Contracting with Non-Exponential Discounting: Moral Hazard and Dynamic Inconsistency*

Doruk Cetemen[†] Felix Zhiyu Feng[‡] Can Urgun[§]

August 2019

Abstract

We develop a framework for dynamic moral hazard problems with dynamic inconsistencies resulting from general, non-exponential discount functions. We derive the principal-optimal contract as a Markov perfect Nash equilibrium of the game played between the agent's and the principal's future selves. Such contract exists even when both contracting parties have dynamically inconsistent discount functions, and can be characterized via a system of differential equations rather than the classical Hamilton-Jacobi-Bellman equation. We demonstrate the applicability of our framework by solving two examples in closed form: one with quasi-hyperbolic discounting and one with anticipatory utility.

Keywords: Continuous-time contracting, dynamic inconsistency, HJB system, non-atomic games

JEL code: D81, D86, D91

*We would like to thank Paulo Barelli, Alessandro Bonatti, Jess Benhabib, Gonzalo Cisternas, Dino Gerardi, Hari Govindan, Johannes Hörner, Ryota Iijima, Roberto Marfè, Bruno Strulovici, Leeat Yariv, Stephen Morris, Juuso Toikka, Pietro Ortoleva, Rohit Lamba, Romans Pancs and Debraj Ray for discussions, and seminar participants at Bilkent University, Collegio Carlo Alberto, London School of Economics, Princeton University, Queen Mary University of London, University of Bristol, University of Warwick, Yale University, conference participants at NASMES 2019, ESSET(Gerzensee) 2019, SAET 2019, and SITE 2019 for helpful comments.

[†]Collegio Carlo Alberto. Email: doruk.cetemen@carloalberto.org

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

[§]Princeton University. Email: curgun@princeton.edu

1 Introduction

In a dynamic principal-agent contracting relationship, how each party discounts his or her future returns to their actions critically shapes the outcome of the contract.[1] Most existing dynamic contracting models assume that contracting parties have exponential discounting. This is analytically convenient but potentially restrictive, because exponential discounting implies that the discount rate between two adjacent periods is constant regardless of how distant into the future those periods are, and calendar dates are mostly irrelevant.

We relax the exponential discounting assumption and develop a framework for dynamic moral hazard problems under general discounting. We model discounting between two points in time through a bi-function that takes two dates as inputs and yields the relative ratio of one util as the output. This generalization captures many cases that are widely applicable in economic research and in practice, such as *present bias* (e.g. Ainslie (1975), Thaler (1981) Jackson and Yariv (2015)) or *anticipatory utility* (e.g. Loewenstein (1987), Caplin and Leahy (2001), Brunnermeier et al. (2016)), as well as inherent value attachment to specific dates (such as birthdays, anniversaries, etc).

Despite the extensive evidence – anecdotal and empirical – that "time preferences can be non-exponential" (Laibson (1997)), the formal analysis of problems with non-exponential discounting is known to be generally difficult because of the resulting *dynamic inconsistencies*. Under dynamic inconsistency, future plans that look "optimal" at present may ultimately be suboptimal. In particular, if decision makers are given a chance to re-evaluate their plan, they might do so continuously instead of following the original plan made before. In this paper we assume that parties are sophisticated about both their own and their counterparties' dynamic inconsistencies and only make plans that will be followed by their future "selves". This approach to resolve dynamic inconsistency is akin to the seminal works of Strotz (1955) and Pollak (1968) and has been widely adopted since (e.g. Ekeland and Pirvu (2008), Ekeland and Lazrak (2010), Björk et al. (2017)). We follow a similar approach and our framework characterizes a principal optimal *equilibrium contract* as the Markov perfect

[1] In DeMarzo and Sannikov (2006), for example, the difference in discount rates between the principal and the agent determines when payments to the agent are actually made. Lehrer and Pauzner (1999), Ray (2002), and Opp and Zhu (2015) demonstrate various effects of differential discount factors in dynamic games.

Nash equilibrium of the game played between the agent, the principal, and their future selves. We follow a similar approach and characterize a principal optimal *equilibrium contract* as the Markov perfect Nash equilibrium of the game played between the agent, the principal, and their future selves.

The contribution of this paper is three-fold: first, we establish the existence of the principal optimal contract and provide a characterization of such contract. The characterization reveals a somewhat surprising result: compared to the case in which discounting is exponential, the incentive compatibility condition for the agent does not change qualitatively, whether one or both contracting parties have general non-exponential discounting and dynamically inconsistent preferences. In particular the principal provides the incentives via a local pay-performance-sensitivity to each self instead of relying on equilibrium dynamics between different selves. Of course, how each party and their different selves *value* a given set of incentives still varies according to their respective discounting functions, resulting in different optimal contracts. In other words, the impact of non-exponential discounting is consequential for the principal's problem but inconsequential for the agent's problem. Second, using our characterization, we explicitly solve the optimal contract under two special discount functions that have been widely studied in economic research: quasi-hyperbolic preferences (as in Harris and Laibson (2012)), and anticipatory preferences (as in Loewenstein (1987)). The closed form of the contracts enable us to provide some comparative statics and highlight the specific effects of the form of discounting on contract outcomes. Third, we make a methodological contribution by providing a novel approach for tackling the existence of the dynamically inconsistent equilibrium contract. We believe our approach can shed light on the existence of equilibrium strategy for other dynamically inconsistent stochastic control problems beyond contracting.

To balance the richness of economic insights with analytical tractability, we cast our model in continuous time. Although the argument for contract existence would be simpler under discrete time due to backward induction, the contract characterization quickly becomes complicated and even intractable for problems that involve multiple periods.[2] In particular,

[2]As noted in Cao and Werning (2016) even simple dynamic savings games in discrete time may be quite ill-behaved. Hence, "these models are relatively intractable and make it difficult to characterize equilibria". See Peleg and Yaari (1973) for examples of non-existence of Markov perfect equilibrium and Bernheim and

our explicitly solved examples reinforce the importance of the contracting horizon in dynamic inconsistent problem (as previously highlighted in Gottlieb (2008) and Gottlieb and Zhang (2018)). Thus, the continuous-time framework, which allows for an arbitrary length of the contracting horizon, enables us to demonstrate several economic insights that are difficult to obtain from discrete-time models.

We also relax the assumption of full commitment power adopted in most existing contracting models. In the case of full commitment, dynamic inconsistency of the principal has little impact on the structure of the contract as only the time 0 "self" is relevant. In practice though, contracts that are supposed to involve long-term commitments are usually embedded with clauses that permit certain degree of alterations, such as the commonly observed *refinancing* of loans. During refinancing, the total amount of the loan often remains the same but the lender and the borrower agree to a different schedule of repayments.[3] In light of these observations, we make a minimal departure from full commitment by allowing the principal to offer an alternative contract at any point in time, subject to the agent having the same continuation value from the newly proposed contract. We call such an offer a *restructuring* as it does not change the continuation payoff of the agent, only how that continuation payoff is delivered. In other words, restructuring implies commitment to continuation payoff but flexibility in delivery.[4] Such an extension enables us to explore the intrapersonal conflict of the principal and how it impacts the structure of the contract while simultaneously staying close to the canonical long-term commitment setup. Indeed, if both parties are dynamically consistent, restructuring will never be utilized and the optimal equilibrium contract becomes identical to the one under full commitment.

The equilibrium optimal contract is characterized by a system of non-linear partial differential equations, instead of the classical Hamilton-Jacobi-Bellman (HJB) equation. This is

Ray (1986) for the existence of Markov Perfect equilibrium in discrete time consumption models.

[3]A typical example is mortgage refinancing, which usually involves changes in the length and/or the interest rate but not the total amount due. See Mian and Santos (2018), Xu (2018)) for empirical evidence on the prevalence of refinancing.

[4]Restructuring is a limit form of renegotiation because the principal can *commit* to the value of her promises. Full renegotiation allows for contracting parties to change the promised continuation utility as well as its delivery, which removes commitment completely and is beyond the scope of this paper. Our notion is closer to Miller et al. (2018) in the sense of having a *disagreement point* as the previous contract in the contractual equilbrium explored in their paper.

because the Bellman optimality principle applicable in dynamically consistent benchmarks (e.g. Sannikov (2008)) is no longer valid. In particular, an individual's self tomorrow does not solve the continuation problem of the self today. Thus, a system of equations that captures both the different selves' valuations and the equilibrium of the game between them is necessary. Such a system is called an *extended HJB system*, which consists of a forward part and a backward part. The forward part is akin to the HJB equation representing the current self solving an otherwise standard maximization problem but with additional terms – particularly, terms that capture the "equilibrium value" of the game between the selves and the effect the current self on different selves. The backward part captures the best responses of all the different selves in a "Stackelberg" fashion and feed into the forward part through those additional terms. On the agent side, the game between the selves hinges critically on the incentive-compatibility conditions. However, due to the lack of bargaining power and the impermanence of actions, the contract only need to provide incentive to each self via a *local* pay-performance-sensitivity similar to that in Sannikov (2008). Surprisingly, the evolution pattern of each selves' continuation utility turns out to be similar to the dynamically consistent benchmark as well, albeit the "equilibrium value" of the continuation utility must be different, because both the self and the time are changing simultaneously. Importantly, we show that the equilibrium value is also a process and can be utilized as a relevant state variable for the design of the principal-optimal equilibrium contract.

