti, Gonzalo Cisternas, Dino Gerardi, Hari Govindan, Johannes Hörner, Ryota Iijima, Rohit Lamba, Erik Madsen, Roberto Marfè, Stephen Morris, Pietro Ortoleva, Romans Pancs, Wolfgang Pesendorfer, Debraj Ray, Bruno Strulovici, Juuso Toikka, and Leeat Yariv for useful discussions, and seminar participants at Bilkent University, Collegio Carlo Alberto, London School of Economics, New York University, Princeton University, Queen Mary University of London, University of Bristol, University of Warwick, Yale University, conference participants at ESSET(Gerzensee) 2019, NASMES 2019, SAET 2019, and SITE 2019 conference on Dynamic Games, Contracts and Markets for helpful comments. Part of this research was conducted while Doruk Cetemen was affiliated with Collegio Carlo Alberto.

†City, University of London. Email: doruk.cetemen@city.ac.uk

‡University of Washington. Email: ffeng@uw.edu

§Princeton University. Email: curgun@princeton.edu

1 Introduction

Dynamic contracting explores how two parties can use inter-temporal incentives to mitigate agency frictions. Most existing contracting models assume that inter-temporal preferences take the special form of exponential discounting because of its technical conveniences, such as dynamic consistency and stationarity. However, there is extensive evidence – both anecdotal and empirical – that "time preferences can be non-exponential" (Laibson, 1997). Laboratory and field studies find that discount rates are much greater in the short-run than in the long-run (Harris and Laibson, 2012). When a group of decision makers (such as a board of directors, team, or committee of experts) follow unanimous decisions, they can collectively exhibit non-exponential, present-biased behaviors even when each individual follows exponential discounting (Jackson and Yariv, 2015). Therefore, it is natural to ask: how would dynamic contracting be different with non-exponential discounting?

This paper aims to provide a general framework to answer this question. We begin with a standard continuous-time dynamic moral hazard model in which a principal hires an agent to manage a project over a finite horizon. The project's outcome is noisy, and the agent controls its drift with private actions. We include three critical elements in this general framework: first, we allow the principal and the agent to have different, generic, time-varying, and non-exponential discount functions. Second, the principal and the agent are fully rational and sophisticated: they make decisions knowing exactly how their future preferences will change due to non-exponential discounting. Third, renegotiation is allowed as long as both parties can agree on the proposed change of terms. An example of this setting is the contracting between a CEO and a board of directors. At the onset of their relationship, the two parties consent to a long-term contract that provides the CEO with incentives to exert the effort that is beneficial to the firm. At any point in time, the board can renegotiate by offering a new contract that replaces the old one if the CEO agrees. If the CEO disagrees, the old contract stays in place, and the same protocol applies to any future possible renegotiation.

The importance of considering sophisticated contracting parties and renegotiation stems from the fact that non-exponential discounting generates *dynamic inconsistency*: a contract that appears "optimal" to the board or the manager given their current preferences may appear sub-optimal in the future, not because of any inefficient punishments or constraints, but simply due to changes in their discounting. If either the manager or the board is naive, or if renegotiation is not allowed, then one party may regret having previously agreed to the contracting terms at some point, calling into question whether the previously agreed-on contract was actually "optimal". Instead, we assume that both contracting parties are sophisticated and the contract can be renegotiated, and we provide an intuitive definition for a *renegotiation-proof, incentive-compatible, optimal long-term contract* under such a setting. In particular, the dynamic contracting relationship can be formulated as an inter- and intra-personal game, in which each player is a *time-t self* of the principal or the agent and the optimal, renegotiation-proof contract corresponds to an appropriately defined equilibrium of the game. By exploiting the link between dynamic contracting and static non-atomic games, we formally characterize the optimal, renegotiation-proof contract and prove its existence under fairly general conditions.

Our first finding is that under our framework, the dynamic inconsistency on the agent's side turns out to have much less impact than a time-inconsistent principal does. Despite the fact that the agent is also playing an intra-personal game with his future selves, his incentive compatibility condition is characterized by a local pay-performance sensitivity similar to that in time-consistent benchmarks (e.g., Sannikov, 2008). The principal only needs to provide incentives for the current agent and not his future selves. This provides a much-needed simplification, allowing us to take the promised equilibrium value to the agent as a state variable and focus on the effect of dynamic inconsistency on the principal's side.

With the agent's problem solved, we present our main theorem characterizing the incentive-compatible, renegotiation-proof, optimal long-term contract with as broad a generality as possible. The theorem includes two innovations. First, we demonstrate how to characterize the optimal contract using recursive techniques when the usual tool of Bellman equation is no longer applicable. We argue that given the agent's local incentive compatibility and our notion of renegotiation-proofness, the optimal contract is equivalent to the equilibrium of a simpler auxiliary game in which each t-self of the principal can only influence the agent's action, consumption, and the evolution of continuation utility at time t. The recursive characterization yields an *extended Hamilton-Jacobi-Bellman (HJB) system* with additional terms compared to the usual HJB equation for time-consistent benchmarks. These additional terms capture the equilibrium incentives of the principal, who takes into account the impact of the contract policies on the payoff for all her future selves.

The second innovation of our main theorem is the proof that such an optimal contract exists. This is known to be a thorny problem because the well-posedness of the extended HJB system is not well understood even in the mathematical literature.[1] We prove that such a system must have a solution by exploiting the unique connection between partial differential equations and static non-atomic games of incomplete information. This novel connection enables us to utilize known existence results in non-atomic games to establish

[1] Harris and Laibson (2012) and Bernheim, Ray, and Yeltekin (2015) show the existence of solutions for special cases only. Björk, Khapko, and Murgoci (2017) and Yong (2012) propose the extended HJB systems for more general cases but leave the crucial question of existence open and provide only a partial verification theorem.

the existence of a solution to our extended HJB system. We also show that any equilibrium of the principal's auxiliary game must also be a contract that solves the extended HJB system. That is, our characterization of the optimal contract is complete. The generality of our framework opens the door for a broad range of applications.

We demonstrate the applicability of our general framework by solving a more specific class of contracting problems. For simplicity, we assume the agent has constant-absolute-risk-aversion (CARA) utility and exponential discounting, and the set of discount functions of the principal belongs to a *time-difference family*: i.e., the discount factor between any two dates t, s is only a function of their difference $s - t$. This applies to some of the most commonly-studied examples of non-exponential discounting in economics, such as hyperbolic discounting and anticipatory utility. Importantly, we are able to obtain closed-form solutions for this class of simplified problems and highlight the specific effects brought by non-exponential discounting (on the principal's side). In particular, we illustrate a "deadline" effect and the resulting non-monotonic paths of equilibrium actions and incentives under *quasi-hyperbolic* discounting (as in Harris and Laibson, 2012).[2] We also demonstrate how these theoretical predictions can help reconcile the observed managerial compensation practices, such as the incentive power of the CEOs as they approach retirement.

Literature Review: This paper is broadly related to several strands of research. First, it belongs to the thriving literature of continuous-time dynamic contracting. Benchmark models with time-consistent preferences such as DeMarzo and Sannikov (2006), Biais, Mariotti, Plantin, and Rochet (2007), and Sannikov (2008) demonstrate the analytical convenience of a continuous-time formulation, which allows the derivation of economic insights from problems that are otherwise challenging to solve in discrete time.[3] Our paper differs in that we explicitly model dynamic inconsistency from non-exponential discounting for both the principal and the agent. Moreover, dynamic inconsistency necessitates a formal discussion of contract renegotiation, which is often ruled out or assumed away in benchmark models that assume full commitment power for the principal.

Our definition and analysis of renegotiation are related to both contract renegotiation

[2]More specifically, we show that under the optimal contract, the equilibrium paths of actions and contract incentives are monotonic if the principal is time-consistent or if the contracting horizon is infinite. However, if the principal has quasi-hyperbolic discounting and the contract horizon is finite, the equilibrium paths of actions and incentives will display a reversal near the end of the contracting horizon.

[3]The power of continuous-time tools has been demonstrated in other problems. For instance, Faingold and Sannikov (2011) analyze a continuous-time reputation model and are able to prove the uniqueness of the equilibrium. Georgiadis (2014) demonstrates the tractability of continuous-time tools in delivering sharp comparative statics in public good contribution games – an intractable problem in discrete-time. Cisternas (2018) expands the linear career concerns model of Holmström (1999) to a general non-linear setup in continuous-time. Recently, Kolb and Madsen (2019) show that a continuous-time setting enables a tractable analysis of both multiple actions and imperfect monitoring in information design problems.

in dynamic moral hazard problems as well as equilibrium renegotiation in dynamic games. In particular, we adopt the same procedure of renegotiation as that considered in Watson, Miller, and Olsen (2019): once a renegotiation is proposed, if that proposal is rejected, the previous contract remains in place. The setting of Watson, Miller, and Olsen (2019) is a more general principal-agent setup that incorporates an explicit renegotiation phase in which both parties have different bargaining power. In comparison, we make the simplifying assumption that the principal has full bargaining power. In that regard, our model also resembles Fudenberg and Tirole (1990), who explore contract renegotiation in a one-time setting with the principal having full bargaining power. In addition, our renegotiation-proof contract is characterized as a Markov perfect equilibrium, which is closely related to the notion of internal consistency in Bernheim and Ray (1989) for repeated games as well as Ray (1994) and Strulovici (2020) for dynamic games.

Our study is also related to contracting problems with behavioral preferences.[4] In particular, we assume the contracting parties are sophisticated regarding their time-inconsistency. This resembles Galperti (2015), which focuses on the optimal provision of commitment devices with sophisticated agents. In contrast, a substantial number of studies on behavioral preferences assume the contracting parties are naive or partially naive. For example, Gottlieb (2008) studies the optimal design of non-exclusive contracts and identifies different implications of immediate-cost goods and immediate-reward goods for dynamically inconsistent but naive consumers. Gottlieb and Zhang (2020) study repeated contracting between a risk neutral firm and dynamically inconsistent but partially naive consumers, and find that at-will terminations may improve welfare if the level of dynamic inconsistency is sufficiently high. These studies focus on adverse selection problems (reporting), while we focus on a moral hazard (hidden effort) problem. Moreover, DellaVigna and Malmendier (2004) analyze the optimal two-part tariff of a firm facing a partially naive consumer with present-biased preferences. Heidhues and Kőszegi (2010) study a similar setting with naive agents and show that simple restrictions on contract forms may significantly improve welfare.

In general, the idea that preferences may be dynamically inconsistent as the result of non-exponential discounting can be traced to early work of Strotz (1955) and Pollak (1968). Since then, many studies have explored dynamically inconsistent preferences in various settings: including consumption-saving problems (Laibson, 1997; Krusell and Smith, 2003; Bernheim, Ray, and Yeltekin, 2015; Ray, Vellodi, and Wang, 2017; Bond and Sigurdsson, 2018; Cao and Werning, 2018); investment and asset allocation (Caplin and Leahy, 2001; Grenadier and Wang, 2007; Brunnermeier, Papakonstantinou, and Parker, 2016); monetary policy (Kydland and Prescott, 1977); fiscal policy (Halac and Yared, 2014); procrastination (O'Donoghue

[4]See Kőszegi (2014) and Grubb (2015) for a survey of this topic.

and Rabin, 1999), public finance (Bisin, Lizzeri, and Yariv, 2015; Harstad, 2016), etc. These studies are typically limited to a single party being dynamically inconsistent. In particular, Li, Mu, and Yang (2016), Liu, Mu, and Yang (2017), Wang, Huang, Liu, and Zhang (2020) and Rivera (2022) expand standard dynamic moral hazard settings (DeMarzo and Sannikov, 2006; DeMarzo, Fishman, He, and Wang, 2012; and Sannikov, 2008, respectively) to explore the differences in the optimal contract if the agent has quasi-hyperbolic discounting (while the principal still has normal exponential discounting). Consistent with what we demonstrate in our paper, time-inconsistency does not significantly affect the solution to the agent's problem: the incentive compatibility conditions in these studies are identical to their respective benchmarks. In contrast, we focus on the case in which the principal is time-inconsistent. This makes the contracting problem significantly different, and consequently much more involved. One exception in this line of research is Chade, Prokopovych, and Smith (2008), who provide a recursive characterization of a repeated game in which all players share the same quasi-hyperbolic discount function. In comparison, we allow generic, non-exponential, and potentially different discounting for all players, including quasi-hyperbolic discounting as a special case.

Our paper also synthesizes and utilizes some of the latest advancements in mathematical finance research. We use the extended HJB system developed in Björk, Khapko, and Murgoci (2017) to characterize the equilibrium strategies in time-inconsistent problems. However, Björk, Khapko, and Murgoci (2017) studies a single-agent problem without moral hazard or contracting, and leaves the existence of the extended HJB system as an open question.[5] Our solution technique partly follows Yan and Yong (2019), which offers two different but related approaches: the *open-loop* strategy which stems from the stochastic maximum principle, and the *closed-loop* strategy which discretizes time into a mesh and defines the equilibrium as the individual mesh size goes to zero. The former approach yields solutions with known properties such as existence but are difficult to interpret or directly utilize. The latter approach is more natural in most economics settings and can lead to HJB equations with Markov controls. However, Yan and Yong (2019) notes that when the diffusion terms are also part of the controls, as in the case of dynamic contracting problems, the limit of such HJB equations as the mesh size goes to zero is not necessarily well posed. Thus, the existence of the solution is not always guaranteed.

[5]The more recent study of Lindensjö (2019) extends Björk, Khapko, and Murgoci (2017) by showing that the solution to the extended HJB has certain regularity (smoothness) properties, *if* such a solution exists. Meanwhile, He and Jiang (2019) strengthens the equilibrium notion in Björk et al. (2017) to a *regular* one that requires each self in the intra-personal game to behaves optimally in every conceivable infinitesimal length of time. We adopt the original definition of equilibrium in Björk et al. (2017) but show that, under a non-restrictive assumption on the utility function, our notion of the equilibrium also conforms to the regular equilibrium defined in He and Jiang (2019).

Our paper is the first that integrates these individual results into a principal-agent model and solves a general renegotiation-proof optimal contract with non-exponential discounting. We also expand this literature with novel analytical results that are potentially applicable to future studies of related topics. First, we demonstrate how the extended HJB system in Björk, Khapko, and Murgoci (2017) and the solution techniques of Yan and Yong (2019) can be applied when the diffusion terms of the state variables (from the solutions to the agent's problem) are also part of the (principal's) controls.[6] Second, we prove the existence of the solution to the extended HJB system by exploiting the connection between differential equations and non-atomic games of incomplete information. In fact, the novelty of this proof of existence is independent of the specific contractual setting used in the paper and can thus be applied to a broad class of problems involving dynamic inconsistency and intra-personal games. Finally, we demonstrate the applicability of our methodology in economics by explicitly solving a special class of non-exponential discounting functions commonly used in economic research. In particular, our examples include quasi-hyperbolic discounting, which has been used as the pre-requisite of many of the aforementioned studies. Our paper thus provides both convenient results for researchers wishing to adopt quasi-hyperbolic discounting as one of their model elements and a general framework for those interested in exploring the impact of non-exponential discounting beyond quasi-hyperbolic discounting.

2 General Framework

In this section, we present the general framework which introduces generic, non-exponential discounting functions into an otherwise standard dynamic principal-agent model. We lay out the foundations, solve the agent's problem, and discuss the precise role of commitment and renegotiation in the model.

2.1 Basic Environment

A risk-neutral principal (she) contracts with a risk-averse agent (he) over a fixed-time horizon $T < \infty$.[7] Time is continuous and indexed by t. There is a probability space Ω equipped with a filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ and the associated measure $\mathbb{P}$. $\{Z_t^0\}_{t\in[0,T]}$ is a $\mathcal{F}_t$-measurable stochastic

[6] In doing so, we utilize the results in Brunick and Shreve (2013), also known as the Gyongy-Dupire formula pioneered by Gyöngy (1986) and Dupire et al. (1994), for the construction of diffusion processes out of adapted processes; and Wang and Yong (2019b), for translating the BSVIE resulting from the stochastic maximum principle to a diffusion inside a PDE.

[7] The assumption of a finite contracting horizon offers several critical analytical conveniences. It allows us to represent continuation utility processes as solution to a series of backward-stochastic-differential equations (BSDE) with known regularity properties, such as the smoothness.

process representing the Brownian motion under $\mathbb{P}$. The agent affects the monetary output M_t by selecting a hidden action $\hat{a}_t$ (i.e., his effort) from a compact set $[\underline{a}, \bar{a}]$ at each moment in time, such that M_t is characterized by a drift plus a Brownian motion $\{Z_t^{\hat{a}}\}_{t\in[0,T]}$ under the measure $\mathbb{P}^{\hat{a}}$. That is,

$$dM_t = \hat{a}_t dt + \sigma dZ_t^{\hat{a}} \ , \tag{1}$$

where $M_0 = 0$, and $Z_t^{\hat{a}} = Z_t^0 - \int_0^t \frac{\hat{a}_s}{\sigma} ds$ and $\sigma > 0$ is a known constant parameter. At $t = 0$, the principal and the agent have outside options worth $\underline{V}$ and $\underline{W}$, respectively, which they lose once the contract commences. The principal also has unlimited liability — she can tolerate arbitrarily large losses resulted from the Brownian shocks. She observes the output M_t, but not the agent's action $\hat{a}_t$.

Between time t and t' the principal's and the agent's discount rates are given by *self-indexed* functions $R^t(t')$ and $r^t(t')$, respectively. Such a specification allows for discounting to vary with both the current self t and an arbitrary future date t', not just the time difference. In particular, at time t, the principal uses the mapping defined by $R^t : [t, \infty) \to [0, 1]$ as the discount function. One can interpret $R^t(s)$ as if there are infinitely many discounting functions, one corresponding to each point t in time. Similarly, at time t, the agent uses the mapping defined by $r^t : [t, \infty) \to [0, 1]$. We assume that both the agent and the principal are sophisticated regarding their time preferences and the time preferences are common knowledge.

We impose the following conditions on $R^t(\cdot)$ and $r^t(\cdot)$:

Assumption 1 *For all $t \geq 0$, the discount functions $R^t(\cdot), r^t(\cdot)$ satisfy:*

1. $R^t(s), r^t(s) = 1$ *for all* $s \leq t$ *and* $\lim_{s\to\infty} R^t(s) = 0$*,* $\lim_{s\to\infty} r^t(s) = 0$*.*
2. $R^t(s), r^t(s) > 0$ *for all* $s > t$*.*
3. $\int_t^\infty R^t(s)ds < +\infty$ *and* $\int_t^\infty r^t(s)ds < +\infty$ *a.s.*
4. *Regarded as bi-functions of t and s, $R^t(s)$ and $r^t(s)$ are uniformly Lipschitz continuous and three times differentiable in both arguments. i.e., there exist constants $K_r, K_R > 0$ such that, for all $(s_1, t_1, s_2, t_2) \in [0, \infty)$ where $s_1 > t_1$ and $s_2 > t_2$,*

$$\begin{aligned} |r(t_1, s_1) - r(t_2, s_2)| &< K_r \left(|t_1 - t_2| + |s_1 - s_2|\right) \ , \\ |R(t_1, s_1) - R(t_2, s_2)| &< K_R \left(|t_1 - t_2| + |s_1 - s_2|\right) \ . \end{aligned}$$

The first condition states that the current payoff is not discounted, while the payoff in the infinitely far future has a present value of 0. The second condition states that any return in a finite future has some positive value, albeit potentially very small. The third condition ensures that the discounted value of any bounded stream of consumption remains finite. The last condition is a technical one precluding drastic changes and guaranteeing sufficient differentiability in the discount functions across and over time.

Overall, Assumption 1 describes a fairly broad class of discount functions (including those that are not monotonic). One subset of those is the *time-difference* family in which the discounting between any two dates s and t is a function of the difference $(s-t)$ only. This family is of particular interest because it incorporates both the standard exponential discounting as well as some of the special non-exponential discount functions most commonly used in the literature. We explore a few of those examples in Section 4. Assumption 1 can be narrowed down to the time-difference family by imposing the restriction $r^t(s) = r^{t+k}(s+k) = r(s-t)$ and $R^t(s) = R^{t+k}(s+k) = R(s-t)$ for all t, s, k. In such a case, the term "discount functions" refers to $r(s-t)$ and $R(s-t)$.

The agent's utility depends on his action $\hat{a}$ and consumption c received from the contract. We assume his utility function has the following properties:

Assumption 2 *The agent has a weakly risk averse utility function of instantaneous consumption and action, $u(c, \hat{a})$, such that*

1. *u is strictly convex and decreasing in $\hat{a}$, concave and increasing in c.*
2. *u is twice differentiable with bounded derivatives.*
3. *$\lim_{c\to\infty} u(c, \hat{a}) = \infty$ and $\lim_{c\to-\infty} u(c, \hat{a}) = -\infty$ for all $\hat{a} \in [\underline{a}, \bar{a}]$.*

The first two requirements are fairly standard. The last is a technical one facilitating the proof of the general existence theorem (Theorem 1) and is only sufficient but not necessary. For special cases in which more specific properties of the other elements of the model are imposed (such as those analyzed in Section 4), our results can be derived under more generalized utility functions.

Since the contracting horizon is finite, we also assume that the agent derives terminal utility $U(c_T)$ from any terminal consumption c_T. $U(\cdot)$ has the standard properties such as being increasing, concave, and unbounded from below. The agent is not allowed to save, unless otherwise stated.

A contract $\mathcal{C}$ specifies a consumption process $\{c_t\}_{t\in[0,T]}$ to the agent and a sequence of recommended actions $\{a_t\}_{t\in[0,T]}$, both measurable with respect to $\mathbb{P}^{\hat{a}}$. Given the contract, the agent's time-0 expected payoff, if he follows the recommended actions, is

$$E^{\mathcal{C}}\left[\int_0^T r^0(s)u(c_s, a_s)ds + r^0(T)U(c_T)\right].$$

The principal's time-0 payoff equals the expected cash flows net of the consumption to the agent:

$$E^{\mathcal{C}}\left[\int_0^T R^0(s)\left(dM_s - c_s ds\right) - R^0(T)c_T\right] = E^{\mathcal{C}}\left[\int_0^T R^0(s)\left(a_s - c_s\right)ds - R^0(T)c_T\right].$$

Because both the agent and the principal can have time-inconsistent preferences, for each point in time $t > 0$, we must treat the principal and the agent as different players (selves), to whom we refer as the "agent's t-self" and the "principal's t-self". Given any contract $\mathcal{C}$ and the agent's action stream $\{\hat{a}_t\}_{t\in[0,T]}$, the expected discounted utility for the agent's t-self at any time $k \geq t$ is defined as

$$\tilde{W}(k,t) = E_t^{\mathcal{C}}\left[\int_k^T r^t(s)u(c_s,\hat{a}_s)ds + r^t(T)U(c_T)\right], \quad \forall k \geq t.$$

In particular, $\tilde{W}(t,t)$ is simply known as *agent-t's continuation utility*, where the first t denotes the time and the second t denotes the self. The conditional expectation (both here and later) is taken under $\mathcal{F}_t$ unless otherwise explicitly stated.

Using these notations we can define an incentive compatible contract as follows:

Definition 1 *A contract $\mathcal{C} = \{a_t, c_t\}_{t\in[0,T]}$ is incentive compatible if at each instant t, the effort choice of agent's t-self is the action suggested by the principal, (i.e., $\hat{a}_t = a_t$) and maximizes $\tilde{W}(t,t)$ assuming that all other selves of the agent follow the actions recommended by the contract $\mathcal{C}$.*

When a contract $\mathcal{C}$ is incentive compatible, we use

$$f(k,t) = E_t^{\mathcal{C}}\left[\int_k^T R^t(s)\left(a_s - c_s\right)ds - R^t(T)c_T\right], \quad \forall k \geq t.$$

to denote the expected discounted payoff of the principal's t-self at any time $k \geq t$, where $f(t,t)$ is simply *principal-t's continuation payoff.*

2.2 The Agent's Problem

We first solve the agent's problem for a fixed contract $\mathcal{C}$. In general, the agent's consumption at any time t could depend on the entire path of the outputs. We resolve this history dependence by taking the key state variable to be the probability density of the output process (M_t) instead of M_t itself. This so-called "change of measure" technique is a common approach in dynamic contracting, most notably used in studies that involve persistent private information, such as Williams (2011, 2015), He, Wei, Yu, and Gao (2017), Marinovic and Varas (2019), Bloedel, Krishna, and Strulovici (2020), Feng (2021), etc. Specifically, for any sequence of actions $\{\hat{a}_t\}_{t\in[0,T]}$, define the following family of $\mathcal{F}_t$ predictable processes:

$$\Gamma_t \equiv \exp\left(\int_0^t \frac{\hat{a}_s}{\sigma} dZ_s^0 - \frac{1}{2}\int_0^t \left|\frac{\hat{a}_s}{\sigma}\right|^2 ds\right),$$

with $\Gamma_0 = 1$. The process $\{\Gamma_t\}_{t\in[0,T]}$ captures the evolution of the probability density of the output when the agent takes private actions $\hat{a}$ and evolves according to

$$d\Gamma_t = \left(\frac{\hat{a}_t}{\sigma}\right)\Gamma_t dZ_t^0, \quad \Gamma_0 = 1, \tag{2}$$

i.e., Γ_t is a controlled Markov process. Under this change of measure (from $Z_t^{\hat{a}}$ to Z_t^0), the agent's payoff from taking any sequence of actions can be written as:

$$W(t,\Gamma,t) = E_t^{Z^0}\left[\int_t^T \Gamma_s(\hat{a}_s)r^t(s)u(c_s,\hat{a}_s)ds + \Gamma_T r^t(T)U(c_T)\right], \tag{3}$$

subject to (2).[8] Thus, given any contract, Γ_t is the relevant state variable in evaluating the continuation utility generated by the actions of the agent's each self at any given time.

We can now analyze the agent's optimal actions induced by a given contract. Because each self of the agent is infinitesimal and makes his own decision, all the agent's selves collectively play an *intra-personal game* with each other, taking the contract terms as given. The optimal actions of each player (each of the agent's selves) must constitute an equilibrium of the intra-personal game. In particular, we look for an equilibrium that involves a Markov strategy for each player that depends on (t,Γ_t), where Γ_t is the value of the Γ-process at time t.[9] Formally, such equilibrium is defined as follows:

Definition 2 *Consider a contract $\mathcal{C}$ and a Markov strategy profile $\{\hat{a}(t,\Gamma_t)\}_{t\in[0,T]}$ taken by the agent. For any initial point and state (t,Γ_t) and a small increment of time Δ, define a "deviation" strategy a_Δ as:*

$$a_\Delta = \begin{cases} \tilde{a} & \text{for } t \le t' < t+\Delta \\ \hat{a}(t',\Gamma_{t'}) & \text{for } t+\Delta \le t' \le T \end{cases},$$

[8] Recall that Z_t^0 is the Brownian motion under the measure $\mathbb{P}$ and $Z_t^{\hat{a}}$ is the Brownian motion under $\mathbb{P}^{\hat{a}}$, with $Z_t^{\hat{a}} = Z_t^0 - \int_0^t \frac{\hat{a}_s}{\sigma} ds$. Thus $E_t^{Z^0}$ represents the expectation taken under $\mathbb{P}$.

[9] If the agent is dynamically consistent, any adapted strategy of the agent can be likewise captured. The Markov restriction is only relevant due to the existence of different selves.

*where $\tilde{a} \in [\underline{a}, \bar{a}]$, is some action potentially different from the one dictated by the strategy $\hat{a}(t, \Gamma_t)$. Then $\hat{a}(t, \Gamma_t)$ is called an **agent equilibrium** if*

$$\liminf_{\Delta \to 0} \frac{W(t, \Gamma, t) - W_{a_\Delta}(t, \Gamma, t)}{\Delta} \geq 0 \tag{4}$$

*for all a_Δ, where W_{a_Δ} denotes the valuation from following a_Δ. Under an agent equilibrium, $W(t, t) \equiv W(t, \Gamma, t)$ is called an **equilibrium value process**.*

In words, for a given contract, each self of the agent at time t given the state Γ_t explores whether a deviation can be made for a small time interval Δ that yields a higher payoff for him compared to his valuation from following the strategy $\hat{a}(t, \Gamma_t)$. If no such profitable deviations can be found for each self when taking the limit $\Delta \to 0$, then the entire strategy profile of all agent's selves is called an intra-personal equilibrium. Based on this notion of the equilibrium and the equilibrium value process, the following lemma characterizes the agent's continuation utility and the incentive compatibility condition for a given contract:

Lemma 1 *Under Assumptions 1 and 2, given any contract $\mathcal{C} = \{a_t, c_t\}_{t \in [0,T]}$ and any sequence of the agent's choices, there exists a flow of processes $\tilde{\psi}(s, t)$ and an equilibrium value process W such that*

$$\tilde{W}(k, t) = r^t(T)U(c_T) - \int_k^T r^t(s)u(c_s, \hat{a}_s)ds + \int_k^T \tilde{\psi}(s, t)dZ_s^{\hat{a}}, \tag{5}$$

for each t-self of the agent and each $k > t$. The equilibrium value process satisfies

$$W(t) = \tilde{W}(t, t) \ . \tag{6}$$

The contract $\mathcal{C}$ is incentive compatible if and only if

$$\psi_t = \frac{\tilde{\psi}(t, t)}{\sigma} = u_a(c_t, a_t). \tag{IC}$$

Compared to benchmark models with time-consistent preferences (e.g., Sannikov, 2008), the implications of Lemma 1 are two-fold: on the one hand, unlike the benchmark, the time-inconsistent preference implies that the agent's continuation utility becomes a *flow* of processes, one for each t-self of the agent. $\tilde{W}(k, t)$ represents the value of continuation from period k onward for the t-self. Setting $k = t$ yields $W(t)$, the continuation utility of the t-self agent, taking into account not only the change in time but also the change in the agent's future preferences. Consequently, $W_t \in \mathbb{R}$ which is the realization of $W(t)$ at time t is the relevant state variable for the principal when designing the optimal contract. On the other hand, despite the existence of future selves, the IC condition (IC) is a local constraint similar to the time-consistent benchmark. This is because the agent's action does not have a persistent effect, and each self wants to maximize his own utility.

Remark 1 *Broadly speaking, there are two types of solution strategies to a general stochastic control problem: open-loop, and closed-loop, with the former (which is only progressively measurable) being a superset of the latter (Markov) controls (see Sun, Li, and Yong, 2016 and Zhang and Li, 2018 for discussions). Open-loop controls are widely used in time-inconsistent problems as the existence of equilibrium is relatively straightforward.*[10] *As Lemma 1 shows, the agent's continuation utility in our setting is summarized by a Backward Stochastic Volterra Integral Equation (BSVIE), which captures a flow of continuation payoffs with flows of sensitivities to the Brownian motion which follows from a stochastic maximum principle. The challenges of the open-loop strategies is that the stochastic maximum principle yields additional co-state variables, of which one must keep track in order to characterize the dynamics of the equilibrium value process. However, as we show in the proof of Theorem 3 later, the simplicity of the agent's problem in our setting allows the dynamics of the equilibrium value process* $W(t)$ *to be represented by a diffusion (without the co-state variables) and* $W(t)$ *to be used as the state variable in the principal's problem. Despite the difficulty of tractability, the co-state variables have the advantage of easier interpretation as the sensitivities of the agent's equilibrium continuation value. In comparison, a closed-loop strategy usually yields a diffusion for the equilibrium value and Markov controls directly, which are easier to analyze but lacking the interpretation of co-state variables as sensitivities. For the agent's problem in our setting, these subtle differences between the two types of controls do not pose an issue. It can be shown that any one-dimensional process with adapted coefficients has a diffusion representation.*[11] *Thus, we can without the loss of generality state Definition 2 directly from a Markov control, and adopt open-loop controls for the agent's problem taking advantage of its easier interpretation.*

2.3 Renegotiation and Commitment

The introduction of time-inconsistent preferences necessitates some formal discussion of the principal's ability to commit to the contract. In particular, a future self of the principal with time-inconsistent preferences may find full commitment to be sub-optimal, as the resulting contract would prevent potentially beneficial alterations, once the principal's own preference

[10] To our knowledge, for closed-loop controls our Theorem 1 is the only result on the existence of the solution when the diffusion term is controlled.

[11] See, e.g., Brunick and Shreve (2013) Corollaries 3.7 and 3.13. Similar results are broadly known as Gyongy-Dupire formulas as discussed in Gyöngy (1986) and Dupire et al. (1994).

has changed over time. To address this issue, we consider contract renegotiation in our model as follows: the principal can offer a new long-term contract at any point in time. If the agent rejects the newly offered contract, the old contract stays in place, but can still be changed in the future by a different self (as in Watson, Miller, and Olsen, 2019). We further assume that the agent breaks indifference in favor of acceptance.[12]

A renegotiation is *feasible* if the agent is willing to accept the contract given his preferences at the time and the contract is incentive compatible with all future selves. Formally:

Definition 3 *An incentive compatible contract $\mathcal{C}'$ is a feasible renegotiation of an incentive compatible contract $\mathcal{C}$ at time t valued at W given information $\mathcal{F}_t$ if the following inequality holds*

$$E_t^{\mathcal{C}'}\left[\int_t^T r^t(s)u(c_s', a_s')ds + r^t(T)U(c_T')\right] \geq E_t^{\mathcal{C}}\left[\int_t^T r^t(s)u(c_s, a_s)ds + r^t(T)U(c_T)\right] = W.$$

We will call $\mathcal{C}'$ a feasible renegotiation of $\mathcal{C}$ at time t that promises $W' \geq W \in \mathbb{R}$ if, in addition,

$$E_t^{\mathcal{C}'}\left[\int_t^T r^t(s)u(c_s', a_s')ds + r^t(T)U(c_T')\right] = W'.$$

Based on this definition, the new set of actions and consumption of each feasible renegotiation contract must generate weakly higher continuation utility for the agent.[13] Note that a renegotiation could increase the payoff not only to the current self but also to the future selves, because those payoffs represent the valuations of the same stream of consumption and actions, albeit with different weights. However, our next proposition shows that, for each feasible and incentive compatible contract $\mathcal{C}'$, there exists a *principal-preferred alternative* $\mathcal{C}''$ that is both feasible and incentive compatible, inducing the same actions and prescribing the same consumption to the agent as $\mathcal{C}'$ does, while keeping the currently renegotiating agent's continuation utility unchanged by lowering only the final consumption c_T. Formally,

[12] By restricting the agent's decision to whether to accept the contract or not we effectively assign the principal with all the bargaining power. It is possible to allow the agent and the principal to bargain over the total surplus at the outset of the contract. This changes the minimum utility that the agent can extract from the contract and thus possibly the optimal initial continuation utility W_0. Assigning the agent bargaining power during the renegotiation process after $t > 0$ though introduces a number of analytical challenges and is thus left to future research.

[13] Our notion of renegotiation defines an incentive compatible contract for the remainder of the contractual relationship. This is a less restrictive notion than the recursively defined set of feasible renegotiations (as in, e.g., Benoit and Krishna, 1993), which explicitly includes the strategies of the principal and the agent's future selves. However, imposing such equilibrium restrictions on the *set* of feasible renegotiations does not change the renegotiation-proof contract because of the focus on a principal intra-personal equilibrium. More discussion can be found in Section 3.1.

Proposition 1 *Let $\mathcal{C}$ be an incentive compatible contract. For any feasible renegotiation $\mathcal{C}'$ of $\mathcal{C}$ at time t given information $\mathcal{F}_t$, there exists a principal-preferred alternative $\mathcal{C}''$ such that $(c'_t, a'_t) = (c''_t, a''_t)$ for all $t < T$ and*

$$E_t^{\mathcal{C}}\left[\int_t^T r^t(s)u(c_s, a_s)ds + r^t(T)U(c_T)\right] = E_t^{\mathcal{C}''}\left[\int_t^T r^t(s)u(c''_s, a''_s)ds + r^t(T)U(c''_T)\right].$$

Moreover, all selves of the principal weakly prefer $\{c''_t, a''_t\}_{t\in[0,T]}$ over $\{c'_t, a'_t\}_{t\in[0,T]}$.

Proposition 1 implies that there is no *Pareto improvement* for the principal's selves by offering the agent higher continuation utility. Furthermore, because $\mathcal{C}''$ induces the same paths of actions and prescribes the same consumption as $\mathcal{C}'$ does except for a lower final consumption, $\mathcal{C}''$ is "preferred" not only by the current principal, but by all the principal's past and future selves. This result is a generalization of Theorem 1 in Fudenberg, Holmstrom, and Milgrom (1990) with similar intuition: while a renegotiation can induce different paths of consumption and actions with different probabilities, the principal-preferred alternative keeps those paths and their probabilities unchanged while decreasing the terminal payments on each path, so that the terminal utility decreases by a fixed amount. This does not change the incentives for any self of the agent because his ranking of the histories remains the same. The principal can reduce the promised utility in this way until it reaches the level of the existing contract. This increases the payoff to all principals' future selves, because the only payoff in the utility function of all her future selves is the terminal one.

In the case of a dynamically consistent preference, Proposition 1 is sufficient to imply that no renegotiation with a higher payoff to the agent is also a Pareto improvement. In the case of a dynamically inconsistent preference, we must also consider the strategic incentives for renegotiation. That is, whether a renegotiation, if accepted, can potentially alter the renegotiation between the principal and agent's future selves. Our next result, which is not covered in Fudenberg, Holmstrom, and Milgrom (1990), demonstrates that renegotiation cannot achieve any *strategic improvement* for either the principal or the agent: neither of them can utilize renegotiation to expand or restrict the set of implementable actions and consumption available to their future selves. Formally:

Proposition 2 *Given any value of continuation utility $W \in \mathbb{R}$, let $\xi_t(W)$ represent the set of feasible distributions over $\{a_s, c_s\}_{s\in[t,T)}$ except the final action, consumption pair (a_T, c_T) that can be generated by a feasible renegotiation that promises the same value W to the t self of the agent. Formally:*

$$\xi_t(W) = \Big\{\{a_s, c_s\}_{s\in[t,T)} \in \mu(t) \ \Big|\ \exists c_T \in \mathbb{R} \ s.\ t.\ \ \mathbb{E}\left[\int_t^T r^t(s)u(c_s, a_s)ds + r^t(T)U(c_T)\right] = W\Big\},$$

where $\mu(t)$ denotes the of distributions over $[0,T)\times[-\infty,\infty]\times[\underline{a},\bar{a}]$. For any $t\in[0,T)$ and any arbitrary pair of W',W'', ($W'\neq W''$) the set $\xi_t(W')$ is equal to $\xi_t(W'')$, $\xi_t(W')=\xi_t(W'')$.

Proposition 2 argues that regardless of how a feasible renegotiation changes the level of continuation utility of the time-t agent, the available set of actions and consumption paths until T that are incentive compatible in the future remains intact. Therefore, each principal's time-t self cannot use renegotiation to strategically influence the behavior and policies of her future selves. Similarly, each agent's time-t self cannot use the acceptance or rejection of a renegotiation to strategically influence the behavior and policies that the principal's future selves can offer. The intuition follows that for Proposition 1. Suppose a principal's current self is considering a feasible renegotiation that is not principal-preferred for the purpose of forestalling a particular renegotiation in the future. Similar to Proposition 1, what the principal's future self could do is simply increasing the final payment on every possible path such that the continuation payoff at the time of renegotiation attains this higher level. In that way, the future self can implement the consumption and action path she desires regardless of the continuation utility carried over, and she would indeed do so. Thus, any renegotiation that is not principal-preferred reduces the payoff for all of the principal's future selves without limiting the set of feasible action paths for any of her future selves.

The absence of this strategic role of renegotiation implies that in our setup, we can without the loss of generality limit our attention to principal-preferred renegotiations. That is, we only need to consider renegotiations that do not change the agent's continuation utility, only how such continuation utility is delivered. Any other renegotiation that would result in a strictly higher continuation utility to the agent is never optimal for the principal and thus never offered.

Remark 2 *Besides applying our framework, one can address the issue of renegotiation in alternative, albeit more restrictive settings. One option is to directly rule out renegotiation by imposing the assumption of full commitment. In this case, the contract is never renegotiated but is also not necessarily renegotiation-proof, because the principal's future selves may evaluate the same payoff streams differently than her time-*0 *self. Another option is to assume special forms of non-exponential discounting that still yield dynamically consistent preference. Then, the contract is (trivially) renegotiation-proof because what is optimal to the principal at time-*0 *remains optimal to her future selves. However, this setting requires specific weights of the discount functions of each self on different dates. Formally, it requires $\frac{R^t(s)}{R^t(s+k)}=\frac{R^{t'}(s)}{R^{t'}(s+k)}$ and $\frac{r^t(s)}{r^t(s+k)}=\frac{r^{t'}(s)}{r^{t'}(s+k)}$ for all t,t',s, and k. In contrast, our objective is to explore the optimal long-term contract in a general setting in which the need for renegotiation is a natural consequence of dynamic inconsistency.*

3 The Optimal Contract

In this section, we define and solve the optimal contract in several steps. First, we establish a formal definition of what the "optimal contract" refers to in our setting. Next, we provide a heuristic derivation of the extended HJB system that characterizes such a contract. Finally, we state our main theorem regarding the characterization and existence of the optimal contract, and compare it to the optimal contract in dynamically consistent benchmarks.

3.1 Intra-personal Game and the Optimal Equilibrium Contract

In this subsection, we establish a formal and economically appropriate notion of what "optimal contract" refers to in this paper. First and foremost, as in the agent's problem, dynamic inconsistency and the possibility of contract renegotiation imply that the contracting problem can be formulated as an intra-personal game played by the different selves of the principal and the agent. However, while the intra-personal game among the agent's selves take the contract terms as given, the intra-personal game of the principal is substantially more involved, as it must take into account the equilibrium of the agent's intra-personal game induced by the principal's contract offers. In light of Lemma 1, we focus on the Markov Perfect Equilibrium (MPE) with $W(t)$ being the state variable. That is, we look for an equilibrium in which the principal's t-self's strategy is a mapping of $W(t)$ onto a contract — a contingent path of recommended actions and consumption that promises $W(t)$ – and the agent's t-self's strategy is to reject or accept the contract and choose an action.

Given Propositions 1 and 2, it is sufficient to verify that a contract is renegotiation-proof as long as there is no deviation that the principal prefers. This implies that a renegotiation-proof contract corresponds to an MPE, because it satisfies the intra-personal equilibrium of the principal and is incentive compatible for all selves of the agent. Altogether, these observations imply that the intra-personal game among the different selves of the principal and the agent can be simplified to an auxiliary game. The time-t principal takes the equilibrium continuation process as given (which depends on the other principals) and chooses the contract terms, assuming she affects only the process at time-t subject to the incentive compatibility constraint identified in Lemma 1. Let

$$f_{\mathcal{C}}(t, W, t) = E^{\mathcal{C}}_{t,W}\left[\int_t^T R^t(s)\,(a_s - c_s)\,ds - R^t(T)c_T\right] \tag{7}$$

denote the continuation payoff of the t principal's time-t self from period t onward given an agent equilibrium under contract $\mathcal{C}$ with a current promised value W. With a slight abuse of notation we let $\mathcal{C}(t, W)$ denote the actions for the principal under the contract $\mathcal{C}$ when time is t and the agent$'s$ continuation payoff under $\mathcal{C}$ is W. Analogous to Definition 2, an *equilibrium contract* of this auxiliary game of the principal is given as follows:

Definition 4 *Consider an incentive compatible contract process* $\mathcal{C}$. *For any initial point and state* (t, W) *and a "small" increment of time* Δ, *define a "deviation" strategy* $\mathcal{C}_\Delta$ *as:*

$$\mathcal{C}_\Delta = \begin{cases} \hat{\mathcal{C}}(t', W') & \text{for } t \leq t' < t + \Delta \\ \mathcal{C}(t', W') & \text{for } t + \Delta \leq t' \leq T \end{cases},$$

where $\hat{\mathcal{C}} \neq \mathcal{C}$ *is another incentive compatible contract and* W' *is the state corresponding to the future date* t'. $\mathcal{C}$ *is called an* ***equilibrium contract*** *if*

$$\liminf_{\Delta \to 0} \frac{f_{\mathcal{C}}(t, W, t) - f_{\mathcal{C}_\Delta}(t, W, t)}{\Delta} \geq 0 \tag{8}$$

for all $\hat{\mathcal{C}}$. *For an equilibrium contract* $\mathcal{C}$, *the corresponding value function* $f_{\mathcal{C}}(t, W, t)$ *is called as an* ***equilibrium value function***.