The basic idea behind how the extended HJB system captures the solution to dynamic inconsistent problems is economically intuitive and has been recently studied in more general stochastic control settings such as Yong (2012), Wang and Wu (2015), Lindensjö (2016), Björk et al. (2017), Wei et al. (2017), etc. However, "[t]he task of proving existence and/or uniqueness of solutions to the extended HJB system seems (...) to be technically extremely difficult" (Björk et al. (2017)). Indeed, absent from specially designed utility/cost functions or other very restrictive assumptions, the existence of solutions to extended HJB systems is an open question. We introduce a novel methodology that proves the existence of the solution to a broad class of problems with dynamic inconsistencies arising from non-exponential discounting. Instead of trying to establish the well-posedness and compatibility of the components of the extended HJB system, we translate the system into an appropriate static game of incomplete information between non-atomic players (à la Schmeidler (1973), Mas-Colell (1984), Balder (1991)), and prove the existence of the equilibrium of this non-atomic game to establish the existence of a solution to the extended HJB system.

Following the characterization and the existence of the optimal contract, we demonstrate the applicability of our framework by explicitly solving the optimal contract for two cases formulated under well-established economic theories. In the first case, the principal has *quasi-hyperbolic* discounting. We adopt the quasi-hyperbolic discount function of Harris and Laibson (2012) with the same interpretation that the "hyperbolic shock" (β in β-δ preferences of discrete time) to the principals' discounting arrives at an exponential rate (as opposed to arriving in the next period in the traditional β-δ case). In such a setting using our methodology we solve the optimal contract in closed form which enables us to explore the implications of the quasi-hyperbolic discounting function on dynamic moral hazard.

We find that under quasi-hyperbolic discounting the optimal effort path is not necessarily monotonic, in contrast to the standard case. The contract takes shape according to two factors; how likely the principal perceives the shock to her discounting and, how large the drop after the shock. We demonstrate that there is a "deadline effect" whereby the likelihood of shock happening before the end of the contracting horizon (the deadline) plays a critical role in shaping the optimal contract.

Our second example explores the case in which the principal has *anticipatory utility*. We adopt the model from Loewenstein (1987) and assume that at each date the principal receives utility from the anticipation of future utility in addition to her immediate consumption. Such utility function implies a "double counting" of future returns and yields a compounded discount factor. Again, using our methodology we solve the contract in closed form which enables us to explore the implications of anticipatory utility on dynamic moral hazard.

We find that under anticipatory utility the optimal effort path is again potentially non-monotonic. However, unlike quasi-hyperbolic discounting, the driving force behind the non-monotonicity under anticipatory utility is the contracting horizon: if the horizon is too short there is not much to anticipate, but once the contracting horizon becomes long enough, anticipatory feelings can overwhelm immediate profits, causing an incentive to back load effort.

The remainder of the paper is structured as follows: after a brief literature review, we establish the general framework in Section 2. Section 3 first characterizes the agent's problem and the relevant game between the agent's selves, then proceeds to characterize the game between the principal's selves and finally delivers the optimal contract. We then solve the two explicit examples (quasi-hyperbolic and anticipatory utility) in Section 4. Section 5 concludes the paper. All proofs are relegated to the appendix.

1.1 Literature Review

The idea that preferences may be dynamically inconsistent as the result of non-exponential discounting stems from the seminal works of Strotz (1955) and Pollak (1968). Since then, many studies have explored dynamically inconsistent preferences in various settings: for example, consumption-saving problems (Laibson (1997), Krusell and Smith (2003), Amador et al. (2006), Bond and Sigurdsson (2017), Cao and Werning (2018)); investment and asset allocation (Caplin and Leahy (2001), Grenadier and Wang (2007), Basak and Chabakauri (2010), Bernheim et al. (2015), Brunnermeier et al. (2016)); monetary policy (Kydland and Prescott (1977)); fiscal policy (Halac and Yared (2014)); capital taxation (Fischer (1980)); criminology (Freeman (1999)); procrastination (O'Donoghue and Rabin (1999)), public finance (Bisin et al. (2015), Harstad (2016)), etc. These studies are typically conducted in discrete time, and are limited to a single party being dynamically inconsistent regardless of the size of the economy.

In this paper, we analyze a continuous-time contracting problem in which both parties can be dynamically inconsistent. Our work relates to studies on contracting and mechanism design with limited rationality: for example Gottlieb (2008), which studies the optimal design of non-exclusive contracts and identifies different implications of immediate-costs goods and immediate-rewards goods for dynamically inconsistent consumers; and Gottlieb and Zhang (2018), which studies repeated contracting between a risk neutral firm and dynamically inconsistent consumers, and finds that allowing consumers to terminate agreements at will may improve welfare if agents are sufficiently dynamically inconsistent. These studies focus on adverse selection problems (reporting) while we focus on moral hazard (hidden effort). Moreover, DellaVigna and Malmendier (2004) analyze the optimal two-part tariff of a firm facing a consumer with known present-biased preferences who makes decisions such as saving or exercising. Heidhues and Kőszegi (2010) and Galperti (2015) explore similar settings; the former show that simple restrictions on the contract form can drastically increase welfare while the latter focuses on the optimal provision of commitment devices. Eliaz and Spiegler (2006) study a setting in which agents differ only in their degree of "sophistication" (their ability to forecast the change in their future tastes.) and characterize the menu of contracts which the principal offers in order to screen the agent's sophistication. See Kőszegi (2014) and Grubb (2015) for a survey of contract theory with behavioral preferences.

Our study differs in several critical dimensions: instead of exploring the implications of time-inconsistency for a certain type of contract or mechanism, we cast our model in an optimal contracting framework, following the now celebrated literature of continuous-time dynamic principal-agent models with transitory private effort.[5] Furthermore, most of the aforementioned studies focus on normative analyses: they either search for ways to implement the first-best result or compare the welfare gains and losses of different mechanisms. In contrast, we provide a positive analysis by demonstrating the effect of non-exponential discounting and different contracting horizons on contract outcomes, and linking them to various observed patterns of contracts in practice.

Our paper also advances the methodology for tackling time-inconsistency problem. Strotz (1955) and Pollak (1968) first suggest the game-theoretic approach that can be applied to solving the optimal dynamic planning problem under time-inconsistency. Kocherlakota (1996) provides a refinement of sub-game perfect equilibrium with dynamically inconsistent preferences. Chade et al. (2008) adopt a Abreu et al. (1990) type recursive characterization of the equilibrium but allow quasi-hyperbolic discounting only. Obara and Park (2017) define the notion of strongly symmetric sub-game perfect equilibrium for individuals with general discounting in repeated games. They explore games with perfect monitoring and their general form of discounting, includes future bias, present bias, and quasi-hyperbolic discounting. Alternatively, Caplin and Leahy (2006), develop a non-strategic, recursive method. The approach is limited to single agent, repeated actions with a specific type of

[5]Such as DeMarzo and Sannikov (2006), Biais et al. (2007), Sannikov (2008), He (2011), etc. These, of course, are all models with exponential discounting and infinite contracting horizon.

objective functions.[6]

Recently, there has been an emergence of mostly mathematical literature of dynamically inconsistent stochastic control, such as Yong (2012), Wang and Wu (2015), Lindensjö (2016), Björk et al. (2017), and Wei et al. (2017), which establish the extended HJB system analogues to us. However, a major caveat of this strand of literature is the lack of the existence to the solution of the extended HJB system, which is a non-trivial problem (See, e.g. Peleg and Yaari (1973) or Bernheim and Ray (1986) for examples in which a solution may not exist.). While each study conjectures various solutions for specific, simplified problems, the existence of solution for the general case remains an elusive, open question.[7] Our novel approach to the existence problem comes from bridging the dynamically inconsistent control problems to studies of non-atomic games such as Schmeidler (1973), Mas-Colell (1984), Balder (1991), Balder (2002), and Khan and Sun (2002). We provide a general theorem describing the scope to which the extended HJB system can be readily applied, thus paving the way for potential future research. Admittedly our model is set in a more restrictive environment compared to the aforementioned mathematical studies, but the restrictions are minimal and are mostly innocuous in economics settings.

2 General Framework

In this section we present the general framework introducing generic, non-exponential discounting functions into an otherwise standard dynamic principal-agent model. We lay out the necessary foundations and describe the dynamic inconsistency problem before providing characterizing the optimal contract in the section.

[6] Instead of maximizing the expected present value of period payoff, they assume entire distribution of future states and choices affect the agent's payoffs

[7] Wei et al. (2017) provides a well-posedness result for uncontrolled diffusion processes. Björk et al. (2017) offers a verification of the extended HJB system, while Lindensjö (2016) proves some regularity properties of the system in Björk et al. (2017), but does not address the existence question either. In particular, the extended HJB system is a characterization of the optimal contract as a system of non-local PDEs and our results prove the existence of a solution to the system of PDEs. Analytical exercises on computations still retains the same challenges as any non local PDE.

2.1 Basic Environment

Time is continuous. A principal (she) contracts with an agent (he) over a fixed-time horizon $T < \infty$. The principal is risk-neutral, with unlimited liability, and has an outside option $\underline{V} \geq 0$. The agent is (weakly) risk-averse and has an outside option $\underline{u}$. The monetary cash flow is given by

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

where $\hat{a}_t$ is the agent's private action (effort), and Z_t is a standard Brownian motion.

Between time t and t' the *principal's* discount rate is given by a bi-function $R(t, t')$. Such specification allows for discounting to vary with both the current date t and an arbitrary future date t', not just the time difference. In particular, at time t, the principal uses the mapping defined by $R^t : [0, T] \to [0, 1]$ as her discount function, where $R^t(s) \doteq R(t, s)$ for all s. One can interpret $R^t(s)$ as that there are infinitely many discounting functions, one corresponding to each point t in time denoted by $R^t(\cdot)$. We impose the following assumptions on $R^t(\cdot)$:

Assumption 1 *For all $\forall t \geq 0$, the principal's discount function $R^t(\cdot)$ satisfies:*

1. $R^t(s) = 1$ *for all* $s \leq t$ *and* $\lim_{s\to\infty} R^t(s) = 0$.