The equilibrium contract in Definition 4 corresponds to an MPE of the game between the agent's selves the principal's selves and is renegotiation-proof. Each self of the principal and the agent has controls over a positive measure of time that vanishes to zero in the equilibrium (when the gains from deviation also vanish). The local optimality corresponds to the best response of each self to the choices of the other selves, thus characterizing the payoffs each self receives in the intra-personal game. This construction is necessary in the continuous-time setup because each self of the agent/principal only has infinitesimal control.[14]

Remark 3 *We apply closed-loop (Markov) controls for the principal's problem for better interpretation and tractability. Open-loop controls (based on the stochastic maximum principle) yield first- and second-order co-states variables that are difficult to interpret economically and make the construction of Markov strategies analytically challenging.*

Our notion of equilibrium can also be further strengthened if needed. For example, one can impose the definition of regularity introduced in He and Jiang (2019): a contract is a regular equilibrium if a "local" deviation to another contract in a however small time interval

[14] This notion of equilibrium is an extension of the so-called "closed loop" controls introduced Ekeland and Lazrak (2010) by Björk et al. (2017). Studies such as Wei et al. (2017) and Li et al. (2017) have shown that along the trajectory, such controls remain time-consistent and locally optimal, and hence constitute an equilibrium if they exist. Since each self of the principal/agent has no direct effect (measure 0), defining MPE directly in this setup would not be instructive due to the trivial multiplicities from measure 0 changes. Thus to avoid the additional notational burden we directly impose our Definitions 2 and 4 as the MPE of our game. Details about the space of Markov strategies and the equilibrium are available upon request. We thank an anonymous referee for suggesting the definition of equilibrium in this fashion.

is not profitable. Formally, given a contract $\hat{\mathcal{C}}$, there exists Δ^* such that $0 < \Delta < \Delta^*$ and $f_{\mathcal{C}}(t, W, t) \geq f_{\hat{\mathcal{C}}}(t, W, t)$ for all Δ. If the agent's utility function $u(c, a)$ is separable in a and c, our equilibrium is also a regular equilibrium as in He and Jiang (2019). If the utility function is not separable then $u_{ac} > 0$ is required for regularity. See Appendix A.7 for details.[15]

We can now define the optimal contract as the contract that maximizes the principal's continuation payoff at each point in time among the set of renegotiation-proof, incentive-compatible equilibrium contracts $\mathcal{C}$. Formally,

Definition 5 *Let $\mathfrak{C}$ denote the set of equilibrium contracts. A contract $\mathcal{C} \in \mathfrak{C}$ is **optimal** if it maximizes $f_{\mathcal{C}}(t, W, t)$ for all t, W.*

The principal's valuation of an optimal equilibrium contract can now be written as:

$$V(t, W) = \sup_{\mathcal{C} \in \mathfrak{C}} f_{\mathcal{C}}(t, W, t) = \sup_{\mathcal{C} \in \mathfrak{C}} E^{\mathcal{C}}_{t,W} \left[\int_t^T R^t(s) \left(a_s - c_s\right) ds + B(t, W_T) \right]. \tag{9}$$

where $B(t, W_T) \equiv -R^t(T) U^{-1}(W_T)$. In contrast to the problem of a single decision maker, the value function in (9) represents optimality as a notion of an equilibrium, not simply what a single self of the principal deems optimal. Taking the time-t principal as given, the optimal contract maximizes her value at all other times. Put differently, Definition 4 introduces the notion of equilibrium, and 5 defines the optimal contract as a *Pareto-efficient* equilibrium.[16]

In summary, this subsection establishes a formal and economically appropriate notion of what "optimal contract" refers to in this paper. To resolve the issue of dynamic inconsistency, we meticulously laid out several critical definitions and results to reach the conclusion that the most appropriate definition of the optimal contract is a *principal-optimal equilibrium contract* (as in Definitions 4 and 5). In the subsequent analysis, we simply refer to this contract as the *optimal contract* or the *equilibrium contract*.

3.2 Optimal Contract: Heuristic Derivation

This subsection provides a heuristic derivation of the optimal contract with a moderate amount of technicality. The main purposes are threefold: first, to introduce some notations that are not required in time-consistent dynamic programming but are necessary when time-inconsistency is involved. Second, to remind readers the standard HJB equation for a time-consistent benchmark under our notations. Third, to heuristically establish the

[15] We thank an anonymous referee for bringing this connection to our attention.

[16] Note that W_t is an endogenous, equilibrium object, as opposed to the exogenous process used in stochastic control studies featuring similar definitions of optimality (e.g., Björk et al., 2017).

extended HJB systems that are critical for characterizing the optimal contract under time-inconsistency, and compare them to the standard HJB equation. The formal characterization of the optimal contract is given in Theorem 1 in Subsection 3.3.

A. Notations

We first introduce a few necessary notations. We have defined the equilibrium value function $f_{\mathcal{C}}(t, W, t)$ in (7) and the optimal equilibrium value function $V(t, W)$ in (9), both of which are equilibrium objects. In particular, a change in t moves both the self and the time in those objects forward. However, when a fixed self makes forward-looking decisions, it is important to keep track of how the continuation value evolves over time for that particular self. We thus distinguish a changing self from a changing time for a fixed self by treating the third entry in $f_{\mathcal{C}}(t, W, t)$ as a parameter, denoted by a superscript $f^t_{\mathcal{C}}(t, W)$, to highlight that the self is fixed. That is, for a fixed self t, the continuation payoff under a contract $\mathcal{C}$ from time t' onward is denoted as

$$f^t_{\mathcal{C}}(t', W) = E^{\mathcal{C}}_{t',W}\left[\int_{t'}^{T} R^t(s)\,(a_s - c_s)\,ds + B(t, W_T)\right], \tag{10}$$

While $f^t_{\mathcal{C}}(t, W)$ looks very similar to $f_{\mathcal{C}}(t, W, t)$ (which coincides with $V(t, W)$ under the optimal contract), they represent different objects and respond differently to changes in the contract. Specifically, following any changes in the contract $\mathcal{C}$, $f^t_{\mathcal{C}}(t+dt, W_{t+dt})$ captures how the t-self's own valuation from time $t+dt$ onward changes; $f_{\mathcal{C}}(t+d_t, W_{t+dt}, t+dt)$ captures how the $t+dt$'th self's valuation in the current equilibrium changes; and $V(t+dt, W_{t+dt})$ captures how the principal's optimal equilibrium value from $t+dt$ onward changes.

Let $f_{\mathcal{C},t}(t, W, t)$ denote the total derivative with respect to t. Ito's lemma implies

$$\begin{aligned}
f^t_{\mathcal{C}}(t+dt, W_{t+dt}) &= f^t_{\mathcal{C},t}(t, W)dt + \mu^{\mathcal{C}}_W f^t_{\mathcal{C},W}(t, W)dt + \frac{1}{2}(\sigma^{\mathcal{C}}_W)^2 f^t_{\mathcal{C},WW}(t, W)dt + f^t_{\mathcal{C}}(t, W),\\
f_{\mathcal{C}}(t+dt, W_{t+dt}, t+dt) &= f_{\mathcal{C},t}(t, W, t)dt + \mu^{\mathcal{C}}_W f_{\mathcal{C},W}(t, W, t)dt + \frac{1}{2}(\sigma^{\mathcal{C}}_W)^2 f_{\mathcal{C},WW}(t, W, t)dt + f_{\mathcal{C}}(t, W, t),\\
V(t+dt, W_{t+dt}) &= V_t(t, W)dt + \mu^{\mathcal{C}}_W V_W(t, W)dt + \frac{1}{2}(\sigma^{\mathcal{C}}_W)^2 V_{WW}(t, W)dt + V(t, W).
\end{aligned}$$

$\mu^{\mathcal{C}}_W$ corresponds to how principal controls the drift of the W. Similarly, $\sigma^{\mathcal{C}}_W$ corresponds to how principal controls the volatility of W. For ease of exposition, we adopt the notation $\mathcal{A}^C$, known in the literature as the *controlled infinitesimal generator*, to simplify the functions above. That is,

$$\mathcal{A}^{\mathcal{C}} f^t_{\mathcal{C}}(t,W) \equiv \lim_{dt\to 0} \frac{f^t_{\mathcal{C}}(t+dt, W_{t+dt}) - f^t_{\mathcal{C}}(t,W)}{dt}, \tag{11}$$

$$\mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}(t,W,t) \equiv \lim_{dt\to 0} \frac{f_{\mathcal{C}}(t+dt, W_{t+dt}, t+dt) - f_{\mathcal{C}}(t,W,t)}{dt}, \tag{12}$$

$$\mathcal{A}^{\mathcal{C}} V(t,W) \equiv \lim_{dt\to 0} \frac{V(t+dt, W_{t+dt}) - V(t,W)}{dt}. \tag{13}$$

Finally, if the principal implements a different contract $\tilde{\mathcal{C}}$ at a single point of time t, we use $\mathcal{A}^{\tilde{\mathcal{C}}} f_{\mathcal{C}}(t,W,t)$ and $\mathcal{A}^{\tilde{\mathcal{C}}} f^t_{\mathcal{C}}(W,t)$ to capture the effect of such deviation from the contract $\mathcal{C}$ on $f_{\mathcal{C}}(t,W,t)$ and $f^t_{\mathcal{C}}(W,t)$, respectively.

B. Reminder: A Time-consistent Benchmark

Consider a simple benchmark of an infinite horizon dynamic contracting problem with exponential discounting. Using our notations, we can state the principal's optimization problem as follows:

$$V(t,W) = \sup_{\mathcal{C}\in\mathfrak{C}} E^{\mathcal{C}}_{t,W}\left[\int_t^T e^{-rs}(a_s - c_s)ds - e^{-(T-t)}c_T\right].$$

where W_t is the agents continuation utility process with value W at time t. The optimal contract can be summarized using the standard HJB equation:

$$rV(t,W) = \sup_{\tilde{\mathcal{C}}} \left\{(\tilde{a}_t - \tilde{c}_t) + \mu^{\tilde{\mathcal{C}}}_W V_W + \frac{1}{2}\left(\sigma^{\tilde{\mathcal{C}}}_W\right)^2 V_{WW}\right\},$$

where $\tilde{C}$ denotes the contemporaneous pair of contract policies: $\tilde{C} \equiv \{\tilde{a}_t, \tilde{c}_t\}$. Importantly, the HJB equation relies on the fact that when *both* the process W_t and the discounting function e^{-rt} are time-homogeneous, i.e., $\partial V/\partial t = -rV(t,W)$. If time is also a state variable of the contracting problem, a more general HJB equation takes the following form:

$$0 = \sup_{\tilde{\mathcal{C}}} \left\{(\tilde{a}_t - \tilde{c}_t) + V_t + \mu^{\tilde{\mathcal{C}}}_W V_W + \frac{1}{2}\left(\sigma^{\tilde{\mathcal{C}}}_W\right)^2 V_{WW}\right\}.$$

In our notation, the HJB equation for a standard time-consistent dynamic contracting problem can be written as

$$0 = \sup_{\tilde{\mathcal{C}}} \left\{(\tilde{a}_t - \tilde{c}_t) + \mathcal{A}^{\tilde{\mathcal{C}}} V(t,W)\right\}. \tag{14}$$

C. Time-inconsistency: The Extended HJB System

We now turn to our setting. Observe that by construction, we have

$$f_{\mathcal{C}}(t,W,t) = (a_t - c_t)\,dt + E^{\mathcal{C}}_{t+dt,W_{t+dt}}\left[\int_{t+dt}^{T} R^t\,(s)\,(a_s - c_s)\,ds + R^t\,(T)\,B(t,W_T)\right],$$
$$= (a_t - c_t)dt + f_{\mathcal{C}}(t+dt, W_{t+dt}, t). \tag{15}$$

Recall from (9) that $V(t,W_t) = \sup_{\mathcal{C}\in\mathfrak{C}} f_{\mathcal{C}}(t,W_t,t)$. Hence, $V(t+dt,W_{t+dt}) = \sup_{\mathcal{C}\in\mathfrak{C}} f_{\mathcal{C}}(t+dt, W_{t+dt}, t+dt)$. Combined, they imply that $\mathcal{C}$ is optimal between t and $t+dt$, therefore:

$$0 = \sup_{\tilde{\mathcal{C}}}\ \Big\{V(t+dt, W_{t+dt})$$
$$-f_{\mathcal{C},t}(t,W,t) - \mu^{\tilde{\mathcal{C}}}_W f_{\mathcal{C},W}(t,W,t)dt - \frac{1}{2}(\sigma^{\tilde{\mathcal{C}}}_W)^2 f_{\mathcal{C},WW}(t,W,t)dt - f_{\mathcal{C}}(t,W,t)\Big\}. \tag{16}$$

Because $\mathcal{C}$ consists of the optimal choices of $\tilde{\mathcal{C}}$, the last few terms can be combined in to $f_{\mathcal{C}}(t+dt, W_{t+dt}, t+dt)$, which yields

$$0 = \sup_{\tilde{\mathcal{C}}}\ \{V(t+dt, W_{t+dt}) - V(t,W) + f_{\mathcal{C}}(t,W,t) - f_{\mathcal{C}}(t+dt, W_{t+dt}, t+dt)\}. \tag{17}$$

Adding and subtracting $f(t+dt, W_{t+dt}, t)$ and using the identity (15) yields

$$0 = \sup_{\tilde{\mathcal{C}}}\ \{V(t+dt, W_{t+dt}) - V(t,W) + (\tilde{a}_t - \tilde{c}_t)\,dt$$
$$+f_{\mathcal{C}}(t,W,t) - f_{\mathcal{C}}(t+dt, W_{t+dt}, t+dt) + f_{\mathcal{C}}(t+dt, W_{t+dt}, t) - f_{\mathcal{C}}(t,W,t)\}. \tag{18}$$

Dividing both sides by dt and taking the limit as $dt \to 0$ this equation can be written as:

$$0 = \sup_{\tilde{\mathcal{C}}}\ \left\{\mathcal{A}^{\tilde{\mathcal{C}}}V(t,W) + (\tilde{a}_t - \tilde{c}_t) - \mathcal{A}^{\tilde{\mathcal{C}}} f_{\mathcal{C}}(t,W,t) + \mathcal{A}^{\tilde{\mathcal{C}}} f^t_{\mathcal{C}}(t,W)\right\}. \tag{19}$$

In particular, $f^t(t,W)$ must be a martingale since $f^s(t+dt, W_{t+dt})$ captures how the principal's s-selves total payoff changes over time.[17] Thus, under the optimal contract $\mathcal{C}$,

$$\mathcal{A}^{\mathcal{C}} f^s_{\mathcal{C}}(t,W) + R^s(t)(a^{\mathcal{C}}_t - c^{\mathcal{C}}_t) = 0, \tag{20}$$

[17] By law of iterated expectations, $f^t(t,w) = E[\int_t^T R^t(s)(a_s - c_s)ds + B(t,W_T)|\mathcal{F}_t]$ and $E[f^t(s,w)|\mathcal{F}_t] = E[E[f^t(s,w)|\mathcal{F}_s]|\mathcal{F}_t] = f^t(t,w)$ where $s > t$.

for all $s \leq t$.Equations (19) and (20) are the two critical components defining the *extended HJB system* used to formally characterize the solution to the optimal contract next.

3.3 Optimal Contract: Formal Characterization and Existence

We are now in the position to present the main result of this paper:

Theorem 1 *Under Assumptions 1 and 2, there exists a principal-optimal renegotiation-proof equilibrium contract (as in Definition 4 and 5)* $\mathcal{C} = \{a_t, c_t\}_{t\in[0,T]}$. *The principal's equilibrium value function* $V(t, W)$ *under this contract is given by:*

$$V(t, W) = E^{\mathcal{C}}_{t,W}\left[\int_t^T R^t(k)\,(a_k - c_k)\,dk + B(t, W_T)\right]. \tag{21}$$

$V(t, W)$ *satisfies the following* ***extended HJB system****:*

$$\sup_{\tilde{\mathcal{C}}}\left\{\mathcal{A}^{\tilde{\mathcal{C}}}V(t, W) + (\tilde{a}_t - \tilde{c}_t) + \mathcal{A}^{\tilde{\mathcal{C}}}f^t_{\mathcal{C}}(t, W) - \mathcal{A}^{\tilde{\mathcal{C}}}f_{\mathcal{C}}(t, W, t)\right\} = 0, \tag{22}$$

subject to (6), the IC condition (IC), and boundary condition:

$$V(T, W) = B(T, W_T) \text{ for all } W. \tag{23}$$

For each fixed s, $f^s(t, W)$ *is defined as the solution of the following equation:*

$$\mathcal{A}^{\mathcal{C}}f^s_{\mathcal{C}}(t, W) + R^s(t)\,(a_t - c_t) = 0, \tag{24}$$

and $f^s_{\mathcal{C}}(t, W) = f_{\mathcal{C}}(t, W, s)$.[18]

Theorem 1 highlights the novel features of the optimal contract under time-inconsistent preferences. While the principal's value function in a time-consistent benchmark is captured by a single HJB equation involving only $(\tilde{a}_t - \tilde{c}_t)$ and $\mathcal{A}^{\tilde{\mathcal{C}}}V$ (i.e., equation 14 in the previous section), the principal's value function with time-inconsistent preferences involves extra terms in (22). One of the extra terms, $\mathcal{A}^{\tilde{\mathcal{C}}}f^t_{\mathcal{C}}(t, W)$, is the solution to a system of backward stochastic equations (24). These extra terms stem from the different objects represented by the value function $V(t, W)$. In a time-consistent case, the value function tracks how the principal

[18]Under the optimal controls $\mathcal{A}^{\tilde{\mathcal{C}}}V_{\check{\mathcal{C}}}(t, W) = \mathcal{A}^{\tilde{\mathcal{C}}}f_{\check{\mathcal{C}}}(t, W, t) \neq \mathcal{A}^{\tilde{\mathcal{C}}}f^t_{\check{\mathcal{C}}}(t, W)$, which allows a cancellation in Equation (22). The cancellation is useful for the proof of the existence of the solution (which we utilize in the Appendix) but is actually inconvenient for deriving the value function, because the extended HJB system still has both forward and backward components. For specific utility or discount functions, more convenient simplifications of the entire system are possible, as shown in Section 4.

evaluates her discounted future payoff. Changes in time do not change the principal's selves or how the discounting is made. In our setup, the value function tracks how the principal *at each time* (i.e., each of her selves) evaluates the discounted future payoff *in an optimal equilibrium* of her intra-personal game. Changes in time change the principal's relevant self, and the value function is consequently linked to the equilibrium of the intra-personal game between the different selves.

When the principal's t-self evaluates the contract terms, her evaluation is based on her current preferences, which yields the term $f^t_{\mathcal{C}}(t, W)$. Meanwhile, when evaluating a small and potentially sub-optimal change in the contract terms, the principal's t-self must also consider the effect of such a change on her future selves, especially the immediately following self. Such consideration is reflected by the term $f_{\mathcal{C}}(t, W, t)$, which treats the selves as a changing variable. Finally, along the equilibrium path, all selves of the principal correctly anticipate the behavior of her future selves, and her evaluation takes those future behaviors into account. This backward induction logic of the intra-personal game implies that the backward equation system (24) behaves as a martingale and hence always equals to 0.

Besides the heuristic explanation of the extended HJB system, we must establish three additional results to complete the theorem: first, at least one solution to the system exists; second, if multiple solutions exist, they must all be captured by the system we propose; finally, we need a verification theorem showing that the solution to the extended HJB is the value function of the underlying contracting problem. The existence problem is known to be particularly challenging among existing studies. As noted in Björk et al. (2017), "[t]he task of proving existence and/or uniqueness of solutions to the extended HJB system seems (...) to be technically extremely difficult". To address this challenge, we adopt a novel approach connecting the extended HJB system to a static game of incomplete information. We prove that the game has an equilibrium and therefore the system has a solution.[19]

More specifically, because the extended HJB system is the result of the intra-personal game among the principal's different selves, we can treat this system as a static non-atomic game in which the utility function of each player incorporates the solutions of each of the backward systems (equations 24). This is a game of incomplete information with a particular structure: first, the random variable of interest is the entire realization of the Brownian Motion. Second, the game is populated by a mass of players all with different *characteristics* (their discount function R^t) and *information* (the filtration $\mathcal{F}_t$) but the same action set. Third, each player's utility is a function of his or her own action, characteristics, and

[19]The solution may not be unique under the general framework, a common property of dynamic programming problems in which the uniqueness of the solution usually requires problem-specific boundary conditions. Nevertheless, the solution is (without the loss of generality) optimal following Definition 5.

information, as well as an aggregation of other players' actions, while an equilibrium of the game is a distribution over the players' utility functions, strategies, and information.

In the equilibrium, the marginal distribution over the characteristics of the players is uniform over $[0, T]$. Thus, each t-self is a player whose action is the best response to the expected valuations identified by the backward equation given the information $\mathcal{F}_t$. This novel transformation from the intra-personal game to a non-atomic game with incomplete information converts the original dynamic problem into a static one, and the dynamic optimization of the original game is reduced to the information set of each player in the static game.[20] Utilizing Balder (1991), this non-atomic static game with incomplete information has an equilibrium in distributional strategies. Thus, the existence of an equilibrium in this non-atomic game implies the existence of a solution to the extended HJB system. We provide a verification argument for Theorem 1: any contract $\mathcal{C}$ that solves the extended HJB system must be an equilibrium contract. We prove that the converse is also true: any equilibrium contract $\mathcal{C}$ must also solve the extended HJB system. Consequently, Theorem 1 captures all principal-optimal renegotiation-proof equilibrium contracts in our model.