2. $R^t(s) > 0$ *for* s.

3. $\int_t^\infty R^t(s)ds < +\infty$ *a.s.*

4. $R^t(\cdot)$ *is uniformly Lipschitz continuous.*

The first part of the assumption states that starting from any period any payoff in the current period is not discounted, while any payoff in the infinitely far future has a present value of 0. The second part of the assumption states that any return in a finite future has some positive value, albeit potentially very small. The third part ensures that the discounted value of a bounded stream of consumption along any path remains finite. The final part is a standard technical assumption.

Similarly, between time t and t' the *agent's* discount rate is given by a bi-function $r(t, t')$. In particular, at time t, the agent uses the mapping defined by $r^t : [0, T] \to [0, 1]$ where $r^t(s) \doteq r(t, s)$ for all s. We require the following assumptions on $r^t(\cdot)$:

Assumption 2 *For all $\forall t \geq 0$, the agent's discount function $r^t(\cdot)$ satisfies:*

1. $r^t(s) = 1$ *for all* $s \leq t$ *and* $\lim_{s\to\infty} r^t(s) = 0$.
2. $r^t(s) > 0$ *for* s.
3. $\int_t^\infty r^t(s)ds < +\infty$ *a.s.*
4. $r^t(\cdot)$ *is uniformly Lipschitz continuous.*
5. $\frac{R^t(s)}{r^t(s)}$ *is increasing in* s.

Agent's discounting is analogous to the principal's discounting introduced in assumption 1 with one addition: the principal is weakly more patient then the agent. The comparison of patience in the case of discounting bi-functions was first introduced in Quah and Strulovici (2013) from which we borrow directly. This is a fairly standard assumption that prevents the principal from increasing promises until infinity.

Remark 1 *It is important to underline the difference between the bi-function approach we adopt here and a discounting function that relies only on "time difference". Exponential discounting is a particular time difference discount function. In particular, the familiar form $e^{-\gamma(s-t)}$ for some $\gamma > 0$ does not depend on s or t but only on $s - t$. The bi-functions we explore here can readily capture discount functions that only depend on time difference by the simple transformations $r^t(s) = r(s - t)$ and $R^t(s) = R(s - t)$. In such a case we will use the term discount function instead of discount bi-function.*

We assume the following forms of utility functions and risk preferences:

Assumption 3 *The agent has a weakly risk averse utility function that is continuous in both arguments and twice differentiable for evaluating instantaneous consumption and action denoted by $u(c_t, a_t)$, we assume that u is convex and decreasing in a, concave and increasing in c. $u(\cdot, \underline{a})$ is invertible with a continuous inverse function.*

The first requirement on the utility function is fairly standard: effort a is costly with increasing marginal cost and consumption is valuable. The rest of the requirements are largely technical to facilitate the proof of the general existence theorem (Theorem 1), and are sufficient but certainly not necessary conditions. In more specific cases the utility function of the agent can be generalized further. For the principal side, we assume she is risk neutral and maximizes the discounted streams of profits which in this setting equivalent to cash flows net of the consumption process offered to the agent.

Denote the probability space as $(\Omega, \mathcal{F}, \mathcal{P})$ and the associated filtration as $\{\mathcal{F}_t\}_{t\geq 0}$. Contingent on the filtration, a contract specifies a consumption process $\{c_t\}_{t\geq 0}$ with final payment denoted c_T to the agent and a sequence of recommended actions $\{a_t\}_{t\geq 0}$. The set of feasible effort levels comes from a compact set $\mathbb{A} := [\underline{a}, \bar{a}]$. All quantities are assumed to be integrable and progressively measurable under the usual conditions.

Next we define the principal's and the agent's time 0 payoff:

Definition 1 *Principal's time 0 payoff under the contract* $\{c_t, a_t\}_{t\geq 0}$,

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

Agent's time 0 payoff under the contract $\{c_t, a_t\}_{t\geq 0}$,

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

Without exponential discounting, the valuation of future streams of actions and consumption might vary over time. As a result, it is plausible that there are additional gains to be realized by altering the terms (the consumption process and recommended actions) of an existing contract. We call such an alteration a *restructuring* of the contract. Clearly an agent will not be willing to take any restructuring of the contract that delivers lower expected continuation utilities. Therefore, we formally define a feasible restructuring as follows:

Definition 2 *A contract* $\{c'_t, a'_t\}_{t\geq 0}$ *is a restructuring of the contract* $\{c_t, a_t\}_{t\geq 0}$ *if*

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

Contract restructuring is a critical pre-requisite for capturing the principal's dynamic inconsistency in this paper. It translates to partial commitment on the principal's side, as opposed to the standard full commitment assumption which trivializes the search for optimal contracts between sophisticated, dynamically inconsistent parties. There is broad evidence of contract restructuring in practice, where the principal remains *committed* to the promised continuation value.[8] Note that restructuring is different from contract renegotiation. Renegotiation is a more general concept, as it allows the parties to revise past promises as well. In particular, most of the literature on contract renegotiation (with Dewatripont (1988), Aghion et al. (1994) as canonical references) explores cases in which the contract yields suboptimal payoffs to both the agent and principal, and they can renegotiate to re-establish efficiency. Suboptimal payoffs can happen due to reasons such as inefficient punishment for agent's deviations or lack of information. Restructuring, however, is a particular form of renegotiation: it holds the agent's continuation utility constant, but changes how the continuation utility is delivered. In the absence of time-inconsistency, restructuring never happens under a principal-optimal contract.

Remark 2 *Restructuring would be equivalent to renegotiation proofness with full bargaining power on the principal side if the utility frontier is decreasing. See Fudenberg et al. (1990) for details.*

3 Equilibrium Contract and Intrapersonal Games

We now proceed to solve the optimal contract. We adopt a game-theoretic approach following Björk et al. (2017) by constructing a dynamic game between the principal's and the agent's future selves, and look for Markov perfect equilibria. This approach is used in other dynamically inconsistent problems, notably Strotz (1955), Ekeland and Pirvu (2008), Ekeland and Lazrak (2010), Yong (2012), Lindensjö (2016), Wei et al. (2017). Importantly, given the agent and the principal are sophisticated, any choice the principal or the agent makes has to also factor in what the parties might do in the future under different preferences.

[8] See, for example, Mian and Santos (2018) for stylized, empirical observations on the restructuring of bank loans, and Xu (2018) for similar observations on corporate bonds.

As a first step we define the agent's t *self continuation utility*, which is the present value of his future utility under incentive compatibility.

Definition 3 *The agent's t self continuation utility at time t under an incentive compatible contract is:*

$$\tilde{W}(t,t) = \mathrm{E}_t\left[\int_t^T r^t(s)u(c_s,a_s)ds + r^t(T)u(c_T,\underline{a})\middle|\,\mathcal{F}_t\right],$$

where $\mathcal{F}_t$ is the natural filtration generated by Z_t. A contract is incentive compatible if at every instant t agent's effort choice maximizes $W(t,t)$ by choosing $\{\hat{a}_t\}_{t\geq 0} = \{a_t\}_{t\geq 0}$.

Similarly we define the principal's t self *continuation utility*, which is the present value of her future utility.

Definition 4 *The principal's t self continuation utility at time t under an incentive compatible contract is:*

$$f(t,t) = \mathrm{E}_t\left[\int_t^T R^t(s)\left(a_s - c_s\right)ds - R^t(T)c_T\middle|\,\mathcal{F}_t\right],$$

where $\mathcal{F}_t$ is the natural filtration generated by Z_t.

We can now construct the intra-personal dynamic game and its equilibrium, the basic idea of which can be summarized as follows.[9]

- The principal and the agent are playing a non-cooperative game with their future selves. For each point in time t, treat the principal and the agent as different players, whom we refer to as "Agent t" and "Principal t".

- For each fixed t, Agent t and Principal t can only influence the process M_t exactly at time t. Principal t chooses suggested action $\hat{a}_t$ and the compensation c_t and agent t chooses the action a_t.

[9]Full details of the game and the equilibrium strategies are mathematically intricate and therefore left in the Appendix.

- At each point in time, each player t's best response (in a Markov Perfect Equilibrium sense) to his/her different selves forms an intra-personal equilibrium, taking the other parties' intra-personal equilibrium as given. We call each intra-personal equilibrium as an equilibrium control for the agent/principal respectively.[10]

3.1 The Agent's Problem

We now solve the agent's problem. In general, the agent's consumption at any instance could depend on the entire path of outputs. We follow Williams (2015) and introduce a change of measure to resolve the issue of history dependence. We first fix an equilibrium contract process for the principal and denote it by $\hat{c}_t$. For any $\hat{a} \in [\underline{a}, \bar{a}]$, define the family of $\mathcal{F}_t$ predictable processes:

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

with $\Gamma_0(\hat{a}) = 1$. Suppressing $\hat{a}$, we introduce the following change of measure:

$$\mathrm{E}^{\mathbb{P}^{\hat{a}}}\left[\int_t^T r(t,s)u(a_s,\hat{c}_s)ds + R(t,T)u(\underline{a},\hat{c}_T,)\right] = \mathrm{E}^{\mathbb{P}^0}\left[\int_t^T \Gamma_s r(t,T)u(a_s,\hat{c}_s)ds + \Gamma_T R(t,T)u(\underline{a},\hat{c}_T)\right].$$

The optimal control problem of the dynamically inconsistent agent under the measure $\mathbb{P}^0$ is given by:

$$J^A(t,\Gamma) = \sup_{a_s,\, s\in[t,T]} \mathrm{E}^{\mathbb{P}^0}\left[\int_t^T \Gamma_s r(t,s)u(a_s,\hat{c}_s)ds + \Gamma_T r(t,T)u(\underline{a},\hat{c}_T)\right], \quad (2)$$

$$d\Gamma_s = \Gamma_s\left(\frac{a_s}{\sigma}\right)dZ_s^0. \quad (3)$$

The solution to this agent's problem resembles a textbook Pontryagin approach (as in Wang and Yong (2019) and Hamaguchi (2019)). The notable difference compared to the classical problem is that our problem implies a flow of adjoint processes, which determines the

[10] In discrete time, it is possible to show that every subgame perfect Nash equilibrium must be Markovian. In continuous time though, this definition is not sufficient, given Player t can only choose the exactly at time t, which only influences the control on a time set of Lebesgue measure zero. Put differently, the control chosen by an individual player will have no effect on the dynamics of the process. We therefore adopt a modified definition of the equilibrium concept following Björk et al. (2017). See the Appendix for details.

evolution of the continuation utility of each agent $t \in [0, T]$.

The next lemma characterizes the continuation utilities and the incentive compatibility condition:

Lemma 1 *Under assumptions 1, 2, 3, given any contract $\hat{a}, \hat{c}$ and any sequence of the agents choices, there exists a flow of processes $\tilde{\psi}(t,s)$ and an equilibrium value process W such that for each self of the agent we have:*

$$\tilde{W}(t,k) = r(t,T)u(0,c_T) - \int_k^T r(t,s)u(a_s,\hat{c}_s)ds + \int_k^T \tilde{\psi}(t,s)dZ_s^{\hat{a}} \tag{4}$$

The equilibrium value process satisfies:

$$W(t) = r(t,T)u(0,c_T) - \int_t^T r(t,s)u(a_s,\hat{c}_s)ds + \int_t^T \tilde{\psi}(t,s)dZ_s^{\hat{a}} \tag{5}$$

The contract is incentive compatible if and only if

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

Under an incentive compatible contract, the agent's effort choice becomes a Markov Perfect equilibrium in which the strategies are Markovian with respect to Γ and t. In contrast to Sannikov (2008), we have a flow of continuation utilities, one for each self of the agent. The process $\tilde{W}(t,k)$ pins down self ts continuation from period k onwards, whereas the process $W(t)$ pin downs the equilibrium value.

For each self t, $\tilde{W}(t,k)$ is a process. However, as the selves evolves over time, $\tilde{W}(t,k)$ becomes a *flow* of processes. Consequently, instead of a backward stochastic differential equation that would arise in a dynamically consistent environment (as in Cvitanic and Zhang (2012)), the representation in equation 4 is a backward stochastic Volterra integral equation. At each time of decision making, the t self is the relevant self and the process $\tilde{W}(t,t)$ is the relevant continuation utility. $\tilde{W}(t,t)$ takes the self t as a parameter and is a backward stochastic differential equation much like the dynamically consistent case (e.g Sannikov (2008)). This leads to an identical incentive constraint as in Sannikov (2008), which provide the agent with local incentives via $\psi_t = \tilde{\psi}(t,t)$. Nevertheless, the derivation of such close resemblance is a non-trivial task, which we lay out in detail in the Appendix.

Setting the agent's self and time the same (i.e. $k = t$), as shown in Yan and Yong (2019), leads to the integral valued process $W(t)$. The equality of these two processes were first shown in Yong (2008). Notice that unlike the t self valuation, $W(t)$ captures not only the time change but also the change in the agent's preferences. Therefore, $W(t)$ is the relevant variable for the principal when designing the optimal contract. In the appendix we show that the integral valued process $W(t)$ also has a forward representation.

3.2 The Principal's Problem and the Optimal Contract

The principal's problem is significantly more involved. Before we delve into the full characterization let us make some observations about this problem.

First, at the end of the contract the principal can always retire the agent with a lump sum payment to deliver a promise W_T. Since the principal's valuation of payments vary over time we denote the principal's *retirement benefit* $B(t, W_T)$ evaluated at time t, as

$$B(t, W_T) = -R^t(T)c_T \tag{6}$$

where c_T is chosen such that $u^{-1}(c_T, \underline{a}) = W_T$.

Next, observe that the (dynamically inconsistent) principal's optimal control problem is:

$$J^P(t, W) = \sup_{(a_s, c_s, \psi_s)\ s\in[t,T]} \mathrm{E}_{t,W} \left[\int_t^T R^t(s)\,(a_s - c_s)\,ds + B(t, W_T) \right]. \tag{7}$$

Thus we define the optimal contract as:

Definition 5 *A contract is optimal if it maximizes $J^P(t, W)$ for all t, W over the set of contracts that are incentive compatible.*

Finally, we introduce the following notation of a controlled infinitesimal generator:[11]

[11] This notation comes from Björk et al. (2017). The only difference compared to the standard notation in stochastic calculus is that it acts on the time variable as well as on the space variable. (i.e. it includes the term $\partial/\partial t$)

Definition 6 *Suppose* $x = [x_1 \ x_2 \ ... \ x_n]$ *is a* n*-dimensional vector of state variables and* $Z_t^j, \ j = 1, 2, ..., n$ *are independent Brownian motions such that*

$$dx_j = \mu_j(t, x)dt + \sigma_j(t, x)dZ_t^j. \tag{8}$$

For any process $G(t, x)$ *and control* h *the operator* $\mathcal{A}^h$ *is defined by*

$$\mathcal{A}^h G = \frac{\partial G}{\partial t} + \sum_{j=1}^{n} \mu_j^h(t, x) \frac{\partial G}{\partial x_j} + \frac{1}{2} \sum_{i,j}^{n} \sigma_j^h(t, x) \sigma_i^h(t, x) \frac{\partial^2 G}{\partial x_j x_i}. \tag{9}$$

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

Theorem 1 *Under assumptions 1, 2, and 3, there exists an optimal contract. The principal's equilibrium value function under this contract* $V(t, W)$ *with optimal controls* $h_t = (a^*, c^*, \psi)$*, is given by:* [12]

$$V(t, W) = \mathrm{E}_{t,W}^h \left[\int_t^T R^t(k) \left(a_k^* - c_k^* \right) dk + B(t, W_T) \right]. \tag{10}$$

The equilibrium value function satisfies the following extended HJB system:

$$\sup_h \{ \mathcal{A}^h V(t, W) + (a_t - c_t) - \mathcal{A}^h f(t, W, t) + \mathcal{A}^h f^t(t, W) \} = 0. \tag{11}$$

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

$$\mathcal{A}^{h^*} f^s(t, W) + R^s(t) \left(a_t^* - c_t^* \right) = 0, \tag{12}$$

where for each fixed s *the function* $f^s(t, W) = f(t, W, s)$ *equation (12) is a Kolmogorov backward equation with the probabilistic interpretation:*

$$f^s(t, W) = \mathrm{E}_{t,W}^{h^*} \left[\int_t^T R^s(k) \left(a_k^* - c_k^* \right) dk + B(t, W_T) \right] \tag{13}$$

[12]From here on we will suppress the h in V whenever it is clear.

subject to (4), the IC condition (IC) and boundary conditions:

$$V(T, W) = B(T, W_T) \text{ for all } W \tag{14}$$

where $B(t, W)$ is defined by (6).[13]

Theorem 1 highlights two important differences compared to the standard optimal contract under exponential discounting which, under our notation, would be captured by a single Hamilton-Jacobi-Bellman (HJB) equation involving just $\mathcal{A}^h V$ and $(a_t - c_t)$.[14]. The first difference is that there is now a system of equations, instead of one single HJB equation. The additional equation (12) is a Kolmogorov Backward equation defined using the infinitesimal generator, and captures the "backward induction" intuition that the principal must follow when playing a game against her future selves.[15]

Second, $f(t, W, t)$ and $f^t(t, W)$ are the "extra" terms in the HJB equation.[16] In exponential discounting benchmark models, the principal's immediate utility $((a_t - c_t))$ and her value function $(V(t, W))$ are sufficient to capture the immediate and future incentives for any action of the agent, and the suprema determines the optimal actions. However, in our model due to dynamic inconsistency we not only need additional terms $(f(t, W, t)$ and $f^t(t, W))$ but also need to identify these terms through another equation (12). Under dynamic consistency, $\mathcal{A}^h f(t, W, t)$ and $\mathcal{A}^h f^t(t, W)$ always cancel out exactly, leaving only the first two terms in (11) in the standard HJB equation.

To understand those extra terms and how they are determined, first observe that the equilibrium value function $V(t, W)$ yields the value of each time t self for each potential state W and relies on what each self will do at all $t' > t$. This dependence on "far" terms introduces a non-local PDE since different selves discount future differently, instead of the usual local PDE from an HJB equation. To pin down this PDE, another non-local function which considers the valuations of different selves is necessary. The term $f(t, W, s)$ is indeed

[13]Note that operator $\mathcal{A}$ only acts on the objects inside the parenthesis. Moreover, on the equilibrium by construction we have $V(t, W) = f(t, W, t)$.

[14]The HJB equation is commonly written as (suppressing the notation) $rV = \sup\{a_t - c_t + \mu_W V_W + \frac{1}{2}\sigma_W^2 V_{WW}\}$. Under our definition of the infinitesimal generator $\mathcal{A}$, it simply equals $\sup \mathcal{A}^h V - (a_t - c_t) = 0$.

[15]Iijima and Kasahara (2016) shows another application of backward stochastic differential equations in dynamic games.