The proof of existence and the verification theorem are critical components of Theorem 1 but are technically involved. Readers who are only interested in applying our framework to solve problems with dynamically inconsistent decision makers can adopt the results in Theorem 1, knowing that a solution exists and is well behaved. In Section 4, we apply Theorem 1 to a specific family of discount functions in which we obtain closed-form solutions and conduct straightforward comparative statics.

Remark 4 *Before closing this section, we wish to highlight that the following seemingly natural approach of proving Theorem 1 will encounter serious technical hurdles: selecting an arbitrary incentive compatible contract $\mathcal{C}$, defining $f^t(t, W)$ via the backward equations (24), solving $V(t, W)$ by maximizing the HJB equation, finding $\mathcal{C}'$ (the contract that generates the solved $V(t, W)$), and starting this process over from $\mathcal{C}'$ in an attempt to find a fixed point using standard arguments (e.g. Schauder fixed-point theorem). However, such an approach requires establishing technical conditions on the mapping of $\mathcal{C}$ to $\mathcal{C}'$, which are endogenous objects. Typically, it is difficult to ascertain which combination of the model primitives will guarantee those conditions.*

[20] To the best of our knowledge, such transformation is new and allows us to establish the existence of a solution to similar systems of PDEs. Thus, we believe this proof technique can be fruitfully used to prove the existence of a larger class of dynamic systems. Readers interested in applying this proof of existence strategy can find the details of this technique in the Appendix.

4 Application: Time-difference Discount Functions

In this section we demonstrate the applications of our general framework. We show that a particular class of discounting functions: the *time-difference discount family*, combined with a widely-used set of simplifying assumptions on the agent's side: CARA utility plus hidden savings, delivers closed-form solutions to the optimal contract (Section 4.1). We then further explore a specific example in this discount family, quasi-hyperbolic discounting, and offer a comparison of the resulting optimal contract with that under exponential discounting (Section 4.2). Finally, we discuss the practical implications of our results (Section 4.3).

4.1 General Time-difference Discounting

The time-difference discount family is defined as follows:

Definition 6 *A family of discount functions $R^t(s)$ is called a* time-difference discount family *if $R^t(s) = R^{t+k}(s+k) = R(s-t)$ for all s, t, k. We call the function $R(s-t)$ as a time difference discount function.*

The unique feature of this type of discount function is that it relies on $s-t$ only and not on t or s individually, thus the notation $R(s-t)$ in lieu of $R^t(s)$. This discount family includes some of the well-studied time-inconsistent preferences, such as quasi-hyperbolic discounting (as in e.g., Harris and Laibson, 2012) and anticipatory utility (as in e.g., Loewenstein, 1987). Restricting the principal's discount function to within this family means that $\mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}(t, W, t)$ and $\mathcal{A}^{\mathcal{C}} f^t_{\mathcal{C}}(t, W)$, the two critical terms in the extended HJB system in Theorem 1 can be simplified as follows

$$\begin{aligned}\mathcal{A}^{\tilde{\mathcal{C}}} f_{\mathcal{C}}(t, W, t) &= E^{\mathcal{C}}_{t,W}\left[\int_t^T \Big(R(s-t)\tilde{\mathcal{A}}^{\mathcal{C}}(a_s - c_s) - R'(s-t)(a_s - c_s)\Big)\, ds\right] \\ &- R(T-t)\tilde{\mathcal{A}}^{\mathcal{C}} E^{\mathcal{C}}_{t,W}\left[c_T\right] + R'(T-t)E^{\mathcal{C}}_{t,W}\left[c_T\right],\end{aligned}$$

and

$$\tilde{\mathcal{A}}^{\mathcal{C}} f^t_{\mathcal{C}}(t, W) = E^{\mathcal{C}}_{t,W}\left[\int_t^T R(s-t)\mathcal{A}^{\tilde{\mathcal{C}}}(a_s - c_s)ds\right] + R(T-t)\tilde{\mathcal{A}}^{\mathcal{C}} E^{\mathcal{C}}_{t,W}\left[c_T\right].$$

Therefore, the extended HJB system is reduced to the following *single HJB equation*:

$$\sup_{\tilde{\mathcal{C}}} \{\mathcal{A}^{\tilde{\mathcal{C}}} V + (\tilde{a}_t - \tilde{c}_t) + E^{\mathcal{C}}_{t,W}\left[\int_t^T R'(s-t)(a_s - c_s)ds\right] - R'(T-t)E^{\mathcal{C}}_{t,W}\left[c_T\right]\} = 0. \quad (25)$$

This single HJB equation (25) can be further simplified by imposing some structures on the agent's side. Doing so is without the loss of generality because, as we have shown in Section 2, the main impact of dynamic inconsistency in our model is on the principal side. In particular, we introduce the following widely-used simplifying set of assumptions on the agent's side:

Assumption 3 *The agent has the following exponential discounting function:*

$$r^t(s) = e^{-\gamma(s-t)}, \forall t, s \in [0, T]$$

where $\gamma > 0$ *and a constant-absolute-risk-aversion (CARA) utility function:*

$$u(c, a) = -\frac{1}{\eta} e^{-\eta\left(c - \frac{k}{2}a^2\right)},$$

where $\eta > 0$ *measures his degree of risk-aversion and* $k > 0$ *captures the marginal cost of effort. The agent also has access to a private savings account of which the balance grows at rate* γ*.*

Assumption 3 brings two sets of benefits: first, the principal optimal long-term contract would be renegotiation-proof if the principal were also an exponential discounter (as in Holmstrom and Milgrom, 1987). This allows us to isolate the impact of renegotiation-proofness for the ensuing analysis regarding non-exponential discounting. Second, Assumption 3 implies a simplification of the agent's problem that is commonly used in the contracting literature.[21] In our model, this simplification manifests as follows:

Lemma 2 *Under Assumption 3 the agent's continuation utility satisfies*

$$dW_t = \psi_t(dM_t - a_t dt) \ , \tag{26}$$

$$u(c_t, a_t) = \gamma W_t \ . \tag{27}$$

The agent's incentive compatibility condition becomes

$$\psi_t = k a_t. \tag{28}$$

[21]Studies that utilize the same CARA utility plus hidden savings combination include He (2011), Williams (2015), Gryglewicz and Hartman-Glaser (2020), Marinovic and Varas (2019), and Bloedel, Krishna, and Strulovici (2020), etc. For simplicity we also assume $U = u(\cdot, 0)$, that is equivalent to normalizing the discount rate to 1 when the relationship ends. Other normalizations follow analogously and yield qualitatively similar results. Note that because there exists a well-behaved process W_t (given in Lemma 2 below) representing the agent's continuation utility here theorem 1 is still applicable under this simplified environment despite the fact that CARA and hidden savings are not part of Assumption 2.

The evolution of $\ln(W)$ *is given by*

$$E\left[\ln(-W_t)\right] = \ln(-W_0) + \frac{1}{2}\int_0^t \eta^2\gamma^2\sigma^2 \ k^2 a_s^2 ds. \tag{29}$$

Finally, there is no private savings on the equilibrium path.

Here, ψ_t represents the contract's incentive power: the variations in the agent's continuation utility following the realized cash flows. The dynamics of ψ_t has important practical implications which we discuss in Sections 4.2 and 4.3. Substituting the results from Lemma 2 into the single HJB equation (25) implies the principal's value function $V(t, W)$ solves

$$\begin{aligned} 0 = \sup_{a_t} \ & V_t + a_t - \left[\frac{k}{2}a_t^2 - \frac{\ln(\eta\gamma)}{\eta} - \frac{1}{\eta}\ln(-W)\right] + \frac{1}{2}\left(a_t k\gamma\eta W\sigma\right)^2 V_{WW} \\ & + \int_t^T R'(s-t)\left(a_s - \left[\frac{k}{2}a_s^2 - \frac{\ln(\eta\gamma)}{\eta} - \frac{1}{\eta}\ln(-W_s)\right]\right) ds \\ & + \frac{\ln(-W_T)}{\eta} R'(T-t), \end{aligned} \tag{30}$$

subject to (26), the incentive compatibility condition (28), and the boundary condition:

$$V(T, W) = B(T, W) \text{ for all } W. \tag{31}$$

We guess and verify that $V(t, W) = A_t \ln(-W) + B_t$, for some undetermined coefficients A_t and B_t. Substituting this guessed $V(t, W)$ into (30) and matching the coefficients yields the following solution to the optimal contract:

Proposition 3 *If the principal has a time-difference discount function* $R(t-s)$ *and the agent's discounting and preference follows that given in Assumptions 3, then under the optimal contract,*

$$\psi_t = a_t = \frac{1}{k + K + \eta k^2\gamma^2\sigma^2 \int_0^t [R(T-s) + R'(T-s)]ds}, \tag{32}$$

where K *is a constant of integration, and*

$$c_t = \frac{k}{2}a_t^2 - \frac{\ln(\gamma\eta)}{\eta} - \frac{1}{\eta}\ln(-W_t). \tag{33}$$

Proposition 3 immediately implies the following results:

Corollary 1 *The equilibrium effort path is monotonic if*

- $R(s) + R'(s) < 0 \;\; \forall s \in [0, T]$ *(increasing path).*

- $R(s) + R'(s) > 0 \;\; \forall s \in [0, T]$ *(decreasing path).*

The equilibrium effort path is non-monotone if the sum $R(\cdot) + R'(\cdot)$ ever changes sign.

These features of the optimal contract demonstrate that the equilibrium path of the agent's actions is deterministic, driven by the discount function R and its first derivative R' only. Note that exponential discounting is also a special case of the time-difference discount family. Thus, the equilibrium path of effort and contract incentive under exponential discounting is always monotonic in time (increasing if $r < 1$ and decreasing if $r > 1$). In the next subsection, we explore specific (time-inconsistent) examples in the time-difference discount family and compare the resulting optimal contracts with that under exponential discounting.

The theoretical predictions of Proposition 3 can be further sharpened by assigning explicit functional forms to the discount function. In this paper we solve two specific examples with wide economic applications: quasi-hyperbolic discounting (following Harris and Laibson, 2012), and anticipatory utility (following Loewenstein, 1987). We present the details of the former case below and demonstrate how its predictions may shed light on empirical observations on managerial compensation in practice. For the latter case we first demonstrate how anticipatory utility leads to the specific discounting function above, and then derive and discuss its properties. The derivation and discussion are straightforward following the general results of Proposition 3 but are left in appendix in the interest of space.

4.2 Quasi-Hyperbolic Discounting

In this section we further explore a specific time-inconsistent discount function that belongs to the time-difference family: the quasi-hyperbolic discounting, widely-featured in various studies including O'Donoghue and Rabin (1999), Thaler and Benartzi (2004), Harris and Laibson (2012), Jackson and Yariv (2014), Jackson and Yariv (2015), Bisin, Lizzeri, and Yariv (2015), etc. Here, we follow Harris and Laibson (2012) and define a quasi-hyperbolic discount function as a convex combination of a short-term discount function and a long-term discount factor. Formally, we assume

Assumption 4 *The principal has the following time difference discount function:*

$$R^t(s) = R(s-t) = (1-\beta)e^{-(\rho+\lambda)(s-t)} + \beta e^{-\rho(s-t)} \tag{34}$$

with $\beta \in (0, 1)$ and $\gamma > \rho + \lambda$.

The representation above is a deterministic characterization of a principal who values "near present" returns with a larger discount factor (discounted by $e^{-\rho(s-t)}$), and "far future" returns with a smaller factor (discounted by $\beta e^{-\rho(s-t)}$). That is, the principal becomes less patient over time. $\beta < 1$ captures the size of the drop in the discount factor in the far future. The switch between the "near present" and the "far future" occurs stochastically at the rate of λ, and the overall discount function $R^t(s)$ incorporates this expected drop.

Because $R(T-s)$ is an exponential function in s, substituting it back into (32) yields a closed-form solution, which is given in Appendix A.6. This closed-form solution implies the following analytical characterization of the properties of the optimal contract:

Proposition 4 *Under Assumption 4, the optimal renegotiation-proof contract derived in Proposition 3 has the following properties*

1. *If $\lambda + \rho \geq 1$, the incentive power ψ_t and wage c_t of the optimal contract decreases if β decrease, for all t and all W_t.*
2. *$\lambda + \rho < 1$ then there exists $T(\lambda, \rho) > 0$ such that*
   (a) *If $T \leq T(\lambda, \rho)$, the incentive power ψ_t and wage c_t of the optimal contract decreases if β decrease, for all t and all W_t.*
   (b) *If $T > T(\lambda, \rho)$ then there exists $t(\lambda, \rho) < T(\lambda, \rho)$ such that*
      i. *The incentive power ψ_t and wage c_t of the optimal contract increases if β decrease, for all $t < t(\lambda, \rho)$ and all W_t,*
      ii. *The incentive power $\psi_{t'}$ and wage $c_{t'}$ of the optimal contract decreases if β decrease, for all $t' \geq t(\lambda, \rho)$ and all $W_{t'}$.*

The proposition demonstrates the effect of quasi-hyperbolic discounting and the associated present bias on the dynamic characteristics of the optimal contract. In the case of exponential discounting, the discount factor ρ is commonly interpreted as the death rate or the hazard rate of the principal. Under quasi-hyperbolic discounting, the principal also faces the additional risk λ of being replaced by a more impatient self. Thus, her overall survival probability is negatively correlated with the sum $\lambda + \rho$. The higher the sum, the lower the survival probability of the time-t principal. Note that this survival probability is not equivalent to a higher discount factor but a combination of termination (standard discount) and devaluation (hyperbolic drop).

A decrease in β, or a larger hyperbolic drop, has two effects: first, an *internalization effect*, whereby the current self of the principal internalizes the decrease in patience of her future self, and all her selves become more impatient; second, an *equilibrium effect*, whereby the principal front-loads contract incentives, because she correctly predicts that she cannot credibly promise sufficient dynamic incentives in equilibrium in the future, and thus front-loads the incentives before the impatient self arrives. If the survival probability is low ($\lambda + \rho \geq 1$) or the horizon T is short, the internalization effect dominates. The wage and incentive power of the contract decrease unambiguously with the level of patience of the replacing self. However, if T is long enough and the current self has a sufficiently high probability of survival ($\lambda + \rho < 1$), the equilibrium effect results in the front-loading of contract wage and incentive power. The prominence of the front-loading increases as β decreases. That is, the more impatient the principal's future self, the more her current self shifts wage and incentive power of the contract towards the early period of the contracting relationship while lowering them towards the end of the contracting horizon.

We also make the following observations regarding how close the optimal contract resembles a time-consistent benchmark:

Proposition 5 *Under Assumption 4 the optimal renegotiation-proof contract derived in Proposition 3 has the following properties:*

1. *If $\beta = 1$ (or $\beta = 0$), the optimal contract is identical to that for a dynamically-consistent principal with discount rate ρ (or $\rho + \lambda$), and the incentive power ψ_t is (weakly) monotonic in time t.*

2. *As the contracting horizon becomes infinitely long (i.e., $T \to \infty$), the optimal contract converges to that for a time consistent-principal with discount rate $\frac{\rho(\lambda+\rho)}{\rho+\beta\lambda}$.*[22]

3. *For every $\rho \in (0, 1)$, the optimal incentive power ψ_t and the equilibrium actions a_t are non-monotone in t as long as λ is high enough.*

The first property is straightforward: if $\beta = 1$ ($\beta = 0$) the principal has the same discount rate for "near present" payoff and "far future" payoff, which is equivalent to having exponential discounting. Following Holmstrom and Milgrom (1987), the optimal incentive power is monotonic in time, regardless of the contracting horizon.[23]

The second property demonstrates that the effect of quasi-hyperbolic discounting becomes indistinguishable from exponential discounting if the time horizon becomes arbitrarily large. If the contracting horizon is infinitely long, the solution to the dynamically inconsistent

[22] More precisely, a contract $\mathcal{C}$ converges to another contract $\mathcal{C}'$ if the processes ψ_t, a_t, c_t under $\mathcal{C}$ converge pointwise to the processes under $\mathcal{C}'$.

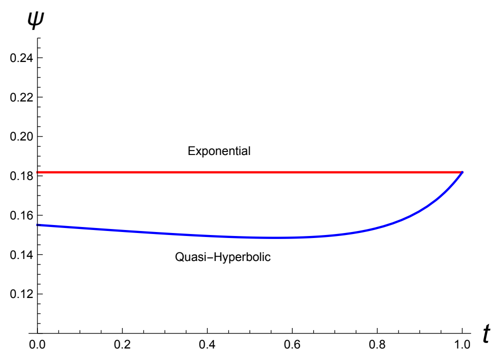
[23] Whether incentive power is monotonically increasing, decreasing, or constant over time is the result of model-specific boundary conditions. See He (2011), He, Wei, Yu, and Gao (2017), and Marinovic and Varas (2019) for some recent examples.

problem agrees with that of a dynamically consistent one with *exactly the same instantaneous utility function* but a *different exponential discount factor*. Intuitively, as the horizon increases, the problems faced by each of the principal's time-t selves become more similar. When the horizon is infinitely long, the principal's problem becomes completely stationary. Under quasi-hyperbolic discounting, stationarity is sufficient to ensure that the resulting optimal contract converges to that of a principal with the exact same instantaneous utility function but a different discount rate.

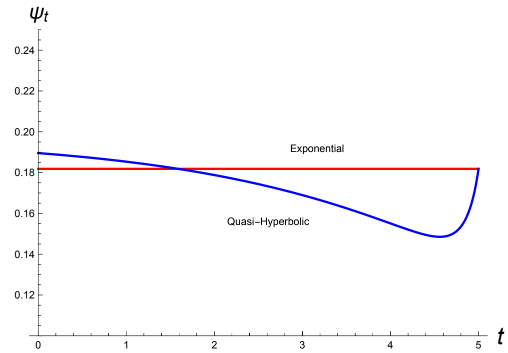
The third property highlights the possibility of non-monotonic incentive power under finite horizon. Based on the first two properties, we know that the incentive power is (weakly) monotonic in time without quasi-hyperbolic discounting or without a finite horizon. Under quasi-hyperbolic discounting and a finite horizon, however, a "near-future" principal with the short-term discount factor and a "far-future" principal with the long-term discount factor may prefer different levels of incentive power. Suppose that the former prefers a high power contract and the latter prefers low power. Because the switch between the two types of principals occurs stochastically, the sophisticated "near-future" principal anticipates that her less patient "far-future" self will arrive at some point. Thus, she designs a path of incentive power that gradually decreases towards the level preferred by her "near-future" self. However, at some point, as she approaches the end of the contracting horizon (the deadline), the probability of the switch happening before the deadline decreases over time as the contract "runs out of time." Consequently, the principal resembles more closely her "near future" patient self and reverts the incentive power of the contract until it converges to the time-consistent benchmark level at the deadline, causing a U-shaped path. We refer to this reverting of incentive power towards the end of the contracting horizon as the "deadline effect". The turning point is determined by λ, the arrival intensity of the drop in discount. The higher the λ, the sooner the action path changes course. If λ is sufficiently low, the drop in discount is very remote such that the "deadline effect" does not occur, and the path of optimal incentive power converges to the (weakly) monotonic case.

We illustrate the properties above and the comparison between quasi-hyperbolic discounting and the time-consistent benchmark in Figure 1.[24]

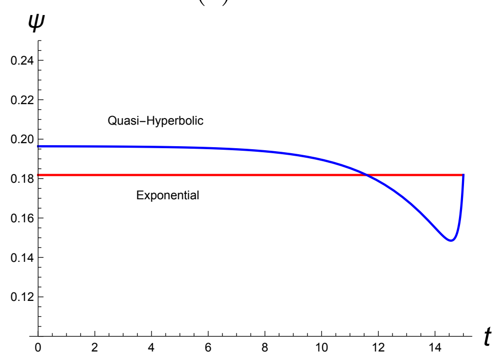
[24]When T is finite, discounting becomes irrelevant at $t = T$, and the optimal incentive power at that point (ψ_T) must be the same under both types of discounting. As T increases, the difference in the contracts induced by the two types of discounting is pushed further to the future. When $T \to \infty$, such difference is pushed infinitely far away, and the optimal contract under quasi-hyperbolic discounting coincides with that under exponential discounting but with a different discount factor (i.e., Property 2 in Proposition 5).



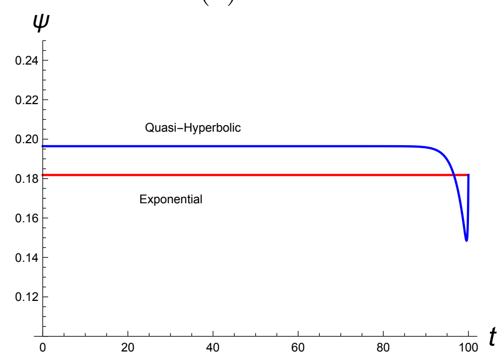
(a) $T = 1$



(b) $T = 5$



(c) $T = 15$



(d) $T = 100$

Figure 1: These plots illustrate the paths of the incentive power (also the paths of agent's actions given Proposition 3) under the optimal contract. Blue lines indicate the paths under quasi-hyperbolic discounting with $\eta = 2, k = 1, \sigma = 1, \beta = 0.5, \lambda = 5, \rho = 0.5, \gamma = 1.5$. Red lines indicate the paths under exponential discounting with the same parameters except for $\rho = 1$. Each plot corresponds to a different time horizon (different T). If $\rho > 1$ ($\rho < 1$), the equilibrium path under exponential discounting would be monotonically increasing (decreasing).

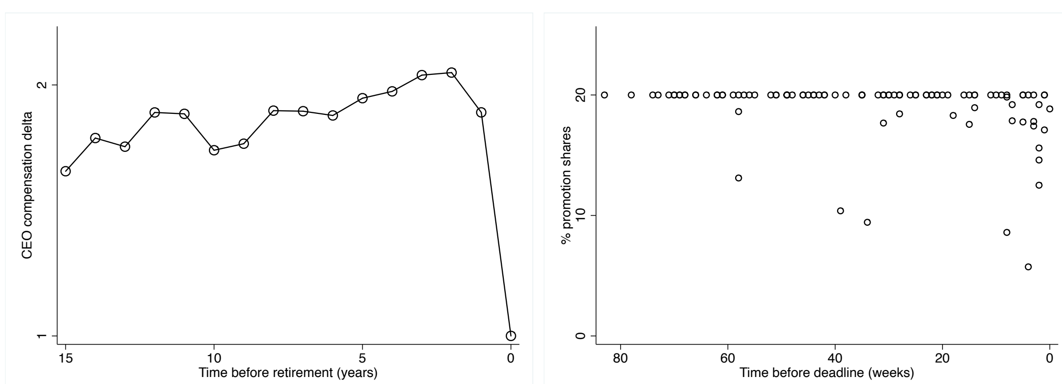


Figure 2: The panel on the left illustrates the average CEO compensation delta as a function of the years before CEO retirement. Data come from Compustat Execucomp from 1992 to 2018. The delta of each CEO in each year is scaled by the value of delta in the final year. The panel on the right illustrates the percentage of the promotion shares given to the sponsors of special purpose acquisition companies (SPAC) as a function of the time before the deal deadlines from 2010-2019. Each circle represents one completed SPAC deal.

4.3 Empirical Implications

In this section we briefly discuss the practical application of the theoretical results given in Propositions 4 and 5 and illustrated in Figure 1. While those theoretical results summarize both the recommended actions and the incentive power prescribed by the optimal contract, the latter is arguably much easier to observe in practice. Furthermore, Property 2 of Proposition 5 highlights that quasi-hyperbolic discounting and exponential discounting produce different predictions only in the presence of a finite contracting horizon. Hence, we examine below practical settings in which a contracting horizon can be fairly clearly identified and the incentive power of the contract can be reasonably measured.

Our first example explores the time-series variations in CEO compensation delta as the CEO approaches retirement. Delta measures the change in the CEO's compensation if the firm's stock price changes by 1% and is a commonly used empirical proxy for the incentive power of the CEO contract (e.g., Core and Guay (2002), Coles, Daniel, and Naveen (2006), and the references within). We calculate the average CEO compensation data as a function of the time until retirement for firms in the Compustat Execucomp database from 1992 to 2018. Retirement in our sample is identified in two steps: first, if the variable *reason* (for turnover) in Execucomp equals "retired". However, less than 20% of the CEOs in the Execucomp database are assigned an explicit value for this variable. Therefore, for the observations with missing *reason*, we identify retirement as CEOs who are more than 65-year-old when their contracts end and have not appeared to be CEOs for a different firm afterward in the database. To eliminate the impact of outliers, we only keep CEOs whose tenure before the retirement is longer than 6 years. Finally, because the CEOs of larger firms receive higher compensation and thus have a mechanically higher delta, we scale the delta of each CEO in each year prior to retirement by the delta of the final year of the CEO's contract. We then plot the sample-average delta as a function of the time before retirement in the left panel of Figure 2. As the plot shows, on average, compensation delta is constant except in the two years before retirement, where delta declines significantly. This pattern illustrates a clear "deadline" effect in contrast to the prediction under exponential discounting shown in Figure 1 but is consistent with the prediction based on quasi-hyperbolic discounting.

Another setting to which our theoretical predictions might by applicable is the incentives to the managers of special purpose acquisition companies, or SPACs. A SPAC is a shell company listed on a public exchange without any concrete business. Its sole purpose is to find and merge with a private company and take the latter public, and its popularity has exploded since the outbreak of the coronavirus pandemic. The SPAC managers, known as the sponsors, are delegated with the task of finding the private company target to merge with, and their effort in this process is compensated with a fraction of the SPAC shares.

These promotion shares can therefore be regarded as the sponsor's incentives similar to the role of inside equity in dynamic moral hazard models (e.g., DeMarzo and Sannikov, 2006). The sponsor is typically given two years after the SPAC IPO to find and complete the merger. Otherwise, the SPAC is dissolved and the capital returned to the investors. The deadline can be extended following majority shareholder approval and proper filing with the SEC, which means that the actual deadline is publicly known and observable in the data. Using hand-collected data from the SEC filings, we plot the percentage of the sponsor's promotion shares as a function of the time before the deadline in the right panel of Figure 2 for 102 complete SPAC combos during 2010 and 2019. As the plot shows, in most deals the incentive shares are 20% according to the industry standard. However, there are substantial variations especially near the deadline, which is once again more consistent with the predictions based on quasi-hyperbolic discounting.[25]

5 Conclusion

Non-exponential discounting has been widely observed in practice and extensively studied in economic research. However, dynamic moral hazard and long-term contracting between sophisticated parties with non-exponential discounting has thus far eluded formal analysis due to several challenges. First, individuals with non-exponential discounting make dynamically inconsistent plans, and what currently appears optimal may become sub-optimal in the future. In such a scenario, what is the appropriate notion of optimality? Moreover, how does the contract provide incentive for a dynamically inconsistent agent, and how to ensure that a contract agreed by the agent and the principal today will actually be carried out or renegotiated in the future after their preferences have changed? Finally, does a long-term optimal contract always exist? If so, how is such a contract different from the optimal contract under exponential discounting?

In this paper, we answer these questions by establishing a broad framework of an *incentive compatible, renegotiation-proof, optimal long-term* contract under dynamically inconsis-

[25] The consistency with the theoretical predictions notwithstanding, the empirical patterns documented in Figure 2 should *not* be viewed as definitive proofs for the validity of the model. Our theoretical results are derived under the assumption that the principal has quasi-hyperbolic discounting, which is possible when the principal is represented by the board of directors or a group of investors (Jackson and Yariv, 2015). However, whether quasi-hyperbolic discounting is an accurate summary of the average board of directors for firms in the Excucomp database or the SPAC investors requires substantial empirical work and is beyond the scope of this paper. It is also possible to explain the specific patterns shown in Figure 2 with reasons other than quasi-hyperbolic discounting. For example, the precipitous drop in the incentive power for the SPAC sponsors near the deadline could be due to the sponsors taking a cut of their own pay in order to push the proposed deals through. Overall, we urge caution in interpreting any empirical observations based solely on the predictions of this model.

tent preferences resulting from non-exponential discounting. We formulate the contracting problem as a dynamic, intra-personal game played among the agent, the principal and their respective future "selves". This allows us to derive the optimal contract as the equilibrium of intra-personal game and prove its existence. We demonstrate the applicability of our general framework through the example of time-difference discount family, which yields closed-form solutions and generates testable implications, such as the precise impact of the "deadline effect" on the incentive power and the compensation scheme of the optimal contract.

The agency friction of our general framework is adopted from the dynamic hidden effort problem in Sannikov (2008). A related strand of literature formulates the agency friction as a cash-flow diversion (CFD) problem. The analytical tools developed in this paper can be easily adapted to CFD problems and their various extensions to highlight the impact of non-exponential discounting on a number of applications, including but not limited to security design (Biais, Mariotti, Plantin, and Rochet, 2007), risk management (Biais, Mariotti, Rochet, and Villeneuve, 2010), liquidity management (Bolton, Chen, and Wang, 2011), and resource allocation (Feng and Westerfield, 2021).

Our framework can be expanded in different directions. In particular, the extended HJB system derived in this model and its variations can arise in many dynamically inconsistent optimal control problems for reasons other than non-exponential discounting. Our main theorem, which bridges dynamically inconsistent control problems with non-atomic games, may be modified to establish the solutions to these problems at both theoretical and practical levels. Examples of the practical applications include but are not limited to ambiguity, habit formation and mean-variance risk preferences. Finally, exploring the impact of bargaining power during the renegotiation process under time-inconsistent preferences can also yield fresh insights. We leave these topics for future research.

A Appendix

A.1 Proof of Lemma 1

Consider a contract $\mathcal{C}$ offered by the principal. In general, the consumption and action paths induced by $\mathcal{C}$ could depend on the entire history of the outputs. To resolve the history dependence, we follow Williams (2011, 2015) and introduce a change of the probability measure. We differentiate using the conditional expectations $E^{Z^{\hat{a}}}$ and E^{Z^0}, respectively. That is, the optimal control problem of the agent under the new measure induced by Z_t^0 can now be formulated as follows:

$$J(t,\Gamma) = \sup_{\hat{a}_s,\, s\in[t,T]} E^{Z^0}\left[\int_t^T \Gamma_s r^t(s)u(c_s,\hat{a}_s)ds + \Gamma_T r^t(T)U(c_T)\right], \tag{35}$$

$$d\Gamma_s = \left(\frac{\hat{a}_s}{\sigma}\right)\Gamma_s dZ_s^0, \quad \Gamma_0 = 1. \tag{36}$$

where

$$E^{Z^0}\left[\int_t^T r^t(s)u(c_s,\hat{a}_s)ds + r^t(T)U(c_T)\right] = E^{Z^{\hat{a}}}\left[\int_t^T \Gamma_s r^t(s)u(c_s,\hat{a}_s)ds + \Gamma_T r^t(T)U(c_T)\right].$$

Note that, although the optimal control problem is still time-inconsistent, the change of measure resolves the history dependence. This is a common technique used in dynamic contracting studies especially those with time-consistent preference (i.e., exponential discounting) and persistent private information, such as (in addition to Williams, 2011, 2015) He, Wei, Yu, and Gao (2017), Marinovic and Varas (2019), Bloedel, Krishna, and Strulovici (2020), Feng (2021), etc.

Now consider an arbitrary progressively-measurable control process $\hat{a}_t$. Under $\hat{a}_t$, the evolution of $d\Gamma_s = \left(\frac{\hat{a}_s}{\sigma}\right)\Gamma_s dZ_s^0$ is an Ito process, with a measurable diffusion coefficient, thus by Theorem 3.6 and in particular by Corollary 3.7 of Brunick and Shreve (2013) it is without loss to consider the SDE $d\Gamma_s = E\left(\left(\frac{\hat{a}_s^2}{\sigma^2}\right)\Gamma_s^2|\Gamma_s\right)dZ_s^0$ instead, where the term $E\left(\left(\frac{\hat{a}_s^2}{\sigma^2}\right)\Gamma_s^2|\Gamma_s\right)$ only depends on (s,Γ) because it is a conditional expectation. This representation is unique due to Corollary 3.13 of Brunick and Shreve (2013), and $\hat{a}_s(s,\Gamma)$ can simply be backed out by taking the square root and multiplying by Γ/σ. Thus, any measurable control is equivalent to a Markov control. We can then apply the following theorem of Yan and Yong (2019) to the time-inconsistent control problem defined by (35) and (36):

Theorem 2 (Yan and Yong 2019) *Given a contract $\mathcal{C}(a,c)$ suppose that $(\hat{a},\Gamma)$ is an agent intra-personal equilibrium pair (where each agent takes the action suggested by the principal) and suppose that for any given $t\in[0,T)$, the first order adjoint processes $W(\cdot,t),\tilde{\psi}(\cdot,t)$ and second order adjoint processes $P(\cdot,t),\Lambda(\cdot,t)$ are adapted solutions to the following BSDEs:*

$$dW(s,t) = -\left(\frac{\hat{a}_t}{\sigma}\tilde{\psi}(s,t) + r^t(s)u(c_s,\hat{a}_s)\right)ds + \tilde{\psi}(s,t)dZ_s^0 \quad s\in[t,T],$$

$$W(T,t) = r^t(T)U(c_T),$$

$$dP(s,t) = -\left(\left(\frac{\hat{a}_t}{\sigma}\right)^2 P(s,t) + 2\frac{\hat{a}_t}{\sigma}\Lambda(s,t)\right)ds + \Lambda(s,t)dZ_s^0 \quad s\in[t,T],$$

$$P(T,t) = 0,$$

then almost surely for any deviation $\tilde{a}$ given the suggested path $a = \hat{a}$ we have the following global form of Pontyagrin maximum principle[26]

$$0 \leq \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} E_t \int_t^{t+\varepsilon} \left\langle \tilde{\psi}(s,t), \Gamma_s \left(\frac{\tilde{a} - \hat{a}}{\sigma} \right) \right\rangle ds + r^t(t)\Gamma_t \left(u(c_t, \tilde{a}) - u(c_t, \hat{a}) \right)$$
$$+ \frac{1}{2} \left(\Gamma_t \frac{\tilde{a} - \hat{a}_t}{\sigma} \right)^2 P(t,t). \tag{37}$$

If the agent has exponential discounting, there would be a single adjoint process. Instead, we have a flow of adjoint processes, which jointly describes how the different selves of the agent evaluate the payoff from the contract. Taking the limit as $\varepsilon \to 0$ yields

$$\psi(t,t)\Gamma_t \left(\frac{\tilde{a} - \hat{a}}{\sigma} \right) + r^t(t)\Gamma_t \left(u(c_t, \tilde{a}) - u(c_t, \hat{a}) \right) + \frac{1}{2} \left(\Gamma_t \frac{\tilde{a} - \hat{a}_t}{\sigma} \right)^2 P(t,t) \leq 0. \tag{38}$$

We can now define the Hamiltonian system $\mathcal{H}$ as

$$\mathcal{H} \equiv H - \frac{1}{2} \left(\frac{\hat{a}_t}{\sigma} \Gamma_t \right)^2 P(t,t) + \frac{1}{2} \left(\frac{\tilde{a}_t - \hat{a}_t}{\sigma} \Gamma_t \right)^2 P(t,t). \tag{39}$$

where $H \equiv -r^t(s)\Gamma_s u(c_s, \hat{a}_s) + \Gamma_s \tilde{\psi}(t,s) \frac{\hat{a}_s}{\sigma}$ denotes the local Hamiltonian. Then, (38) is equivalent to the maximum of $\mathcal{H}$ by choosing $\tilde{a}$. Since $[\underline{a}, \bar{a}]$ is convex and the coefficients of Γ, u and $r(t,s)$ are Lipschitz continuous, maximizing $\mathcal{H}$ is equivalent to maximizing H.[27] The first-order condition of H with respect to $\hat{a}_t$ thus yields the incentive compatibility (IC) condition of the agent at time t:

$$u_a(c, \hat{a}) = \frac{\tilde{\psi}(t,t)}{\sigma} = \psi_t. \tag{40}$$

Next, we derive the evolution of the agent's continuation utility. We write the equilibrium value function in the integral form as follows:

$$\tilde{W}(k,t) = r^t(T)U(c_T) - \int_k^T \left(\frac{\hat{a}_s}{\sigma} \tilde{\psi}(s,t) + r^t(s)u(c_s, \hat{a}_s) \right) ds + \int_k^T \tilde{\psi}(s,t) dZ_s^0.$$

[26] In the original reference the term $\frac{1}{2}$ in front of the trace term is missing, which is a typo. See Wang and Yong (2021), Proposition 2.3 for a corrected version of this result.

[27] This follows since $\mathcal{H}$ and H have the same generalized Clark gradient in $\tilde{a}$. For this equivalence see page 138 Lemma 5.1 of Yong and Zhou (1999), for a slightly more detailed proof see Proposition 3.1 of Djehiche and Huang (2016).

Taking $k \to t$, we have

$$\tilde{W}(t,t) = r^t(T)U(c_T) - \int_t^T \left(\frac{\hat{a}_s}{\sigma}\tilde{\psi}(s,t) + r^t(s)u(c_s,\hat{a}_s)\right) ds + \int_t^T \tilde{\psi}(s,t)dZ_s^0.$$

Following Yan and Yong (2019), we can define the following Type-I backward stochastic Volterra integral equation (BSVIE):

$$W(t) = r^t(T)U(c_T) - \int_t^T \left(\frac{\hat{a}_s}{\sigma}\tilde{\psi}(s,t) + r^t(s)u(c_s,\hat{a}_s)\right) ds + \int_t^T \tilde{\psi}(s,t)dZ_s^0,$$

and $W(t) = \tilde{W}(t,t)$ and $W(t) = J(t,\hat{\Gamma})$ for all t as stated by Yan and Yong (2019). Replacing Z_s^0 with Z_t^a yields the representation provided in Lemma 1. Note that $\frac{\hat{a}_s}{\sigma}\tilde{\psi}(s,t)+r^t(s)u(c_s,\hat{a}_s)$ is continuously differentiable, progressively measurable, and uniformly Lipschitz, because u and r are also uniformly Lipschitz, and all other terms are linear. Therefore the conditions in Theorem 4.3 in Hernández and Possamaï (2021) apply and the above Type-I BSVIE is well-posed.

Remark 5 *Before proceeding we need to establish that $W(t)$ above can be used as the state variable to for the principal's problem. Since each agent t only acts once, it is natural to find a forward representation for $W(t)$. In particular, we use the decoupling field of the forward-backward stochastic Volterra integral equation (FBSVIE) to pin down the evolution of $W(t)$. Under Lipschitz continuity of both the forward and backward coefficients, Theorem 1 of Wang and Yong (2019b) shows that decoupling field θ can be represented as follows:*

Theorem 3 (Wang and Yong 2019) *Under Assumption 2, the integral valued process $W(t)$ has a unique representation as a diffusion $\theta(t,t,\Gamma_t)$, where $\theta(t,s,\Gamma)$ is the unique, continuous solution to the following partial differential equation (PDE):*

$$\theta_s(t,s,\Gamma_s) + \frac{1}{2}\left(\frac{\hat{a}_s}{\sigma}\Gamma_s\right)^2 \theta_{\Gamma\Gamma}(t,s,\Gamma_s) + r^t(s)u(c_s,\hat{a}_s) = 0, \tag{41}$$

$$\theta(t,T,\Gamma_T) = r^t(T)U(c_T)\Gamma_T. \tag{42}$$

This theorem requires two conditions noted as $H1$ and $H2$ in Wang and Yong (2019b). Here, $H2$ is directly satisfied by the boundedness of the derivative of the utility function, and the linear growth condition in $H1$ is also satisfied because the difference in volatility of the Γ process is always linearly bounded (by $\bar{a}\Gamma$). Therefore, the above PDE has a unique continuous solution, which is the decoupling field of the BSVIE. Since $\theta(t,t,\Gamma) = W(t)$, Ito's *lemma implies the following diffusion representation of $W(t)$:*

$$dW(t) = \left(\theta_1(t,t,\Gamma_t) + \theta_2(t,t,\Gamma_t) + \frac{1}{2}\left(\frac{\hat{a}_t}{\sigma}\Gamma_t\right)^2 \theta_{\Gamma\Gamma}(t,t,\Gamma_t)\right) dt + \frac{\hat{a}_t}{\sigma}\Gamma_t\theta_\Gamma(t,t,\Gamma_t)dZ_t^0$$
$$= (\theta_1(t,t,\Gamma_t) - u(c_t,\hat{a}_t))\,dt + \frac{a_t}{\sigma}\Gamma_t\theta_\Gamma(t,t,\Gamma_t)dZ_t^0. \quad (43)$$

where second equality follows from Equation 39. Here, $\theta_1(\theta_2)$ denotes the partial derivative with respect to the first (second) argument. $\theta_1(t,t,\Gamma_t)$ and $\theta_{\Gamma\Gamma}$ are the standard components regardless of time-consistency while $\theta_2(t,t,\Gamma_t)$ is the unique term capturing the valuation of the agent's different selves under time-consistent preferences.

A.2 Proof of Proposition 1

Since $\mathcal{C}' = \{c_t', a_t'\}_{t\in[0,T]}$ is a feasible renegotiation, it must be incentive compatible, and

$$E_t^{\mathcal{C}'}\left[\int_t^T r^t(s)u(c_s', a_s')ds + r^t(T)U(c_T')\right] \geq E_t^{\mathcal{C}}\left[\int_t^T r^t(s)u(c_s, a_s)ds + r^t(T)U(c_T)\right].$$

For any given renegotiation $\mathcal{C}'$, the principal-preferred alternative by definition $\mathcal{C}''$ must incentivize the same action paths as $\mathcal{C}'$ but implies the lowest W such that the renegotiation is still feasible. Therefore, if the weak inequality holds with equality then the principal-preferred alternative is the feasible renegotiation itself. If the inequality is strict then let $k > 0$ be the difference in utility from agent t's perspective, that is

$$k = E_t^{\mathcal{C}'}\left[\int_t^T r^t(s)u(c_s', a_s')ds + r^t(T)U(c_T')\right] - E_t^{\mathcal{C}}\left[\int_t^T r^t(s)u(c_s, a_s)ds + r^t(T)U(c_T)\right].$$

Now define $\mathcal{C}''$ as follows: *for any potential history* H_T up to time $(T, \mathcal{F}_T)$, we define c_T'' such that $r^t(T)U(c_T'') = r^t(T)U(c_T') - k$. Such c_T'' exists and $c_T'' < c_T'$ because U is continuous, increasing in c and unbounded below. For all $t < T$, choose $\{c_t'', a_t''\}_{t\in[0,T]}$ such that $(c_t', a_t') = (c_t'', a_t'')$. Observe that compared to $\mathcal{C}'$, the agent's selves *realized* utility under $\mathcal{C}''$ is lower by a constant amount. This amount equals exactly k for the t-self. Given that $\mathcal{C}'$ is incentive compatible, the new contract $\mathcal{C}''$ is also incentive compatible.

A.3 Proof of Proposition 2

Pick any $\{a_s^1, c_s^1\}_{s\in[t,T)} \in \xi_t(W_1)$ that corresponds to an arbitrary history from t to T. Because u is continuous, increasing in c and unbounded, there exists c_T^2 such that $r^t(T)U(c_T^2) = r^t(T)U(c_T^1) + W_1 - W_2$ for every possible c_T^1. Then, let $(c_t^2, a_t^2) = (c_t^1, a_t^1)$ for all $t < T$.

Observe that compared to the contract $(c_t^1, a_t^1)_{t\in[0,T]}$, the contract $(c_t^2, a_t^2)_{t\in[0,T]}$ changes each of the agent's selves' *realized* utility for every path of history by a constant amount. This amount equals exactly $W_1 - W_2$ for the t-self. Given that (c_t^1, a_t^1) is incentive compatible, the new contract (c_t^2, a_t^2) is also incentive compatible. Furthermore, at time t, the contract (c_t^2, a_t^2) delivers W_2 and induces the path $\{c_s^1, a_s^1\}_{s\in[t,T]}$. Thus, $\{c_s^1, a_s^1\}_{s\in[t,T]} \in \xi_t(W_2)$. An identical argument (i.e., switching the superscript 1 and 2 above) shows that any element of $\xi_t(W_2)$ is also an element of $\xi_t(W_1)$.

A.4 Proof of Theorem 1

We first prove the existence of a solution to the extended HJB system, and then provide a verification theorem.