[16]Note that $f(t, W, t)$ is $f(t, W, s)$ and $f^t(t, W)$ is $f^s(t, W)$ both evaluated at $s = t$.

exactly what regards time s-selves as a variable in addition to the usual current time t and state W because intuitively, a change in period t action has an equilibrium effect on *all* selves s. Notice that variations in s also leads to variations in the discount functions as $R(s, \cdot)$ varies in s. $f(t, W, t)$ represents the valuation of the principal where t is the *current* time and t represents a time-t self but a change in t changes the self as well. However, at any particular time s, instead of the equilibrium value a fixed s-self makes a decision with the goal to maximize her own self's payoff, taking herself as given, not a variable. The term $f^s(t, W)$ therefore captures the principal's valuation of future equilibrium actions by her time-s self in state W at time t. $f^t(t, W)$ is the self that is actually doing the maximization and shows up next to $f(t, W, t)$ and $V(t, W)$. It is another non-local PDE but relies only on a single discount function $R^t(\cdot)$. In equilibrium between the different selves a change in action at time t and state W should be perfectly foreseen by the previous time s selves, leading to the backward part (equation 12) behaving like a martingale and hence being equal to 0.

In general, the system in Theorem 1 is very difficult to analyze due to the interaction of non-local PDEs. Indeed, Wei et al. (2017) note the appearance of the backward term poses an essential difficulty. They remark that "at the moment we are not able to overcome the difficulty" and pose it as an open problem. Similarly Björk et al. (2017), whose HJB system is very similar to ours, note that "the task of proving existence and/or uniqueness of solutions to the extended HJB systems seems (at least to us) to be technically extremely difficult". The challenges mainly come in two ways: first, the backward system is actually accompanied by a forward system arising from the state variables, turning it into a forward-backward stochastic differential equation which is known to be very difficult to tackle (Cvitanic and Zhang (2012)). Second, without the solution to the backward system, the value function, which comes from its own non-local PDE, cannot be solved.

We manage to establish the existence of a solution to the backward system in Theorem 1 by taking a unique approach: first, we only establish the well-posedness of the backward system under our smoothness assumptions. Then we turn to the value equation, where the solution from the backward system comes in implicitly. We define the game between the different selves as a game of incomplete information between non-atomic players where the utility function of each of these players incorporates these solutions from each of the backward systems. Of course the backward system without an explicit solution only gives a PDE characterization of the utilities of these players. However, using that PDE analogous to the approach used by Mas-Colell (1984) we look for a distribution over utility functions, strategies and information as an equilibria. Once the contracting problem is translated into this non-atomic game of incomplete information, using existing results from game theory (in particular Balder (1991)), we show that an equilibrium exists, which implies there exists a solution to extended HJB system.

Due to the level of technicality involved, we present the detailed, step-by-step proof of Theorem 1 including the novel existence argument in the Appendix. 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. In Section 4 we demonstrate two special cases in which Theorem 1 admits closed-form solutions and simple comparative statics.

Remark 3 *It is known (Björk et al. (2017)) that if a dynamically-inconsistent optimal control problem has an equilibrium solution, there is a dynamically consistent problem that leads to the same optimal control but with a different instantaneous utility function. Indeed, this would correspond to the case in which one defines a sophisticated decision maker's immediate utility function that directly incorporates the actions of the other selves (as opposed to just* $a_t - c_t$ *or* $u(a_t, c_t)$*). Such an approach is useful for an axiomatic characterization but not very applicable for solving the optimal contracting problem. We explore the connection between this and our approaches in the online appendix.*

4 Applications

In this section we provide two explicitly solved examples of the optimal contract under simplified conditions. Given that the main innovations from introducing dynamic inconsistency come from the principal's problem, we simplify the agent's problem assuming he has CARA utility with time-consistent exponential discounting. We then analyze two kinds of behaviors that lead to non-exponential discounting for the principal: quasi-hyperbolic discounting, following Harris and Laibson (2012); and anticipatory utility following Loewenstein (1987).

In both cases we obtain closed-form solutions of the optimal contracts and offer comparative statics that shed light on the unique implications of dynamically-inconsistent preferences.

Specially, we maintain throughout this section the following assumption that strengthens and replaces assumptions 2 and 3:

Assumption 4 *The agent has the following discount function and CARA utility.*

$$r^t(s) = e^{-\gamma(s-t)} \text{ for all } t$$

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

He also has access to a private savings account, the balance of which grows at rate γ*.*

CARA utility plus private saving is a commonly used technique in the contracting literature to simplify the agent's problem.[17] In our model, they lead to the following results:

Lemma 2 *Under assumption 4 the agent's continuation utility satisfies*

$$dW_t = \gamma(W_t - u(c_t, a_t))dt + \psi_t(dM_t - a_t dt) \ , \tag{15}$$

$$u(c_t, a_t) = \gamma W_t \ . \tag{16}$$

The agent's incentive compatibility condition becomes

$$\psi_t = a_t. \tag{17}$$

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

$$\mathbb{E}\left[\ln(-W_t)\right] = \ln(-W_0) + \frac{1}{2}\int_0^t \eta^2\gamma^2\sigma^2\psi_s^2 ds. \tag{18}$$

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

The proof of lemma 2 is analogous to the proof of He (2011) and hence omitted.

[17] See, e.g. He (2011), Williams (2015), Marinovic and Varas (2018), Gryglewicz and Hartman-Glaser (2019), etc.

4.1 Quasi-Hyperbolic Discounting

Our first example explores a case in which the principal has quasi-hyperbolic discounting. Such discounting has been explored in many examples including but not limited to O'Donoghue and Rabin (1999), Thaler and Benartzi (2004), Harris and Laibson (2012), Jackson and Yariv (2014), Jackson and Yariv (2015), Bisin et al. (2015), etc. Here we follow the form of quasi-hyperbolic discounting introduced in Harris and Laibson (2012) where the quasi-hyperbolic discount function is a convex combination of a short-term discount-function and a long-term discount factor.[18]

Formally, we assume the following:

Assumption 5 *The principal has the following discount function:*

$$R^t(s) = (1-\beta)e^{-(\rho+\lambda)(s-t)} + \beta e^{-\rho(s-t)} \tag{19}$$

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

The representation above is the deterministic characterization of a principal who values "near present" returns with a higher discount factor (discounted by $e^{-\rho(s-t)}$), and "far future" returns with a lower 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 discount factor in the far future. The switch between the "near present" and the "far future" happens stochastically with arrival intensity λ, and the overall discount function $R^t(s)$ incorporates this expected drop. Notice that for the principal the discounting between period t and $s > t$ only relies on $s - t$ but not t or s individually. Thus with a slight abuse of notation, we will denote $R^t(s)$ from equation (19) as $R(s-t)$.

With the simplified agent's continuation utility (from Lemma 2) and the specific discount function for the principal, we now apply Theorem 1 to calculate the optimal contract. The solution turns out to take a closed-form:

Proposition 1 *Under assumptions 4, 5 in the optimal contract a_t and c_t satisfy the follow-*

[18] Pan et al. (2015) also offers a continuous time version of quasi-hyperbolic discounting.

ing:

$$\begin{aligned} a_t = & \left[1 + \tfrac{\eta\gamma^2\sigma^2}{\rho(\lambda+\rho)}\left(-\beta\lambda(1-\beta)\left(e^{-(\lambda+\rho)(T-t)} - e^{-\rho(T-t)}\right) + (1-\eta+\rho\eta+\lambda\eta)\rho + (1-\eta)\beta\lambda\right)\right]^{-1} \\ c_t = & \tfrac{1}{2}a_t^2 - \tfrac{\ln(\gamma\eta)}{\eta} - \tfrac{1}{\eta}\ln(-W_t) \end{aligned}$$

We are now able to analytically explore the implications of quasi-hyperbolic discounting on the optimal contract in detail. We summarize our key findings as follows:

Proposition 2 *The optimal contract derived in Proposition 1 has the following properties:*

1. *If $\beta = 1$ ($\beta = 0$) optimal contract converges to dynamically-consistent principal with discount rate ρ ($\rho + \lambda$) and optimal effort a_t is (weakly) monotonic in time t.*
2. *As $T \to \infty$ optimal contract converges to an optimal contract with time consistent-principal with discount rate $\frac{\rho(\lambda+\rho)}{\rho+\beta\lambda}$.*
3. *For every $\rho \in (0,1)$, if λ is high enough a_t becomes non-monotone function over time.*

The first property is straightforward: if $\beta = 1$ ($\beta = 0$) the principal has the same discount rate for "near present" returns and for "far future" returns, which is equivalent to having exponential discounting. Unsurprisingly, the optimal solution is analogous to the celebrated solution of Holmstrom and Milgrom (1987) in which the path of the optimal effort is monotonic, regardless of the time horizon.[19]

The second property demonstrates that the effect of quasi-hyperbolic exponential discounting becomes indistinguishable if the time horizon becomes arbitrarily large. Björk et al. (2017) shows that for any dynamically inconsistent optimal control problem that has a solution, there is a time-consistent problem that has the same solution with a *different instantaneous utility function*. In comparison, we show that if the contracting horizon is infinitely long, the solution to the dynamically inconsistent problem agrees with that of a time-consistent problem with *exactly the same instantaneous utility function* but a *different exponential discount factor*, which we pin down explicitly. Intuitively, because the principal is sophisticated, as the horizon increases the problems that each of her future selves

[19] Whether effort is monotonically increasing, decreasing, or constant over time is the result of model-specific boundary conditions. See He (2011), Marinovic and Varas (2018), He et al. (2017) for some recent examples.

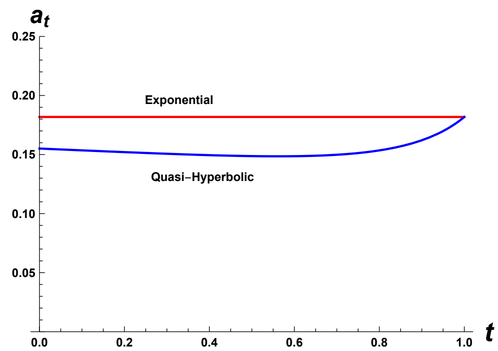
face become similar. When the horizon is infinite, the principal's problem becomes stationary. Under quasi-hyperbolic discounting, stationarity is sufficient to ensure that the solution agrees with that of a principal with the exactly the same instantaneous utility function.