A.4.1 The Existence of a Solution to the Backward System

Begin with an arbitrary incentive compatible contract $\breve{\mathcal{C}}$. For each (t, W), the backward equation of the HJB system

$$\mathcal{A}^{\breve{\mathcal{C}}} f^t_{\breve{\mathcal{C}}}(t, W) + R^t(t)(\breve{a}_t - \breve{c}_t) = 0,$$

is a semi-linear parabolic partial differential equation. If it has a solution $f^t_{\breve{\mathcal{C}}}(t, W)$, then for every s there is a path of W_s that reaches W by time t.[28] Thus,

$$f^t_{\breve{\mathcal{C}}}(t, W) = f^t_{\breve{\mathcal{C}}}(T, W_T) - \int_t^T R^t(s)(\breve{a}_s - \breve{c}_s)ds - \int_t^T Y_s^{\breve{\mathcal{C}}} dZ_s,$$

for any incentive compatible $\breve{\mathcal{C}}$ and some adapted process $Y^{\breve{\mathcal{C}}}$. A solution to the backward system corresponds to a pair of processes $f^t_{\breve{\mathcal{C}}}(t, W)$ and $Y_t^{\breve{\mathcal{C}}}$. However, notice that for a given arbitrary incentive compatible contract $\breve{\mathcal{C}}$, action $\breve{a}_t$ is already pinned down by the IC condition. In particular, according to Assumption 2 for any given $\breve{c}_t$, $\breve{\psi}_t$, and W_t, the incentive compatible action $\breve{a}_t$ is unique. Moreover, given contract $\breve{\mathcal{C}}$, the proof of Lemma 1 shows that the backward system is accompanied by a forward system (43), where θ denotes the decoupling field of the agent's BSVIE. Also note that $f^t_{\breve{\mathcal{C}}}(t, W)$ and $Y_t^{\breve{\mathcal{C}}}$ do not appear in the forward system, hence this is a *decoupled* forward-backward stochastic differential equation (FBSDE).Thus, proving the existence of a solution to the backward system is equivalent to showing that for any incentive compatible control $\breve{\mathcal{C}}$, the FBSDE system is well-posed and therefore has a unique and continuous solution.

[28] The subscript $\breve{\mathcal{C}}$ is used to emphasize the dependence on the control $\breve{\mathcal{C}}$.

Given that the system is decoupled, we first write down the generator of the backward system. Since we only consider Markovian controls that satisfy incentive compatibility, the generator only depends on the forward part (i.e., $R^t(s)(\breve{a}_s - \breve{c}_s)$). Based on Cvitanic and Zhang (2012), Section 9.5, the FBSDE is well-posed if both the forward and backward components have unique solutions and the solution to the forward component satisfies the Markov property. For the forward component, the proof of Lemma 1 shows that the accompanying Type-I BSVIE is well posed and has a unique continue solution. Then, Theorem 3 implies that this well-posed Type-I BSVIE has a unique representation as a diffusion (hence Markovian) identified via the decoupling field. For the backward component, observe that it does not have the backward terms in the generator but only the forward ones, and the generator $R^t(s)(\breve{a}_s - \breve{c}_s)$ is uniformly Lipschitz continuous in $f^t_{\breve{\mathcal{C}}}(t, W)$ and Y_t. Thus, by Theorem 9.3.5 in Cvitanic and Zhang (2012) the backward system has a unique solution.[29] Finally, by Theorem 3, the decoupling field is $C^{1,2}$, and since the discount function is assumed to be C^3, the generator itself is C^3. Thus, by Corollary 2.9 of Pardoux and Peng (1992), $f^t_{\breve{\mathcal{C}}}(t, W)$ has continuous partial derivatives of order 1 and 2 in t and W, respectively.[30]

A.4.2 Finding a Fixed Point to The Extended HJB System

The preceding argument proves that for any incentive compatible control $\breve{\mathcal{C}}$ there exists a unique process $f^t_{\breve{\mathcal{C}}}(t, W)$ satisfying the backward system. Furthermore, if $\breve{\mathcal{C}}$ is optimal for any time t under the optimal controls we have:

$$V(t, W) = f^t_{\breve{\mathcal{C}}}(t, W) = f_{\breve{\mathcal{C}}}(t, W, t), \tag{44}$$

$$\mathcal{A}^{\tilde{\mathcal{C}}} V(t, W) = \mathcal{A}^{\tilde{\mathcal{C}}} f_{\breve{\mathcal{C}}}(t, W, t) \neq \mathcal{A}^{\tilde{\mathcal{C}}} f^t_{\breve{\mathcal{C}}}(t, W). \tag{45}$$

The second line stems from the fact that for $\mathcal{A}^{\tilde{\mathcal{C}}} V(t, W)$ and $\mathcal{A}^{\tilde{\mathcal{C}}} f_{\breve{\mathcal{C}}}(t, W, t)$, the infinitesimal generator $\mathcal{A}^{\tilde{\mathcal{C}}}$ changes both the self and time, whereas $\mathcal{A}^{\tilde{\mathcal{C}}} f^t_{\breve{\mathcal{C}}}(t, W)$ is taken from the perspective of a fixed self.[31] Therefore, the first and last terms in (22) cancel, and the problem of the principal's t-self is reduced to the following static optimization problem:

$$\sup_{\tilde{\mathcal{C}}} \{(\tilde{a}_t - \tilde{c}_t) - \mathcal{A}^{\tilde{\mathcal{C}}} f^t_{\breve{\mathcal{C}}}(t, W)\} = 0, \tag{46}$$

[29] The fact that backward terms do not appear in generator is not surprising when the backward system is a sort of "continuation utility", see Duffie and Epstein (1992) or El Karoui, Peng, and Quenez (1997).

[30] Since the decoupling field is $C^{1,2}$ by construction we do not need to make further assumptions about the drift and variance of the forward process.

[31] Note that $\mathcal{A}^{\tilde{\mathcal{C}}} V(t, W) \neq \mathcal{A}^{\tilde{\mathcal{C}}} f_{\breve{\mathcal{C}}}(t, W, t)$ under any incentive compatible but sub-optimal contract.

for each W. For any given control $\breve{\mathcal{C}}$, the solution to the equation above generates another control $\tilde{\mathcal{C}}$ using the value function generated from the backward system, $f^t_{\breve{\mathcal{C}}}(t, W)$.

Next, we define a probability space $(\Omega, \mathcal{P}, \mathcal{F})$ where $\Omega = Z_{[0,T]}$ and Z is the Brownian motion. We denote ω as an arbitrary realized path of Z_t. Let $S = A \times C \times \mathbb{R}$ denote the space of strategies, $[0, T]$ the space of players, and $\mathcal{P}$ the common knowledge of all players. In particular, we treat the incomplete observations of the realized path by each player as their differential information. The information of a player t is a sub σ-algebra of $\mathcal{F}$, which is denoted by $\mathcal{F}_t$, and corresponds naturally to the filtration $\mathcal{F}$ given that players' different information is a result of the difference in time.

Let $\mathfrak{M}$ denote the space of all measures on S. Let $\mathfrak{S}$ denote the set of all measurable decision rules δ, $\delta : \Omega \to \mathfrak{M}$ (i.e., players are allowed to randomize). We equip $\mathfrak{S}$ with the weak topology, which is the weakest topology in which all functions identified as follows are continuous on $\mathfrak{S}$:

$$\delta \to \int_\Omega \phi(\omega) \left[\int_S c(s)\delta(\omega)ds \right] P(d\omega), \phi \in L^1(\Omega, \mathcal{P}, \mathcal{F}), c \in C_B(S).$$

Here, $L^1(\Omega, \mathcal{P}, \mathcal{F})$ denotes the space of $\mathcal{P}$-integrable functions and $C_B(A)$ the space of bounded and continuous functions. We equip $(\Omega, \mathcal{P})$ with the usual L^1 norm.

With this notation, a pure strategy profile is a distribution over $\mathfrak{S}$, and we denote the set of all probability distributions over $\mathfrak{S}$ as $M(\mathfrak{S})$. For each t, we define $\mathfrak{S}_t \subset \mathfrak{S}$ as the set of all $\mathcal{F}_t$ measurable decision rules δ, $\delta : \Omega \to \mathfrak{M}$.

To proceed, define the set

$$D = \{(\mathcal{F}_t, \delta) : \mathcal{F}_t \in \{\mathcal{F}_t\}_{t\in[0,T]} \text{ and } \delta \in \mathfrak{S}_t\}.$$

Since $\{\mathcal{F}_t\}_{t\in[0,T]}$ is a filtration, $\mathcal{F}$ is the appropriate σ-algebra in which the conditional expectations are measurable. Due to Lemma 2 of Balder (1991), D is $\mathcal{F} \times \mathcal{B}(\mathfrak{S})$ measurable and $\mathfrak{S}_t$ is a compact subset of $\mathfrak{S}$ for every $\mathcal{F}_t \in \{\mathcal{F}_t\}_{t\in[0,T]}$.[32] In other words, player-t only uses information from $\mathcal{F}_t$ when selecting a strategy from D.

Note that by definition any contract $\mathcal{C}$ defined in the main text satisfies $\mathcal{C} \in M(\mathfrak{S})$ with the additional property that $\{a_t, c_t\} \in \mathfrak{S}_t \;\; \forall t$, i.e., player-$t$ indeed only uses information from $\mathcal{F}_t$. For any given t, ω, and $\mathcal{F}_t$, the backward system for any contract $\mathcal{C}$ yields:

$$f^t_{\mathcal{C}}(t, W) = E^{\mathcal{C}}_{t,W} \left[\int_t^T R^t(s)(a_s - c_s)ds + B(t, W_T) \right].$$

[32] Where $\mathcal{B}(\mathfrak{S})$ is the Borel sigma algebra defined on $\mathfrak{S}$.

The above equation only depends on $\mathcal{C}$ and ω and is calculated according to the information at $\mathcal{F}_t$. Meanwhile, for any player t, any $\delta \in \mathfrak{S}$ and $h \in M(\mathfrak{S})$, we can define the utility function as:

$$u(t, \delta, \mathcal{C}) = \mathbb{E}\left[(a(\delta) - c(\delta)) - \mathcal{A}^{(a(\delta), c(\delta))} f^t_{\mathcal{C}}(t, W) | \mathcal{F}_t\right].$$

Observe that u_t is continuous in both its arguments and is measurable by definition. Furthermore, f is continuous and differentiable since the backward system is well-posed. Because $f^t_{\mathcal{C}}(t, W)$ incorporate the randomness generated by Z_t, a player is identified by the characteristics $(t, \mathcal{F}_t)$, where t corresponds to the player's identity and $\mathcal{F}_t$ the player's information. In other words we interpret the uncertainty of the contracting problem as the following structure of the intra-personal game: first, an entire path of Z is realized. Then, there is a population $[0, T]$ of non-atomic players, among which a player t is endowed with information $\mathcal{F}_t$.

Analogous to Balder (1991), we define the *utility function for the game* as a function $U : [0, T] \times D \times M(\mathfrak{S}) \to \mathbb{R}$. Thus, a game is identified by $([0, T] \times \mathcal{F}, \mu, \{\mathfrak{S}_t\}_{t \in [0,T]}, U\}$, where $[0, T] \times \mathcal{F}$ denotes the set of characteristics, μ a given distribution over characteristics, $\mathfrak{S}_t$ the set of strategies available to a player with characteristics $(t, \mathcal{F}_t)$, and U the utility function for the game.

Next, we can identify a *characteristic - strategy* (CS) distribution, which is a distribution over $[0, T] \times D$ specifying how the possible characteristic-strategy combinations are distributed in the game:

Definition 7 *A CS distribution λ is an equilibrium if*

- *The marginal of $\lambda_{|[0,T] \times \mathcal{F}} = \mathcal{P} \times \lambda^T_U$,*
- $\lambda(\{((t, \mathcal{F}_t), \delta) \in [0, T] \times D : \delta \in \arg\max_{\mathfrak{S}_t} u(t, \delta, \lambda_{|\mathfrak{S}})\}) = 1,$

where λ^T_U denotes the uniform distribution over $[0, T]$.

In words, if λ is a CS distribution, then the respective marginals of the distribution $\lambda_{|[0,T] \times \mathcal{F}}$ coincide with the distribution over characteristics and $\lambda_{|\mathfrak{S}}$ coincides with a distribution over the strategies. Observe that among players of different types, the distribution over $[0, T] \times \mathcal{F}$ is identified by the filtered probability space generated by $Z_{[0,T]}$. Because each principal is weighted equally, the distribution over the player type must be uniform. Then, we can invoke Theorem 1 of Balder (1991), which states that the non-atomic game has an equilibrium distribution if the following conditions hold:

1. The metric space S of strategies is separable and complete.

2. S_t is compact for every agent.

3. $U(t,\cdot,\cdot)$ is upper semi continuous.

4. $U(t,s,\cdot)$ is continuous.

5. $U(\cdot,\cdot,\nu)$ is measurable.

We verify that all of these conditions are met in our setting. First, $\mathfrak{S}$ is complete, separable, and compact, as noted above. Second, $U(t,\cdot,\cdot)$ is continuous in both arguments. Finally, $U(\cdot,\cdot,\nu)$ is measurable for all $\nu \in \mathfrak{S}$. Therefore, an equilibrium distribution exists, which implies that there exists a contract $\mathcal{C}^*$ such that

$$(a^*, c^*) \in \arg\sup_{\tilde{\mathcal{C}}}\{\tilde{a}_t - \tilde{c}_t - \mathcal{A}^{\tilde{\mathcal{C}}} f^t_{\mathcal{C}^*}(t,W)\} = 0.$$

for every (t,W). Thus, there exists a solution to the extended HJB system.

A.4.3 Verification Theorem

We divide the verification theorem into two parts. First, we prove that if a contract $\mathcal{C}$ solves the extended HJB system, it must be an equilibrium of the game. Then we prove the reverse: if a contract $\mathcal{C}$ is an equilibrium of the game, then it must also solve the extended HJB system. The following proposition formalizes the verification result.[33]

Proposition 6 *Assume a, c, V and $f_{\mathcal{C}}$ are C^2 with respect to W and C^1 with respect to t, then the following statements hold.*

1. *If a contract $\mathcal{C}$ solves the extended HJB equation, then it must be an equilibrium.*

2. *If a contract $\mathcal{C}$ is an equilibrium, then it also solves the extended HJB equation.*

Proof of the first statement: Applying Ito's formula to $f^s_{\mathcal{C}}(t,W)$ in (24) and using the boundary condition at time T, we obtain the following representation

$$f^s_{\mathcal{C}}(t,W) = E^{\mathcal{C}}_{t,W}\left[\int_t^T R^s(r)(a_r - c_r)dr + B(t,W_T)\right].$$

[33] A recent study by Lindensjö (2019) provides a similar verification theorem. However, the model in Lindensjö (2019) does not involve any moral hazard problem. The state variable evolves exogenously, and only terminal consumption is allowed. In contrast, we consider a principal-agent problem in which the agent's continuation utility evolves endogenously in the equilibrium, and allow for both flow and terminal consumption in the contract.

From the extended HJB system

$$0 = \mathcal{A}^{\mathcal{C}} V(t, W) + \mathcal{A}^{\mathcal{C}} f^t_{\mathcal{C}}(t, W) - \mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}(t, W, t) + a_t - c_t, \tag{47}$$

$$0 = a_t - c_t + \mathcal{A}^{\mathcal{C}} f^t_{\mathcal{C}}(t, W), \tag{48}$$

implying that $\mathcal{A}^{\mathcal{C}} V(t, W) = \mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}(t, W, t)$ for all (t, W). Since V is smooth, by Ito's lemma:[34]

$$EV(T, W_T) = V(t, W) + E\left[\int_t^T \mathcal{A}^{\mathcal{C}} V(s, W_s) ds\right],$$

which can be re-written as

$$EV(T, W_T) = V(t, W) + E\left[\int_t^T \mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}(s, W, s) ds\right].$$

By applying the same reasoning as above to $f^t(t, W)$ and using the boundary conditions for V and f time T, we conclude that $V(t, W) = f_{\mathcal{C}}(t, W, t)$.

Next, we show that $\mathcal{C}$ is an equilibrium. Suppose the agent uses an arbitrary control law $\breve{\mathcal{C}}$ over period length $\Delta > 0$. Let $f_{\mathcal{C}_\Delta}(t, W, t)$ be the payoff to the principal under C_Δ. For $\mathcal{C}$ be an equilibrium, the following condition

$$\liminf_{\Delta \to 0} \frac{f_{\mathcal{C}}(t, W, t) - f_{\mathcal{C}_\Delta}(t, W, t)}{\Delta} \geq 0. \tag{49}$$

must hold for any $\mathcal{C}_\Delta$. Note that,

$$\frac{f_{\mathcal{C}}(t, W, t) - f_{\mathcal{C}_\Delta}(t, W, t)}{\Delta} = (a_t - c_t) - (a^\Delta_t - c^\Delta_t) - \mathcal{A}^{\mathcal{C}_\Delta} f^t_{\mathcal{C}}(t, W) + \mathcal{A}^{\mathcal{C}} f^t_{\mathcal{C}}(t, W).$$

Since $\mathcal{C}$ solves $V(t, W)$ in (22), $(a_t - c_t) + \mathcal{A}^{\mathcal{C}} f^t_{\mathcal{C}}(t, W, t) = 0$. Thus, $(a^\Delta_t - c^\Delta_t) + \mathcal{A}^{\mathcal{C}_\Delta} f^t_{\mathcal{C}}(t, W)$ must be negative, which implies $f_{\mathcal{C}}(t, W, t) - f_{\mathcal{C}_\Delta}(t, W, t) > 0$ is positive. Hence, contract $\mathcal{C}$ must be an equilibrium.

Proof of the second statement: If we can show that $\hat{\mathcal{C}}$ is a continuous equilibrium contract and the corresponding value function V and f are sufficiently smooth (i.e., $V, f \in C^2$), then V solves the extended HJB equation. To do so, we first prove two auxiliary results. For a given contract $\mathcal{C}$, we define $f_{\mathcal{C}}(t, W, t), V_{\mathcal{C}}(t, W), f^t_{\mathcal{C}}(t, W)$ as in (7), (9), and (10), respectively. Using these notation we can reformulate our objective as the following Feynman-Kac form:

[34]See Strulovici and Szydlowski (2015) for the sufficient conditions for the smoothness of the value function and the existence of optimal controls with time-consistent preferences.

Lemma 3 *Consider a continuous contract $\mathcal{C}$ and moreover assume $f_{\mathcal{C}}^t$ is twice continuously differentiable and L^2 integrable. Then for every $s \in [0, T]$ $f_{\mathcal{C}}^s$ is a solution to the following PDE*

$$\mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}^s(t, W) + R^s(t)\,(a_t - c_t) = 0 \quad (t, W) \in [0, T] \times \mathbb{R}.$$

Proof of Lemma 3: By construction,

$$f_{\mathcal{C}}^s(t, W) = E_{t,W}^{\mathcal{C}}\left[\int_t^T R^s(k)\,(a_k - c_k)\,dk + B(t, W_T)\right], \tag{50}$$

with the boundary condition $f^s(T, W) = B(s, W)$, and

$$0 = E_{t,W}^{\mathcal{C}}\left[\int_t^{t+\Delta} R^s(r)\,(a_r - c_r)\,dr\right] + E_{t,W}^{\mathcal{C}}\left[f_{\mathcal{C}}^s(t+\Delta, W_{t+\Delta})\right] - f_{\mathcal{C}}^s(t, W)$$

for a sufficiently small Δ. Applying Ito's formula to the term $f_{\mathcal{C}}^s(t+\Delta, W_{t+\Delta})$ yields

$$\begin{aligned} 0 = {} & E_{t,W}^{\mathcal{C}}\left[\int_t^{t+\Delta} R^s(r)\,(a_r - c_r)\,dr\right] + f_{\mathcal{C}}^s(t, W) + E_{t,W}^{\mathcal{C}}\left[\int_t^{t+\Delta} \mathcal{A}^{\mathcal{C}} f_{C}^s(r, W_r)dr\right] \\ & + E_{t,W}^{\mathcal{C}}\left[\int_t^{t+\Delta} \frac{\partial^2 f}{\partial^2 W}(\sigma_W^{\mathcal{C}})^2 dZ_r\right] - f_{\mathcal{C}}^s(t, W). \end{aligned}$$

Simplifying and using the fact that the expectation of an Ito integral is equal yields

$$0 = E_{t,W}^{\mathcal{C}}\left[\int_t^{t+\Delta} R^s(r)\,(a_r - c_r)\,dr\right] + E_{t,W}^{\mathcal{C}}\left[\int_t^{t+\Delta} \mathcal{A}^{\mathcal{C}} f_{C}^s(r, W_r)dr\right].$$

Dividing both side to Δ and taking the limit $\Delta \to 0$ we have $R^s(t)(a_t - c_t) + \mathcal{A}^{\mathcal{C}} f_{C}^s(t, W) = 0$.

Lemma 3 implies the next result:

Lemma 4 *For any contract $\mathcal{C}$ and deviation $\mathcal{C}_\Delta$,*

$$\lim_{\Delta \to 0} \frac{f_{\mathcal{C}}(t, W, t) - f_{\mathcal{C}_\Delta}(t, W, t)}{\Delta} = -(a_t^\Delta - c_t^\Delta) - \mathcal{A}^{\mathcal{C}_\Delta} f_{\mathcal{C}}^t(t, W).$$

Proof of Lemma 4: Ito's lemma implies:

$$\begin{aligned} \lim_{\Delta \to 0} \frac{f_{\mathcal{C}}(t, W, t) - f_{C_\Delta}(t, W, t)}{\Delta} &= (a_t - c_t) - (a_t^\Delta - c_t^\Delta) - \mathcal{A}^{\mathcal{C}_\Delta} f_{\mathcal{C}}^t(t, W) + \mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}^t(t, W), \\ &= -(a_t^\Delta - c_t^\Delta) - \mathcal{A}^{\mathcal{C}_\Delta} f_{\mathcal{C}}^t(t, W), \end{aligned}$$

where the second line follows Lemma 3.

We now use the two results to complete the proof of Part 2. If $\hat{\mathcal{C}}$ is an equilibrium contract, then $\mathcal{A}^{\tilde{\mathcal{C}}} V_{\hat{\mathcal{C}}}(t,W) = \mathcal{A}^{\tilde{\mathcal{C}}} f_{\hat{\mathcal{C}}}(t,W,t)$. Let $\hat{\mathcal{C}}_\Delta$ denote an arbitrary deviation from $\hat{\mathcal{C}}$ over Δ amount of time. Then, Lemma 4 implies

$$\lim_{\Delta \to 0} \frac{f_{\hat{\mathcal{C}}}(t,W,t) - f_{\hat{\mathcal{C}}_\Delta}(t,W,t)}{\Delta} = -(\hat{a}_\Delta - \hat{c}_\Delta) - \mathcal{A}^{\hat{\mathcal{C}}_\Delta} f^t_{\hat{\mathcal{C}}}(t,W) \geq 0.$$

where the last inequality applies because $\hat{\mathcal{C}}$ is an equilibrium contract. Therefore, $(\hat{a}_\Delta - \hat{c}_\Delta) + \mathcal{A}^{\hat{\mathcal{C}}_\Delta} f^t_{\hat{\mathcal{C}}}(t,W) \leq 0.$, and $\hat{\mathcal{C}}_\Delta$ does not not satisfy (22). In other words,

$$\mathcal{A}^{\tilde{\mathcal{C}}} V_{\hat{\mathcal{C}}}(t,W) + (\tilde{a}_t - \tilde{c}_t) - \mathcal{A}^{\tilde{\mathcal{C}}} f_{\hat{\mathcal{C}}}(t,W,t) + \mathcal{A}^{\tilde{\mathcal{C}}} f^t_{\hat{\mathcal{C}}}(t,W) \leq 0, \quad \forall (t,W)\ .$$

for any other contract $\tilde{\mathcal{C}} \neq \hat{\mathcal{C}}$,

Finally, we verify that $\mathcal{C}$ solves the extended HJB system. By construction, $\mathcal{A}^{\hat{\mathcal{C}}} V(t,W) = \mathcal{A}^{\hat{\mathcal{C}}} f_{\hat{\mathcal{C}}}(t,W,t)$, and $(\hat{a}_t - \hat{c}_t) + \mathcal{A}^{\hat{\mathcal{C}}} f^t_{\hat{\mathcal{C}}}(t,W) = 0$. Therefore,

$$\mathcal{A}^{\hat{\mathcal{C}}} V(t,W) + (\hat{a}_t - \hat{c}_t) - \mathcal{A}\hat{\mathcal{C}} f_{\hat{\mathcal{C}}}(t,W,t) + \mathcal{A}^{\hat{\mathcal{C}}} f^t_{\hat{\mathcal{C}}}(t,W) = 0.$$

This completes the proof of the second statement in Proposition 6.

A.5 Time-difference Discounting Functions

A.5.1 Proof of Lemma 2

The proof is analogous to the proof in Appendix A of Marinovic and Varas (2019) and is hence omitted.

A.5.2 Proof of Proposition 3

First, $u(c,a) = rW$ (from Lemma 2) implies that

$$c = \frac{k}{2} a^2 - \frac{\ln(\eta\gamma)}{\eta} - \frac{1}{\eta}\ln(-W). \tag{51}$$

Next, notice (also from Lemma 2) that dW_t is a martingale with no drift. Therefore,

$$\mathcal{A}^{\mathcal{C}} V = \frac{1}{2}\left(a_t k \gamma \eta W \sigma\right)^2 V_{WW}. \tag{52}$$

Replacing $\mathcal{A}^{\mathcal{C}}V$ and c_s in the single HJB equation (25) with (52) and (51) implies that the value function $V(t, W)$ solves

$$0 = \sup_{a_t} V_t + a_t - \left[\frac{k}{2}a_t^2 - \frac{\ln(\eta\gamma)}{\eta} - \frac{1}{\eta}\ln(-W)\right] + \frac{1}{2}(a_t k\gamma\eta W\sigma)^2 V_{WW}$$
$$+ \int_t^T R'(s-t)\left(a_s - \left[\frac{k}{2}a_s^2 - \frac{\ln(\eta\gamma)}{\eta} - \frac{1}{\eta}\ln(-W_s)\right]\right) ds + \frac{\ln(-W_T)}{\eta}R'(T-t),$$

subject to (26), the incentive compatibility condition (28), and the boundary condition:

$$V(T, W) = B(T, W) \text{ for all } W. \tag{53}$$

Conjecture and verify that the principal's value function has the following functional form: $V(t, W) = A_t \ln(-W) + B_t$, with the boundary condition $A_T = \frac{1}{\eta}$. Given the conjecture, $\dot{V}_t = \dot{A}_t \ln(-W) + \dot{B}_t, V_W = \frac{A_t}{W}, V_{WW} = -\frac{A_t}{W^2}$. Substituting these back into the extended HJB system yields: $a_t = \frac{1}{k+k^2\eta^2\gamma^2\sigma^2 A_t}$. Meanwhile, collecting the log terms in the value function and using $R(0) = 1$ yields

$$\dot{A}_t = \frac{1}{\eta}\left(R(T-t) + R'(T-t)\right). \tag{54}$$

The boundary condition $A_T = \frac{1}{\eta}$ implies

$$A_t = \frac{1}{\eta}\int_0^t \left(R(T-s) + R'(T-s)\right) ds + K, \tag{55}$$

where $K = \frac{1}{\eta}(2 - \int_0^T R(T-s)ds - R(T))$. Plugging this back in to a_t yields:

$$a_t = \frac{1}{k + K + \eta k^2\gamma^2\sigma^2\left[\int_0^t \left(R(T-s) + R'(T-s)\right) ds\right]}. \tag{56}$$

A.6 Proofs for the Case of Quasi-Hyperbolic Discounting

Without the loss of generality we assume $k = 1$. The proof for $k \neq 1$ is identical in spirit. First, substituting R and R' with the quasi-hyperbolic discounting function to (54) yields:

$$\dot{A}_t = \frac{1}{\eta}\left(\beta(1-\rho)e^{-\rho(T-t)} + (1-\beta)(1-(\rho+\lambda))e^{-(\rho+\lambda)(T-t)}\right), \tag{57}$$

with boundary condition $A_T = \frac{1}{\eta}$. Integrating the above equation yields:

$$A_t = \frac{1}{\eta}\left(\frac{\beta(1-\rho)}{\rho}e^{-\rho(T-t)} + (1-\beta)\,\frac{1-(\rho+\lambda)}{\rho+\lambda}e^{-(\rho+\lambda)(T-t)} + \frac{2\rho(\rho+\lambda)-(\beta\lambda+\rho)}{\rho(\rho+\lambda)}\right).$$

Thus equilibrium action is thus given by:

$$a_t = \frac{1}{1+\eta^2\gamma^2\sigma^2\left(\frac{1}{\eta}\left(\frac{\beta(1-\rho)}{\rho}e^{-\rho(T-t)} + (1-\beta)\,\frac{1-(\rho+\lambda)}{\rho+\lambda}e^{-(\rho+\lambda)(T-t)} + \frac{2\rho(\rho+\lambda)-(\beta\lambda+\rho)}{\rho(\rho+\lambda)}\right)\right)}.$$

A.6.1 Proof of Proposition 4

First, note that $\psi_t = a_t = \frac{1}{1+\eta^2\gamma^2\sigma^2 A_t}$. Hence $\frac{\partial\psi_t}{\partial\beta}$ has the opposite sign of $\frac{\partial A_t}{\partial\beta}$, and the latter can be solved explicitly:

$$\begin{aligned}\frac{\partial A_t}{\partial\beta} &= \frac{1}{\eta}\left(\frac{(1-\rho)}{\rho}e^{-\rho(T-t)} - \frac{1-(\rho+\lambda)}{\rho+\lambda}e^{-(\rho+\lambda)(T-t)} + \frac{-\lambda}{\rho(\rho+\lambda)}\right),\\ &= -\frac{((\rho+\lambda)(\rho-1))(e^{-\rho(T-t)} - e^{-(\rho+\lambda)(T-t)}) + \lambda(1-e^{-(\rho+\lambda)(T-t)})}{\rho(\rho+\lambda)}.\end{aligned}$$

The sign of the numerator depends on the parameter values. We discuss their different scenarios below:

Case 1 $\rho+\lambda \geq 1$. There are two sub-cases in this scenario: i): $\rho \geq 1$, ii)$\rho < 1$. In subcase i) it is easy to see that $((\rho+\lambda)(\rho-1))(e^{-\rho(T-t)} - e^{-(\rho+\lambda)(T-t)}) + \lambda(1-e^{-(\rho+\lambda)(T-t)}) \geq 0$. Thus $\frac{\partial A_t}{\partial\beta} \leq 0$ for all t and W_t. In sub-case ii), observe that $0 < e^{-\rho(T-t)} - e^{-(\rho+\lambda)(T-t)} < 1-e^{-(\rho+\lambda)(T-t)}$. Therefore the sign of the numerator is positive if $\lambda > (\lambda+\rho)|\rho-1|$, which is equivalent to $\lambda \geq -\lambda\rho+\lambda-\rho^2+\rho$. This condition simplifies to $\rho+\lambda \geq 1$ which is our initial assumption. This implies $\frac{\partial A_t}{\partial\beta} \leq 0$ for all t and W_t.

Case 2 $\rho+\lambda < 1$. Re-arranging the numerator yields

$$\frac{e^{-\rho(T-t)} - e^{-(\rho+\lambda)(T-t)}}{1-e^{-(\rho+\lambda)(T-t)}} + \frac{\lambda}{(\rho+\lambda)(1-\rho)}. \tag{58}$$

When $\rho+\lambda < 1$, both of terms are always less than 1. In the first term and for any T, one of the following must be true: if $\frac{e^{-\rho(T)}-e^{-(\rho+\lambda)(T)}}{1-e^{-(\rho+\lambda)(T)}} < \frac{\lambda}{(\rho+\lambda)(1-\rho)}$, then $\frac{\partial A_t}{\partial\beta}$ is always negative for all t and all W_t; if $\frac{e^{-\rho(T)}-e^{-(\rho+\lambda)(T)}}{1-e^{-(\rho+\lambda)(T)}} > \frac{\lambda}{(\rho+\lambda)(1-\rho)}$, then there exists some $t < T$ such that for all $t' \leq t$ and all $W_{t'}$, $\frac{\partial A_{t'}}{\partial\beta} \geq 0$ and for all $t'' > t$ and $W_{t''}$ $\frac{\partial A_{t''}}{\partial\beta} < 0$.

To prove the existence of $T(\lambda,\rho)$, observe that $\frac{e^{-\rho(T)}-e^{-(\rho+\lambda)(T)}}{1-e^{-(\rho+\lambda)(T)}} \to 0$ as $T\to\infty$, and $\to \lambda/(1+\lambda+\rho)$ (following L'hopital's Rule) as $\to 0$ with a derivative that does not change sign. Thus, for each ρ and λ, there exists a unique $T(\lambda, \rho)$ large enough such that (58) is equal to 0.

Similarly, given the equilibrium wage $c_t = \frac{1}{2}a_t^2 - \frac{\ln(\gamma\eta)}{\eta} - \frac{1}{\eta}\ln(-W_t)$. With a slight abuse of notation, let $a_t(\beta)$ denote the optimal a_t as a function of β, we have

$$\frac{\partial c_t}{\partial \beta} = 2a_t(\beta)\frac{\partial a_t}{\partial \beta}.$$

Observe that $a_t(\beta) \geq 0$ for all β. Thus, $dc/d\beta$ has the same sign as $\frac{a_t}{\beta}$ which has the opposite sign of A_t. Therefore, c_t and A_t must have opposite signs.

A.6.2 Proof of Proposition 5

Part 1) $\beta = 1$ implies $\dot{A}_t = \frac{1}{\eta}(\rho + \lambda - 1)$, and $\beta = 0$ implies $\dot{A}_t = \frac{1}{\eta}(\rho - 1)$, both equivalent to time-consistent solutions.

Part 2) Fix any $t < T$, $\lim_{T\to\infty} a_t = \frac{1}{1+\eta^2\gamma^2\sigma^2\frac{1}{\eta}\left(\frac{2\rho(\rho+\lambda)-(\beta\lambda+\rho)}{\rho(\rho+\lambda)}\right)}$. which corresponds to the solution in Holmstrom and Milgrom (1987) with a discount rate of $\frac{\rho(\lambda+\rho)}{\beta\lambda+\rho}$.

Parts 3) Recall that the closed-form solution for a_t is:

$$a_t = \frac{1}{1 + \eta^2\gamma^2\sigma^2\frac{1}{\eta}\left(\frac{\beta(1-\rho)}{\rho}e^{-\rho(T-t)} + (1-\beta)\frac{1-(\rho+\lambda)}{\rho+\lambda}e^{-(\rho+\lambda)(T-t)} + \frac{2\rho(\rho+\lambda)-(\beta\lambda+\rho)}{\rho(\rho+\lambda)}\right)}.$$

The dynamics of a_t are pinned down by the term $\beta\frac{(1-\rho)}{\rho}e^{-\rho(T-t)} + (1-\beta)\frac{1-(\rho+\lambda)}{\rho+\lambda}e^{-(\rho+\lambda)(T-t)}$. It's time derivative equals to

$$\beta(1-\rho)e^{-\rho(T-t)} + (1-\beta)(1-(\rho+\lambda))e^{-(\rho+\lambda)(T-t)}. \tag{59}$$

Clearly, if the sign of (59) changes as time changes, it must happen at a unique point. In particular, for t small enough the first term dominates and (59) is always positive. For t large enough, the second term dominates, and if $(1-(\rho-\lambda))(1-\beta)$ is negative, (59) can be negative.

A.7 Equilibrium Regularity

In this section we show that, under minimal additional assumptions, our equilibrium concept can be strengthened to a regular equilibrium as in He and Jiang (2019): any local deviation to another contract, however small, is not profitable.[35] As shown in Lemma 4, we have

$$\liminf_{\Delta \to 0} \frac{f_{\mathcal{C}}(t, W, t) - f_{\mathcal{C}_\Delta}(t, W, t)}{\Delta} = (a_t - c_t) - (a_t^\Delta - c_t^\Delta) - \mathcal{A}^{\mathcal{C}_\Delta} f_{\mathcal{C}}^t(t, W) + \mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}^t(t, W).$$

To apply Theorem 5.2 part iii) of He and Jiang (2019), it is sufficient to show that the minimizer for the *inf* is unique.[36] First, focus on the $\mathcal{A}^{\mathcal{C}_\Delta} f_{\mathcal{C}}^t(t, W) + \mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}^t(t, W)$ terms. Recall that the agents equilibrium continuation payoff W is the unique adapted solution the BSVIE identified by:

$$W(t) = r^t(T)u(c_T, \underline{a}_T) - \int_t^T \left(\frac{a_s}{\sigma} \tilde{\psi}(s, t) + r^t(s)u(c_s, a_s) \right) ds + \int_t^T \tilde{\psi}(t, s) dZ_s^0.$$

Following Definition 4, let $\mathcal{C}_\Delta$ denote a small deviation during some time $[\tau, \tau + \Delta]$. The associated BSVIE satisfies

$$W_\Delta(t) = \left(r^t(T)u(c_T, \underline{a}_T) \right)_\Delta - \int_t^T \left(\frac{a_s}{\sigma} \tilde{\psi}(s, t) + r^t(s)u(c_s, a_s) \right)_\Delta ds + \int_t^T \tilde{\psi}(t, s)_\Delta dZ_s^0$$

where

$$\begin{aligned} \left(r^t(T)u(c_T, \underline{a}_T) \right)_\Delta &= \left(r^t(T)u(c_T^\Delta, \underline{a}_T) \right) I_{[\tau, \tau+\Delta]} + \left(r^t(T)u(c_T, \underline{a}_T) \right) I_{[\tau+\Delta, T]}, \\ \left(\frac{a_s}{\sigma} \tilde{\psi}(s, t) + r^t(s)u(c_s, a_s) \right)_\Delta &= \left(\frac{a_s^\Delta}{\sigma} \tilde{\psi}(s, t) + r^t(s)u(c_s^\Delta, a_s^\Delta) \right) I_{[\tau, \tau+\Delta]} \\ &\quad + \left(\frac{a_s}{\sigma} \tilde{\psi}(s, t) + r^t(s)u(c_s, a_s) \right) I_{[\tau+\Delta, T]} \end{aligned}$$

and I denotes the indicator function. By Proposition 2.2 of Wang and Yong (2019a),

$$\begin{aligned} W_\Delta(t) &= W(t) \quad \tau + \Delta < t \leq T, \\ \tilde{\psi}_\Delta(s, t) &= \tilde{\psi}(s, t) \quad \tau + \Delta < t \leq s \leq T \\ W_\Delta(t) &= \mathbf{W}_\Delta(t) \quad \tau \leq t \leq \tau + \Delta, \\ \tilde{\psi}_\Delta(s, t) &= \tilde{\boldsymbol{\psi}}_\Delta(s) \quad \tau \leq t \leq \tau + \Delta, t \leq s \leq T \end{aligned}$$

[35] See Section 3.1 for the formal definition of a regular equilibrium in our setup.

[36] Theorem 5.2 of He and Jiang (2019) refers to a maximizer as they subtract the equilibrium payoff from the deviation payoff. Instead, we subtract the latter from the former which yields a minimum.

where $\mathbf{W}_\Delta(t)$ is the adapted solution to the following BSDE

$$\begin{aligned}\mathbf{W}_\Delta(t) &= r^t(T)u(c_T, \underline{a}_T)\\ &- \int_t^T \left(\left(\frac{a_s^\Delta}{\sigma} \tilde{\boldsymbol{\psi}}_\Delta(s) + r^t(s)u(c_s^\Delta, a_s^\Delta) \right) I_{[\tau,\tau+\Delta]} + \left(\frac{a_s}{\sigma} \tilde{\boldsymbol{\psi}}_\Delta(s) + r^t(s)u(c_s, a_s) \right) I_{[\tau+\Delta,T]} \right) ds\\ &+ \int_t^T \tilde{\boldsymbol{\psi}}_\Delta(s) dZ_s^0.\end{aligned}$$

Note that even though the deviation equilibrium value process is a BSVIE, it can be represented by a backward stochastic differential equation during the Δ-interval. In particular, since $r^t(t) = 1$, we have

$$d\mathbf{W}_\Delta(t) = -\left(\frac{a_t^\Delta}{\sigma} \tilde{\boldsymbol{\psi}}_\Delta(t) + u(c_t^\Delta, a_t^\Delta) \right) dt + \tilde{\boldsymbol{\psi}}_\Delta(t) dZ_t^0. \tag{60}$$

for $t \in [\tau, \tau + \Delta]$. Now, returning to $(a_t - c_t) - (a_t^\Delta - c_t^\Delta) - \mathcal{A}^{\mathcal{C}_\Delta} f_{\mathcal{C}}^t(t, W) + \mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}^t(t, W)$, we use the identity in Equation (60) where the deviation is from the original equilibrium contract. Here, instead of canceling $(a_t - c_t) + \mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}^t(t, W)$, we can cancel the t derivatives in $\mathcal{A}^{\mathcal{C}_\Delta} f_{\mathcal{C}}^t(t, W)$ and $\mathcal{A}^{\mathcal{C}} f_{\mathcal{C}}^t(t, W)$. Also observe that the diffusion terms cancel each other as well, since they are pinned down by $\boldsymbol{\psi}$, which is the same in both contracts. Then, suppressing the Δ subscript for the equilibrium contract, we are left with the following:

$$\begin{aligned}\liminf_{\Delta \to 0} \frac{f_{\mathcal{C}}(t, W, t) - f_{\mathcal{C}_\Delta}(t, W, t)}{\Delta} &= (a_t - c_t) - (a_t^\Delta - c_t^\Delta)\\ &+ \left(\frac{a_t^\Delta}{\sigma} \tilde{\boldsymbol{\psi}}_\Delta(t) - \frac{a_t}{\sigma} \tilde{\boldsymbol{\psi}}(t) + u(c_t^\Delta, a_t^\Delta) - u(c_t, a_t) \right) f_{\mathcal{C},W}^t(t, W).\end{aligned}$$

Taking the first order derivatives with respect to a_t^Δ and c_t^Δ and setting them equal to 0 yields the following condition

$$\frac{\tilde{\boldsymbol{\psi}}_\Delta(t)}{\sigma} + u_a(c_t^\Delta, a_t^\Delta) = -u_c(c_t^\Delta, a_t^\Delta).$$

The above minimization can have multiple solutions without the incentive compatibility condition. However, recall that the IC condition implies that $u_a(a_t^\Delta, c_t^\Delta) = \tilde{\boldsymbol{\psi}}_\Delta(t)$ for any deviation contract. Thus, under the assumption that $u_{ac}(\cdot, \cdot) > 0$, there is a unique solution that is both incentive compatible and is a deviation contract. Therefore by Theorem 5.2 part iii) of He and Jiang (2019), our equilibrium is also a regular equilibrium.

A.8 Anticipatory Utility

In this section, we explore the applicability of our main theorem in another example in the time-difference discount family: anticipatory utility. That is, in addition to the immediate utility, the principal's well-being also depends on her expectations of future utility. Such expectations may arise from either "savoring" or "dreading" future consumption as in Loewenstein (1987), or from optimism, as in Brunnermeier, Papakonstantinou, and Parker (2016).[37] To maintain tractability we retain the same assumptions made in Section 4 for the problem on the agent's side.