The third property highlights the possibility of non-monotonic actions under finite horizon. From the previous two properties we know that effort is (weakly) monotonic in time without quasi-hyperbolic discounting or without finite horizon. Under quasi-hyperbolic discounting and 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 actions. Suppose that the former prefers high effort and the later prefers low effort. Because the switch between the two types of principals happens 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 effort 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 becomes smaller over time, as the contract is "running out of time." Consequently, the principal acts more likely her "near future" patient self and reverts the recommended effort until it converges to the time-consistent benchmark level at the deadline, causing a U-shaped path. Similarly effort could first increase then decrease towards the end, causing a hump-shaped path. We refer to this reverting of actions 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 so remote that the "deadline effect" does not kick in and the path of optimal actions converges to the (weakly) monotonic case.

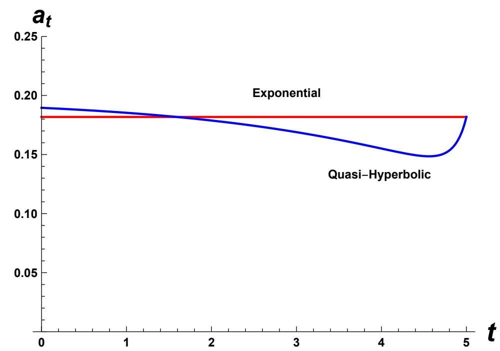
We illustrate the properties above and the comparison between quasi-hyperbolic discounting and the time-consistent benchmark in Figure 1.

4.2 Anticipatory Utility

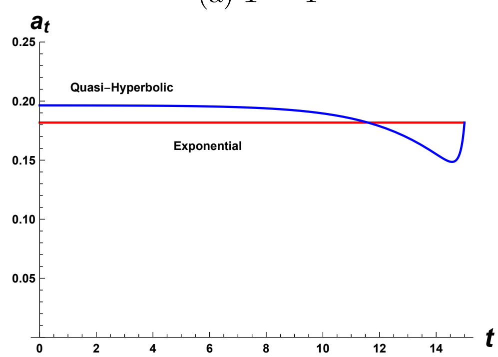
Our second example explores the case in which the principal has an anticipatory utility function. That is, the principal's utility depends on her expectations of the future in addition to immediate utility. Positive utility from future consumption may cause "savoring" or



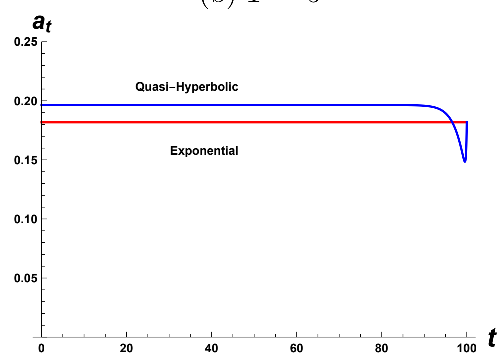
(a) $T = 1$



(b) $T = 5$



(c) $T = 15$



(d) $T = 100$

Figure 1: These plots illustrate the action paths over time recommended by the optimal contract. The blue line indicates the action path under quasi-hyperbolic discounting with parameter values $\eta = 2, \sigma = 1, \beta = 0.5, \lambda = 5, \rho = 0.5, \gamma = 1.5$. The red line indicates the action path under exponential discounting with $\rho = 1$ and the rest of the parameters the same. Each plot corresponds to a different time horizon (different T).

"dread" as introduced by Loewenstein (1987) which effects the intertemporal decisions. As noted in Brunnermeier et al. (2016), such anticipatory emotions are also tied to optimism of the principal. Other related studies of anticipatory utility include but are not limited to Loewenstein (1987), Caplin and Leahy (2001), Loewenstein et al. (2003), Caplin and Leahy (2004), Takeuchi (2011), etc.

We model anticipatory utility following Loewenstein (1987). In addition to utility from profits today, the principal derives $e^{-\zeta t}$ units of utility today from anticipated profits that she expects to receive in future time t. Meanwhile, she discounts utility in the future (whether actual profits or from anticipation of future profits) with a discount rate ρ. Together, they imply the following:

Assumption 6 *The principal values consumption streams by:*

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

where Π_t *is given by*

$$\Pi_t = \mathrm{E}\left[\int_t^T e^{-\zeta(s-t)}(dM_s - c_s)ds|\mathcal{F}_t\right].$$

Here, Π_t captures the utility from anticipated profits, while the remaining terms represent the standard flow utility from actual profits, which equal output net of compensation to the agent, discounted appropriately. Using the Law of Iterated Expectations and changing the order of integration, the principal's valuation of a consumption stream starting from any period t can be written as follows:

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

This means the principal effectively has the following discounting function:

$$R^t(s) = R(s-t) = \frac{e^{-\zeta}(s-t) - e^{-\rho(s-t)}}{\rho - \zeta} + e^{-\rho(s-t)}.$$

The first term captures the discounted anticipation: letting $\zeta \to \infty$ yields the standard discounting as 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.[20]

We again apply Theorem 1 and solve the optimal contract in closed-form:

[20]Although not noted in Loewenstein (1987), an appropriate numerical relationship between ρ and ζ turns out to be economically crucial. Without further restrictions, utility of infinitely far future may still have a positive value today after discounting.

Proposition 3 *Under assumptions 4, and 6, a_t and c_t satisfy the following:*

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

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

Based on these results, the most critical implications of anticipatory utility can be summarized as follows:

Proposition 4 *1. If $\zeta = \infty$ optimal contract converges to dynamically-consistent principal with discount rate ρ.*

2. As $T \to \infty$ optimal contract converges to an optimal contract with time consistent-principal with discount rate $\frac{\rho\zeta}{1+\zeta}$.

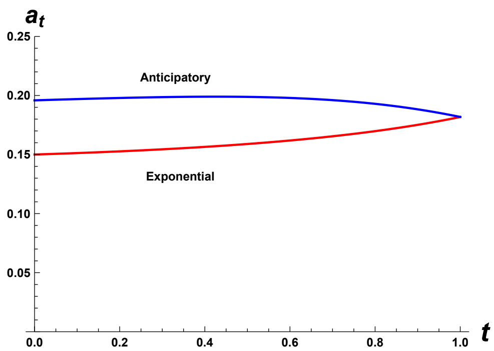
3. For any ρ, ζ there exists a finite T such that a_t is non-monotone.

The first property is straightforward: if $\zeta = \infty$ the principal receives no utility from anticipation. That is, she has exponential discounting. Not surprisingly, the optimal solution is analogous to the celebrated solution of Holmstrom and Milgrom (1987) which induces monotone effort regardless of the time horizon.[21]

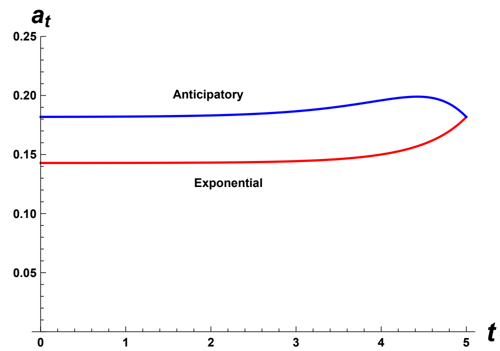
The second property demonstrates that the effect of anticipation disappears if the time horizon becomes arbitrarily large. The intuition is very similar to why the effect of quasi hyperbolic discounting disappears under an infinite time horizon, which we described in Section 4.1. In short, because the principal is sophisticated, as the horizon increases the problem that each of her future selves faces becomes similar as it only relies on the time difference. The principal's problem becomes stationary when the time horizon is infinite. In particular under anticipatory utility, such stationarity is sufficient to ensure that the solution agrees with that of a dynamically consistent principal which has the same instantaneous valuation (albeit with a different discount rate) as the dynamically inconsistent one.[22]

[21] Note that the difference arises due to the boundary conditions. More recently with He (2011), Marinovic and Varas (2018), He et al. (2017) reach similar conclusions.

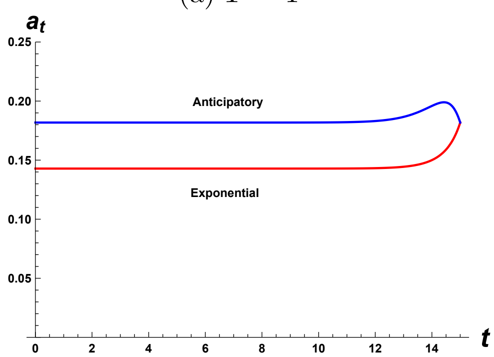
[22] In the appendix we replicate the argument of Björk et al. (2017) which shows that any dynamically inconsistent problem that has a solution is associated with dynamically consistent problem that uses the same controls, but has a different and potentially time dependent instantaneous utility function.



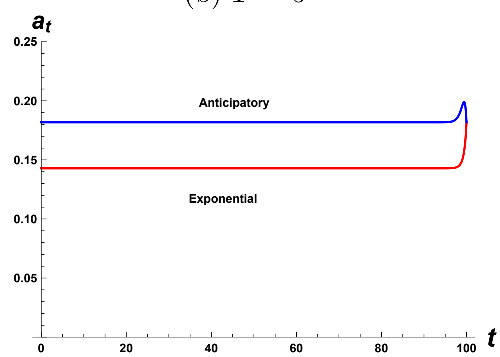
(a) $T = 1$



(b) $T = 5$



(c) $T = 15$



(d) $T = 100$

Figure 2: These plots illustrate the action paths over time recommended by the optimal contract. The red line indicates the action path under anticipatory utility function with parameter values $\eta = 2, \sigma = 1, \rho = 1.5, \zeta = 3, \gamma = 1.5$. Each plot corresponds to a different time horizon (different T).