We model anticipatory utility following Loewenstein (1987). In addition to the current payoff, the principal derives an additional $e^{-\zeta t}$ amount of utility from payoffs she anticipates to receive at a future time t. Her discount of the future utility (both actual future payoffs and anticipation of more distant future payoffs) is ρ. Formally, we assume

Assumption A.5 *The principal values consumption streams by:*

$$E\left[\int_0^T e^{-\rho t}\left(\Pi_t + a_t - c_t\right) dt\right],$$

where Π_t *is given by*

$$\Pi_t = E\left[\int_t^T e^{-\zeta(s-t)}(a_s - c_s)ds\right].$$

Here, Π_t captures the utility from the anticipated future payoff.By the Law of Iterated Expectations and a change in the order of integration, the principal's valuation of a stream of payoffs starting from any period t can be re-written as

$$E^{\mathcal{C}}\left[\int_t^T \left(\frac{e^{-\rho(s-t)} - e^{-\zeta(s-t)}}{\zeta - \rho} + e^{-\rho(s-t)}\right)(a_s - c_s)ds\right].$$

As a result, the principal with the anticipatory utility can be understood as effectively having the following discounting function $R^t(s) = R(s-t) = \frac{e^{-\rho(s-t)} - e^{-\zeta(s-t)}}{\zeta-\rho} + e^{-\rho(s-t)}$. The first term captures the discounted anticipation: letting $\zeta \to \infty$ yields the standard discounting as the utility from anticipation disappears. We also make a technical assumption that $\zeta > \rho > 1$. The first inequality is necessary for transversality and the second is necessary to avoid corner solutions.[38]

[37] Other studies related to anticipatory utility include Caplin and Leahy (2001), Loewenstein, O'Donoghue, and Rabin (2003), Caplin and Leahy (2004), Takeuchi (2011), etc.

[38] Although not noted in Loewenstein (1987), an appropriate numerical relationship between ρ and ζ is economically crucial. Without further restrictions, utility of infinitely far future may still have a positive

Similar to the case of quasi-hyperbolic discounting, because $R(T-s)$ is an exponential function in s, the integral in (32) has a closed-form solution. Based on the closed-form solution, the implications of anticipatory utility can be summarized as follows:

Proposition 7 *Under Assumption A.5 the optimal renegotiation-proof contract derived in proposition 3 has the following properties:*

1. *If $\zeta = \infty$, the optimal contract is identical to that for a dynamically-consistent principal with discount rate ρ and optimal incentive power ψ_t and the equilibrium actions a_t are (weakly) monotonic in time t.*

2. *As the contracting horizon becomes infinitely long (i.e., $T \to \infty$), the optimal contract converges to that for a time consistent-principal with discount rate $\frac{\rho\zeta}{1+\zeta}$.*

3. *For any finite ρ and ζ, there exists $\tilde{T}(\rho,\zeta)$ such that the optimal incentive power ψ_t and the equilibrium actions a_t are non-monotonic in t as long as $\tilde{T}(\rho,\zeta) < T < \infty$.*

Proof: We first derive the equilibrium action path. Under assumption A.5, we can calculate how a payoff in period τ is valued in the stream $a_\tau - c_\tau$ as follows:

$$\int_t^\tau e^{-\rho(s-t)} e^{-\zeta(\tau-s)} E(a_\tau - c_\tau|\mathcal{F}_t) ds = \frac{e^{-\rho(\tau-t)} - e^{-\zeta(\tau-t)}}{\zeta - \rho} E(a_\tau - c_\tau|\mathcal{F}_t).$$

Therefore, anticipatory utility is equivalent to the following discounting $R^t(s) = R(s-t) = \frac{e^{-\rho(\tau-t)} - e^{-\zeta(\tau-t)}}{\zeta-\rho}$. Substituting in R and R' from anticipatory utility to equation (54) yields

$$\dot{A}_t = \frac{1}{\eta}\left(\frac{(1-\rho)e^{-\rho(T-t)} - (1-\zeta)e^{-\zeta(T-t)}}{\zeta-\rho} + (1-\rho)e^{-\rho(T-t)}\right).$$

Using the boundary condition $A_T = 1/\eta$ and integrating with respect to t implies:

$$A_t = \frac{1}{\eta}\left(\frac{(\frac{1}{\rho}-1)e^{-\rho(T-t)} - (\frac{1}{\zeta}-1)e^{-\zeta(T-t)}}{\zeta-\rho} + \frac{(1-\rho)}{\rho}e^{-\rho(T-t)} + \frac{2\rho\zeta - 1 - \zeta}{\rho\zeta}\right).$$

Thus, equilibrium action a_t is thus given by:

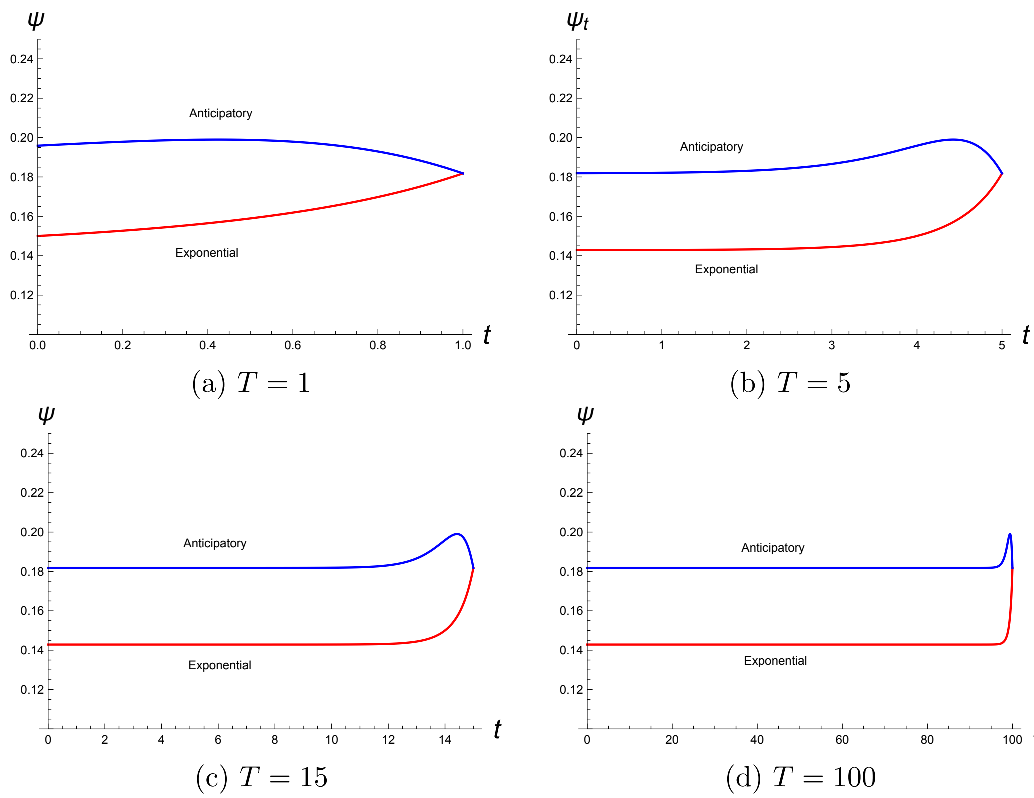
$$a_t = \frac{1}{1+\eta\gamma^2\sigma^2\left(\frac{(\frac{1}{\rho}-1)e^{-\rho(T-t)} - (\frac{1}{\zeta}-1)e^{-\zeta(T-t)}}{\zeta-\rho} + \frac{(1-\rho)}{\rho}e^{-\rho(T-t)} + \frac{2\rho\zeta-1-\zeta}{\rho\zeta}\right)}.$$

We now prove each statement in Proposition 7:

value today after discounting.

Part 1) Fix any $t < T$, then $\lim_{\zeta\to\infty} a_t = \frac{1}{1+\eta\gamma^2\sigma^2 \frac{2\rho-1+(1-\rho)e^{-\rho(T-t)}}{\rho}}$, corresponding to the solution in Holmstrom and Milgrom (1987) with a discount rate ρ.
Part 2) As $T \to \infty$, solving for r yields: $\frac{2r-1}{r} = \frac{2\rho\zeta-1-\zeta}{\rho\zeta} \Rightarrow r = \frac{\rho\zeta}{1+\zeta}$.
Part 3) The dynamics of a_t are pinned down by the term $\frac{(\frac{1}{\rho}-1)e^{-\rho(T-t)}-(\frac{1}{\zeta}-1)e^{-\zeta(T-t)}}{\zeta-\rho} + \frac{(1-\rho)}{\rho}e^{-\rho(T-t)}$; the time derivative equals: $\frac{(1-\rho)e^{-\rho(T-t)}-(1-\zeta)e^{-\zeta(T-t)}}{\zeta-\rho} + (1-\rho)e^{-\rho(T-t)}$. Observe that if the time derivative ever becomes 0 we must obtain $1 = \left(\frac{\zeta-1}{\rho-1}\right)\left(\frac{1}{1+\zeta-\rho}\right)e^{-(\zeta-\rho)(T-t)}$. Next, since we assume that $\zeta > \rho > 1$, we have $\frac{\zeta-1}{\rho-1}\frac{1}{1+\zeta-\rho} > 1$ and $1 \geq e^{-(\zeta-\rho)(T-t)} \geq e^{-(\zeta-\rho)T} > 0$. Observe that if T is small enough, the derivative is always negative. However, for any ζ, ρ if T is sufficiently large, there exists a $t^\star < T$ such that the derivative becomes negative for all $t > t^\star$.

We illustrate the above properties in Figure 3.



(a) $T = 1$ (b) $T = 5$ (c) $T = 15$ (d) $T = 100$

Figure 3: These plots illustrate the paths of the incentive power (also the paths of agent's actions given Proposition 3) under the optimal contract. Blue lines indicate the paths under anticipatory utility with $\eta = 2, \sigma = 1, \rho = 1.5, \zeta = 3, \gamma = 1.5$. Each plot corresponds to a different time horizon (different T).

The first property is straightforward: if $\zeta = \infty$, the principal receives no utility from anticipation. Her discounting is then exponential and the optimal contract follows that of Holmstrom and Milgrom (1987) with monotonic incentive power. The second property demonstrates that the effect of anticipation disappears if the time horizon becomes arbitrarily large. The intuition is identical to why the effect of quasi-hyperbolic discounting disappears: as the horizon increases, the problems faced by each of the principal's t-self become similar. When the horizon is infinitely long, the principal's problem becomes completely stationary. Thus, the solution converges to that of a dynamically consistent principal with a different but constant discount rate. The third property highlights the possibility of non-monotonic incentive power under a finite horizon, hence a "deadline effect". With a finite horizon, the anticipation of future payoff leads the principal to back-load incentive power. However, the exact dynamics depend on the length of the contracting horizon: if the horizon is sufficiently long, back-loading causes incentives to gradually increase in time in the early phases. As the end of the contracting horizon approaches, the effect of anticipation diminishes, causing incentives to decline over time.

References

Balder, E. J. (1991). On Cournot-Nash equilibrium distributions for games with differential information and discontinuous payoffs. *Economic Theory 1*(4), 339–354.

Benoit, J.-P. and V. Krishna (1993). Renegotiation in finitely repeated games. *Econometrica: Journal of the Econometric Society*, 303–323.

Bernheim, B. D. and D. Ray (1989). Collective dynamic consistency in repeated games. *Games and Economic Behavior 1*(4), 295–326.

Bernheim, B. D., D. Ray, and Ş. Yeltekin (2015). Poverty and self-control. *Econometrica 83*(5), 1877–1911.

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

Biais, B., T. Mariotti, J.-C. Rochet, and S. Villeneuve (2010). Large risks, limited liability, and dynamic moral hazard. *Econometrica 78*, 73–118.

Bisin, A., A. Lizzeri, and L. Yariv (2015). Government policy with time inconsistent voters. *American Economic Review 105*(6), 1711–37.

Björk, T., M. Khapko, and A. Murgoci (2017). On time-inconsistent stochastic control in continuous time. *Finance and Stochastics 21*(2), 331–360.

Bloedel, A., R. V. Krishna, and B. Strulovici (2020). Persistent private information revisited. Working paper.

Bolton, P., H. Chen, and N. Wang (2011). A Unified Theory of Tobin's q, Corporate Investment, Financing, and Risk Management. *Journal of Finance 66*, 1545–1578.

Bond, P. and G. Sigurdsson (2018). Commitment contracts. *Review of Economic Studies 85*(1), 194–222.

Brunick, G. and S. Shreve (2013). Mimicking an itô process by a solution of a stochastic differential equation. *The Annals of Applied Probability 23*(4), 1584–1628.

Brunnermeier, M. K., F. Papakonstantinou, and J. A. Parker (2016). Optimal time-inconsistent beliefs: Misplanning, procrastination, and commitment. *Management Science 63*(5), 1318–1340.

Cao, D. and I. Werning (2018). Saving and dissaving with hyperbolic discounting. *Econometrica 86*(3), 805–857.

Caplin, A. and J. Leahy (2001). Psychological expected utility theory and anticipatory feelings. *Quarterly Journal of Economics 116*(1), 55–79.

Caplin, A. and J. Leahy (2004). The social discount rate. *Journal of Political Economy 112*(6), 1257–1268.

Chade, H., P. Prokopovych, and L. Smith (2008). Repeated games with present-biased preferences. *Journal of Economic Theory 139*(1), 157–175.

Cisternas, G. (2018). Two-sided learning and the ratchet principle. *Review of Economic Studies 85*(1), 307–351.

Coles, J. L., N. D. Daniel, and L. Naveen (2006). Managerial incentives and risk-taking. *Journal of Financial Economics 79*(2), 431–468.

Core, J. and W. Guay (2002). Estimating the value of employee stock option portfolios and their sensitivities to price and volatility. *Journal of Accounting research 40*(3), 613–630.

Cvitanic, J. and J. Zhang (2012). *Contract Theory in Continuous-Time Models*. Springer Science & Business Media.

DellaVigna, S. and U. Malmendier (2004). Contract design and self-control: Theory and evidence. *Quarterly Journal of Economics 119*(2), 353–402.

DeMarzo, P., M. Fishman, Z. He, and N. Wang (2012). Dynamic Agency and the Q Theory of Investment. *Journal of Finance 67*.

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

Djehiche, B. and M. Huang (2016). A characterization of sub-game perfect equilibria for sdes of mean-field type. *Dynamic Games and Applications 6*(1), 55–81.

Duffie, D. and L. G. Epstein (1992). Stochastic differential utility. *Econometrica*, 353–394.

Dupire, B. et al. (1994). Pricing with a smile. *Risk 7*(1), 18–20.

Ekeland, I. and A. Lazrak (2010). The golden rule when preferences are time inconsistent. *Mathematics and Financial Economics 4*(1), 29–55.

El Karoui, N., S. Peng, and M. C. Quenez (1997). Backward stochastic differential equations in finance. *Mathematical finance 7*(1), 1–71.

Faingold, E. and Y. Sannikov (2011). Reputation in continuous-time games. *Econometrica 79*(3), 773–876.

Feng, F. Z. (2021). Financing a black box: Dynamic investment with persistent private information. Working paper.

Feng, F. Z. and M. Westerfield (2021). Dynamic Resource Allocation and Hidden Volatility. *Journal of Financial Economics 140*(2), 560–581.

Fudenberg, D., B. Holmstrom, and P. Milgrom (1990). Short-term contracts and long-term agency relationships. *Journal of Economic Theory 51*(1), 1–31.

Fudenberg, D. and J. Tirole (1990). Moral hazard and renegotiation in agency contracts. *Econometrica*, 1279–1319.

Galperti, S. (2015). Commitment, flexibility, and optimal screening of time inconsistency. *Econometrica 83*(4), 1425–1465.

Georgiadis, G. (2014). Projects and team dynamics. *Review of Economic Studies 82*(1), 187–218.

Gottlieb, D. (2008). Competition over time-inconsistent consumers. *Journal of Public Economic Theory 10*(4), 673–684.

Gottlieb, D. and X. Zhang (2020). Long-term contracting with time-inconsistent agents. *Econometrica*. Forthcoming.

Grenadier, S. R. and N. Wang (2007). Investment under uncertainty and time-inconsistent preferences. *Journal of Financial Economics 84*(1), 2–39.

Grubb, M. D. (2015). Overconfident consumers in the marketplace. *Journal of Economic Perspectives 29*(4), 9–36.

Gryglewicz, S. and B. Hartman-Glaser (2020). Investment timing and incentive costs. *Review of Financial Studies 33*(1), 309–357.

Gyöngy, I. (1986). Mimicking the one-dimensional marginal distributions of processes having an itô differential. *Probability theory and related fields 71*(4), 501–516.

Halac, M. and P. Yared (2014). Fiscal rules and discretion under persistent shocks. *Econometrica 82*(5), 1557–1614.

Harris, C. and D. Laibson (2012). Instantaneous gratification. *Quarterly Journal of Economics 128*(1), 205–248.

Harstad, B. (2016). Technology and time inconsistency. Working paper.

He, X. D. and Z. L. Jiang (2019). On the equilibrium strategies for time-inconsistent problems in continuous time. *Available at SSRN 3308274*.

He, Z. (2011). A model of dynamic compensation and capital structure. *Journal of Financial Economics 100*(2), 351–366.

He, Z., B. Wei, J. Yu, and F. Gao (2017). Optimal long-term contracting with learning. *Review of Financial Studies 30*(6), 2006–2065.

Heidhues, P. and B. Kőszegi (2010). Exploiting naivete about self-control in the credit market. *American Economic Review 100*(5), 2279–2303.

Hernández, C. and D. Possamaï (2021). A unified approach to well-posedness of type-i backward stochastic volterra integral equations. *Electronic Journal of Probability 26*, 1–35.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *Review of Economic Studies 66*(1), 169–182.

Holmstrom, B. and P. Milgrom (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica: Journal of the Econometric Society*, 303–328.

Jackson, M. O. and L. Yariv (2014). Present bias and collective dynamic choice in the lab. *American Economic Review 104*(12), 4184–4204.

Jackson, M. O. and L. Yariv (2015). Collective dynamic choice: the necessity of time inconsistency. *American Economic Journal: Microeconomics 7*(4), 150–78.

Kolb, A. and E. Madsen (2019). Leaks, sabotage, and information design. Working paper.

Kőszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature 52*(4), 1075–1118.

Krusell, P. and A. A. Smith, Jr (2003). Consumption–savings decisions with quasi–geometric discounting. *Econometrica 71*(1), 365–375.

Kydland, F. E. and E. C. Prescott (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy 85*(3), 473–491.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics 112*(2), 443–478.

Li, H., C. Mu, and J. Yang (2016). Optimal contract theory with time-inconsistent preferences. *Economic Modelling 52*, 519–530.

Li, X., Y.-H. Ni, and J.-F. Zhang (2017). On time-consistent solution to time-inconsistent linear-quadratic optimal control of discrete-time stochastic systems.

Lindensjö, K. (2019). A regular equilibrium solves the extended hjb system. *Operations Research letters 47*(5), 427–432.

Lindensjö, K. (2019). A regular equilibrium solves the extended hjb system. *Operations Research Letters 47*(5), 427–432.

Liu, B., C. Mu, and J. Yang (2017). Dynamic agency and investment theory with time-inconsistent preferences. *Finance Research Letters 20*, 88–95.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal 97*(387), 666–684.

Loewenstein, G., T. O'Donoghue, and M. Rabin (2003). Projection bias in predicting future utility. *Quarterly Journal of Economics 118*(4), 1209–1248.

Marinovic, I. and F. Varas (2019). CEO horizon, optimal pay duration, and the escalation of short-termism. *Journal of Finance 74*(4), 2011–2053.

O'Donoghue, T. and M. Rabin (1999). Doing it now or later. *American Economic Review 89*(1), 103–124.

Pardoux, E. and S. Peng (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications*, pp. 200–217. Springer.

Pollak, R. A. (1968). Consistent planning. *Review of Economic Studies 35*(2), 201–208.

Ray, D. (1994). Internally renegotiation-proof equilibrium sets: Limit behavior with low discounting. *Games and Economic Behavior 6*(1), 162–177.

Ray, D., N. Vellodi, and R. Wang (2017). Backward discounting. Technical report, Working Paper, New York University.[1546].

Rivera, A. (2022). Contracting with a present-biased agent: Sannikov meets laibson. *Available at SSRN 4009513*.

Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *Review of Economic Studies 75*, 957–984.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies 23*(3), 165–180.

Strulovici, B. (2020). Renegotiation-proof contracts with persistent private information. Working paper.

Strulovici, B. and M. Szydlowski (2015). On the smoothness of value functions and the existence of optimal strategies in diffusion models. *Journal of Economic Theory 159*, 1016–1055.

Sun, J., X. Li, and J. Yong (2016). Open-loop and closed-loop solvabilities for stochastic linear quadratic optimal control problems. *SIAM Journal on Control and Optimization 54*(5), 2274–2308.

Takeuchi, K. (2011). Non-parametric test of time consistency: Present bias and future bias. *Games and Economic Behavior 71*(2), 456–478.

Thaler, R. H. and S. Benartzi (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy 112*(S1), S164–S187.

Wang, H. and J. Yong (2019a). Time-inconsistent stochastic optimal control problems and backward stochastic volterra integral equations.

Wang, H. and J. Yong (2021). Time-inconsistent stochastic optimal control problems and backward stochastic volterra integral equations. *ESAIM: Control, Optimisation and Calculus of Variations 27*, 22.

Wang, T. and J. Yong (2019b). Backward stochastic volterra integral equations—representation of adapted solutions. *Stochastic Processes and their Applications*.

Wang, Y., W. Huang, B. Liu, and X. Zhang (2020). Optimal effort in the principal-agent problem with time-inconsistent preferences. *The North American Journal of Economics and Finance 52*.

Watson, J., D. Miller, and T. E. Olsen (2019). Relational contracting, negotiation, and external enforcement.

Wei, Q., J. Yong, and Z. Yu (2017). Time-inconsistent recursive stochastic optimal control problems. *SIAM Journal on Control and Optimization 55*(6), 4156–4201.

Williams, N. (2011). Persistent private information. *Econometrica 79*(4), 1233–1275.

Williams, N. (2015). A solvable continuous time dynamic principal–agent model. *Journal of Economic Theory 159*, 989–1015.

Yan, W. and J. Yong (2019). *Time-Inconsistent Optimal Control Problems and Related Issues*, pp. 533–569. Cham: Springer International Publishing.

Yong, J. (2012). Time-inconsistent optimal control problems and the equilibrium HJB equation. *arXiv preprint arXiv:1204.0568*.

Yong, J. and X. Y. Zhou (1999). *Stochastic controls: Hamiltonian systems and HJB equations*, Volume 43. Springer Science & Business Media.

Zhang, X. and X. Li (2018). Open-loop and closed-loop solvabilities for stochastic linear quadratic optimal control problems of markov regime-switching system. *arXiv preprint arXiv:1809.01891*.