The third property highlights the possibility of non-monotonic actions under finite horizon. Again, from the previous two properties we know that effort is (weakly) monotonic in time with exponential discounting or without finite horizon. With anticipatory utility and a large enough but finite horizon though, the anticipation of future profits leads the principal to backload effort. However, if the horizon is short enough the anticipation of future decreases and the action path still remains monotone as there is not enough time to backload. The length of the time horizon is critical for the non-monotonic effort path: on the one hand, if the horizon is not long enough, anticipation has no chance to "build up"; on the other hand, if the horizon is infinite then the "build-up" is pushed infinitely far into the future, resulting in a convergence into a monotonic effort path but through a different channel compared to the former case of present-bias utility.

We illustrate the above properties in Figure 2.

5 Conclusion

It is well known in the literature that individuals with dynamically inconsistent preferences make dynamically inconsistent plans, because what appears optimal today may ultimately be suboptimal tomorrow. As a result, is it possible for long-term contracts between parties with such preferences to exist, particularly if the parties are sophisticated in anticipating the changes to their preferences in the future? If so, how do the contracts differ from known time-consistent benchmarks? We answer these questions by establishing a general framework that combines moral hazard with dynamically inconsistent preferences resulting from non-exponential discounting. This contracting problem can be equivalently expressed as a dynamic game played between the agent, the principal and their future selves, although the characterization of its solution remains an open problem in the extant literature. We provide a novel argument for the solution by translating this contracting problem into a non-atomic game with incomplete information and prove the existence of its equilibrium. Given its generality, this approach is potentially applicable in other settings beyond dynamic contracting. We also present two explicitly solved examples capturing two economically important behavioral patterns: present bias and future bias, for which we manage to obtain closed-form solutions. The resulting optimal contracts produce testable implications, such as time-varying, non-monotonic effort and consumption paths.

There are several directions in which our study could be potentially extended. First and foremost, the extended HJB systems and their variations arise in many dynamically inconsistent optimal control problems not necessarily because of non-exponential discounting. We hope that our indirect approach that bridges dynamically inconsistent control to non-atomic games can be further expanded to tackle the solution to the optimal control under those more generic scenarios. Additionally, although the general existence theorem is rather broad, our solved examples highlight the direct applicability of the techniques as well; in particular, it might be interesting to explore dynamically inconsistent contracting problems under other specifications such as ambiguity, habit formation and mean-variance risk preferences. We leave these questions for future research.

A Appendix A

A.1 Equilibrium Control Laws

In this part we define the intra-personal equilibrium generically without referencing the agent or the principal. The definition follows immediately for either contracting party when the relevant process and controls are replaced in definition below (Γ as the relevant process a as control for the agent, and W as the process and (a, c, ψ) as control for the principal). Formally, let $(\Omega, \mathcal{F}, \mathrm{P})$ be a complete filtered probability space on which a standard $d-$dimensional Brownian motion Z is defined and let X be controlled stochastic process defined by a stochastic differential equation:

$$dX_t = \mu(t, X_t, h_t)d_t + \sigma(t, X_t, h_t)dZ_t \quad t \in [0, T]$$

Where μ and σ are suitable deterministic maps,and $h_t \in H \subseteq \mathbb{R}^m$ for all t is an m dimensional control process. Consider a return function $g : [0, T] \times [0, T] \times \mathbb{R} \times H \to \mathbb{R}$ and a final reward function $f : [0, T] \times [0, T] \times \mathbb{R} \to \mathbb{R}$ and an objective defined as:

$$J(t, x, h) = \mathrm{E}\left[\int_t^T g(t, s, x_s, h_s)ds + f(t, T, x_T)\right].$$

Definition 7 *Consider a control process h. For any initial point and state (t, x) and a "small" increment of time Δ, define a "deviation" strategy h_Δ as:*

$$h_\Delta = \begin{cases} \hat{h}(t', x) & \textit{for } t \leq t' < t + \Delta \\ h(t', x) & \textit{for } t + \Delta \leq t' \leq T \end{cases}$$

where $\hat{h} \neq h$ is another control law, and x is the state vector corresponds to the future date t'.[23] *h is called an* ***equilibrium control law*** *if*

$$\liminf_{\Delta \to 0} \frac{J(t, x, h) - J(t, x, h_\Delta)}{\Delta} \geq 0 \tag{20}$$

for all $\hat{h} \neq h$. For an equilibrium control law h, the corresponding value function $J(t, x, h)$, is called as an ***equilibrium value function****.*

A.2 Proof of Lemma 1

Consider an equilibrium contract offered by the principal, the principal at any time observes the entire history of outputs only, and offers a consumption path c_t in return. In general the consumption at any time could depend on the entire history of outputs. We first fix a contract process for the principal and let it be denoted by $\hat{c}_t$. To resolve such a history

[23] In the language of stochastic optimal control h_Δ is the local spike variation of h. Also, this equilibrium approach makes the connection between the discrete time and continuous time clearer. However the full technicality of this connection is beyond the scope of this paper, see Sadzik and Stacchetti (2015) for an excellent treatment of the issue with exponential discounting (time consistent case).

dependence we follow Williams (2015) and introduce a change of measure. Now for any $\hat{a} \in [\underline{a}, \bar{a}]$ define the family of $\mathcal{F}_t$ predictable processes:

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

with $\Gamma_0(\hat{a}) = 1$. Given the linear output, Novikov condition holds and Γ_t is $\mathcal{Q}^0$ martingale. By Girsanov theorem we define the Brownian Motion (Z^a) in the following way:

$$Z_t^a = Z_t^0 - \int_0^t \frac{a_s}{\sigma} ds.$$

Then suppressing $\hat{a}$ we can do the following change of measure as follows:

$$\mathrm{E}^{\hat{a}}\left[\int_t^T r(t,s)u(a_s,\hat{c}_s)ds + R(t,T)u(\underline{a},\hat{c}_T,)\right] = \mathrm{E}^0\left[\int_t^T \Gamma_s r(t,T)u(a_s,\hat{c}_s)ds + \Gamma_T R(t,T)u(\underline{a},\hat{c}_T)\right].$$

Now let us write down the time inconsistent optimal control problem as follows:

$$J(t,\Gamma) = \sup_{a_s,\, s\in[t,T]} \mathrm{E}^0\left[\int_t^T \Gamma_s r(t,s)u(a_s,\hat{c}_s)ds + \Gamma_T r(t,T)u(\underline{a},\hat{c}_T)\right], \tag{21}$$

$$d\Gamma_s = \Gamma_s\left(\frac{a_s}{\sigma}\right) dZ_s^0. \tag{22}$$

Notice that the optimal control problem is still time inconsistent but with the change of measure we have resolved the history dependence. Given the dynamic inconsistency of the problem we will be looking for an Markov perfect (closed loop equilibrium strategy) a^*. Under our regularity conditions (a coming from a compact space, $\mu_\Gamma, \sigma_\Gamma$ being C^2 in a and Lipschitz and) we can apply Yan and Yong (2019) Theorem 1 for the following Pontyagrin type maximum principle for the time inconsistent control problem 21.

Theorem 2 (Yan and Yong 2019) *Suppose Γ^*, a^* is an equilibrium pair for the time inconsistent problem and suppose for any given $t \in [0,T)$ the first order adjoint processes $W(t,\cdot), \tilde{\psi}(t,\cdot)$ and second order adjoint processes $P(t,\cdot), \Lambda(t,\cdot)$ are adapted solutions to the following BSDE's*

$$dW(t,s) = -\left(\frac{a_t^*}{\sigma}\tilde{\psi}(t,s) + r(t,s)u(a_s^*,\hat{c}_s)\right)ds + \tilde{\psi}(t,s)dZ_s^0 \quad s \in [t,T]$$
$$W(t,T) = r(t,T)u(\hat{c}_T,\underline{a})$$
$$dP(t,s) = -\left(\frac{a_t^*}{\sigma^2}\right)^2 P(t,s)ds + 2\left(\frac{a_t^*}{\sigma}\right)\Lambda(t,s)ds + \Lambda(t,s)dZ_s^0 \quad s \in [t,T]$$
$$P(t,T) = 0$$

then almost surely for any control a we have the following global form of Pontyagrin maximum *principle:*

$$0 \leq \langle W(t,t), a - a^* \rangle + \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathrm{E}_t \int_t^{t_\varepsilon} \left[\left\langle \tilde{\psi}(t,s), \Gamma_s^* \left(\frac{a - a^*}{\sigma} \right) \right\rangle ds \right.$$
$$\left. + u(a, \hat{c}_t) - u(a^*, \hat{c}_t) + \left(\Gamma_t^* \frac{a - a_t^*}{\sigma} \right)^2 P(t,t) \right] \tag{23}$$

If agent's had exponential discounting, there will be only one adjoint process. In our setup, we have a flow of adjoint processes. This flow of processes describes how different selves of the agent evaluates the equilibrium contract. Observe that given that a is coming from a convex set, instead of equation 23 following Yong and Zhou (1999) page 120 case 2 (also Peng (1990)) we can use a local version. Thus we can define the generalized Hamiltonian of time t for any time s as:

$$\mathcal{H}_t(s) = -r(t,s)\Gamma_s u(a_s, \hat{c}_s) ds + \tilde{\psi}(t,s)\Gamma_s \left(\frac{a_s}{\sigma} \right) \tag{24}$$

From equation 23 we know the equilibrium actions maximize the generalized Hamiltonian of time t at time t. Thus we must have $\frac{\partial \mathcal{H}_t(t)}{\partial a_t^*} = 0$ which yields the incentive condition of the agent at time t:

$$u_a(a^*, \hat{c}) = \frac{\tilde{\psi}(t,t)}{\sigma}.$$

We have found the IC of the agent now we will pin down how the agent's continuation payoff evolves. Observe that, first order adjoint process that is equal to the equilibrium value function $J(t, \Gamma^*)$. We represent the equilibrium value function in the integral form as follows:

$$W(t,k) = r(t,T)u(0,c_T) - \int_k^T \frac{a_t^*}{\sigma} \tilde{\psi}(t,s) + r(t,s)u(a_s^*, \hat{c}_s) ds + \int_r^T \tilde{\psi}(t,s) dZ_s^0.$$

Taking $k \to t$ we have

$$W(t,t) = r(t,T)u(0,c_T) - \int_t^T \frac{a_t^*}{\sigma} \tilde{\psi}(t,s) + r(t,s)u(a_s^*, \hat{c}_s) ds + \int_t^T \tilde{\psi}(t,s) dZ_s^0.$$

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

$$\bar{W}(t) = r(t,T)u(0,c_T) - \int_t^T \frac{a_t^*}{\sigma} \tilde{\psi}(t,s) + r(t,s)u(a_s^*, \hat{c}_s) ds + \int_t^T \tilde{\psi}(t,s) dZ_s^0,$$

and under the regularity conditions we have $\bar{W}(t) = W(t,t)$ and $\bar{W}(t) = J(t, \Gamma^*)$ for all t. Since each agent t only act once, for the principal relevant object is $\bar{W}(t)$. Natural question becomes whether there is a forward representation for $\bar{W}(t)$ or not. We will follow Wang and Yong (2019) and use the decoupling field of the FBSVIE to pin down the evolution of $\bar{W}(t)$. In particular under the Lipschitz continuity of both the forward and backward coefficients we can represent the decoupling field θ as follows:

Theorem 3 (Wang and Yong 2019) *Let $\theta(t,s,\Gamma)$ be the solution to the following PDE*

$$\theta_s(t,s,\Gamma)+\frac{1}{2}\left(\frac{a_s^*}{\sigma}\Gamma_s\right)^2\theta_{\Gamma\Gamma}(t,s,\Gamma)+\Gamma_s\left(\frac{a_s^*}{\sigma}\right)^2\theta_\Gamma(t,s,\Gamma)+r(t,s)u(a_s^*,\hat{c}_s)=0, \tag{25}$$

$$\theta(t,T,\Gamma)=r(t,T)u(\underline{a},\hat{c}_T). \tag{26}$$

Then we have $\bar{W}(t)=\theta(t,t,\Gamma_t)$.

Notice that the decoupling field allows us to represent the integral equation defining the equilibrium value in terms of the forward component Γ_t. Also as noted in Wei et al. (2017) the returns in the objective, $\int_t^T \Gamma_s r(t,s)u(a_s,\hat{c}_s)ds$, does not rely on the backward terms thus for any *choice* of $\psi(t,t)$ the PDE is linear with t as a parameter.[24] Furthermore, given that the forward part is a standard SDE, it allows us to represent the equilibrium value as a single process. Below we show that the representation PDE can be reduced to a non-homogeneous heat equation. By Ito's rule evolution of $\bar{W}(t)$ can be written as:

$$d\bar{W}(t)=\theta_1(t,t,\Gamma_t)+\theta_2(t,t,\Gamma_t)+\frac{1}{2}\left(\frac{a_s^*}{\sigma}\Gamma_s\right)^2\theta_{\Gamma\Gamma}(t,t,\Gamma_t), \tag{27}$$

where $\theta_1(\theta_2)$ denotes the partial derivative respect to first (second) argument. If we follow the decoupling approach in the Sannikov (2008), we would have end up with the terms $\theta_1(t,t,\Gamma_t)$ and $\theta_{\Gamma\Gamma}$. The term $\theta_2(t,t,\Gamma_t)$ captures the effect how different selves evaluates the equilibrium value.

Transformation of the Decoupling Field to the Non-Homogenous Heat Equation

In this section, we are going to show how to translate the decoupling field θ to a non-homogenous heat equation.
Step 1: First we translate the terminal value problem to an initial value problem.
First we introduce a reversal of time $\tau\to T-s$, and define $\theta(t,T-s,\Gamma)=\tilde{\theta}(t,\tau,\Gamma)$. Then our PDE identifying the decoupling field now has an initial condition instead of a terminal condition and is equivalently represented as:

$$\tilde{\theta}_\tau(t,\tau,\Gamma)=\frac{1}{2}\left(\frac{a_\tau^*}{\sigma}\Gamma_\tau\right)^2\tilde{\theta}_{\Gamma\Gamma}(t,\tau,\Gamma)+\Gamma_\tau\left(\frac{a_\tau^*}{\sigma}\right)^2\tilde{\theta}_\Gamma(t,\tau,\Gamma)+r(t,\tau)u(a_\tau^*,\hat{c}_\tau), \tag{28}$$

$$\tilde{\theta}(t,0,\Gamma)=r(t,T)u(\underline{a},\hat{c}_T)\Gamma. \tag{29}$$

Step 2: In this step we conduct change of variables to reduce our problem to turn our problem to a problem of convective mass transfer (see Polyanin and Nazaikinskii (2015) section 3.1.4)
Let us introduce the change of variable $\Gamma\to e^x$ and consider the function $\tilde{\theta}(t,\tau,\Gamma)=v(t,\tau,x)=v(t,\tau,log(\Gamma))$. Then the PDE identifying the decoupling field is equivalently

[24] We refer the readers to Wei et al. (2017) and Yong (2012) for a lengthier discussion.

represented as:

$$v_2(t,\tau,x) = \frac{1}{2}\left(\frac{a^*_\tau}{\sigma}\right)^2 v_{xx}(t,\tau,x) + \left(\left(\frac{a^*_\tau}{\sigma}\right)^2 - 1\right) v_x(t,\tau,x) + r(t,\tau)u(a^*_\tau, \hat{c}_\tau), \tag{30}$$

$$v(t,0,\Gamma) = r(t,T)u(\underline{a}, \hat{c}_T)e^x. \tag{31}$$

Step 3: In this step we conduct another change of variables to reduce our problem of convective mass transfer to a non-homogenous heat equation

In this step let us introduce another change of variable $x \to \left(z - \left(\left(\frac{a^*_\tau}{\sigma}\right)^2 - 1\right)\tau\right)$ and consider the function $v(t,\tau,\left(x + \left(\left(\frac{a^*_\tau}{\sigma}\right)^2 - 1\right)\tau\right)) = y(t,\tau,z)$. Then the PDE identifying the decoupling field is equivalently represented as :

$$y_2(t,\tau,z) = \frac{1}{2}\left(\frac{a^*_\tau}{\sigma}\right)^2 y_{zz}(t,\tau,z) + r(t,\tau)u(a^*_\tau, \hat{c}_\tau), \tag{32}$$

$$y(t,0,z) = r(t,T)u(\underline{a}, \hat{c}_T)e^z. \tag{33}$$

Notice now the decoupling field described above is a non-homogenous heat equation with the non-homogenous term $r(t,s)u(a^*_s, \hat{c}_s)$ and only an initial condition prescribed. This is not surprising as the first two steps are analogous to the ones that are used to transform the Black-Scholes equation to a heat equation.

A.3 Proof of Theorem 1

The proof of the theorem involves heuristic derivation of the HJB equation, two auxiliary steps and the verification argument.

Derivation of the Extended HJB System

In this section of the appendix we derive the extended HJB equation. This derivation is analogous to the derivation provided in Yong (2012) and Björk et al. (2017) so we provide only a short argument. By construction, of the h_Δ we have

$$\begin{aligned} J(t+\Delta, W_{t+\Delta}, h_\Delta) = J(t,W,h) - \mathrm{E}^{h_\Delta}_{t,W}\left[\int_t^{t+\Delta} R(t,s)\,(a_s - c_s)\,ds\right] \\ + \mathrm{E}^h_{t,W} f(t+\Delta, W^h_{t+\Delta}, t+\Delta) - \mathrm{E}^h_{t,W} f(t+\Delta, W^h_{t+\Delta}, t). \end{aligned}$$

Since $J(t+\Delta, W_{t+\Delta}, h_\Delta) = V(t+\Delta, W_{t+\Delta})$, we can rewrite the equation as follows:

$$\begin{aligned} \mathrm{E}^h_{t,W} V(t+\Delta, W_{t+\Delta}) = V(t,W_t) - \mathrm{E}^{h_\Delta}_{t,W}\left[\int_t^{t+\Delta} R(t,s)\,(a_s - c_s)\,ds\right] \\ + \mathrm{E}^h_{t,W} f(t+\Delta, W^h_{t+\Delta}, t+\Delta) - \mathrm{E}^h_{t,W} f(t+\Delta, W^h_{t+\Delta}, t). \end{aligned}$$

Add and subtract $f(t,W,t)$ to the right hand side rearrange by dividing both sides to Δ and take the limit as $\Delta \to 0$. So, we reach

$$\sup_h \{\mathcal{A}^h V(t,W) + (a_t - c_t) - \mathcal{A}^h f(t,W,t) + \mathcal{A}^h f^t(t,W)\} = 0.$$

Moreover for every fixed W,t, $f^s(t,W)$ corresponds to a martingale therefore it must satisfy the following PDE

$$\mathcal{A}^{h^*} f^s(t,W) + R^s(t)\,(a_t^* - c_t^*) = 0.$$

In addition it must satisfy the following boundary conditions

$$\begin{aligned} V(T,W) &= B(T,W_T) \text{ for all } W, \\ f(T,W,T) &= B(T,W_T) \text{ for all } W. \end{aligned}$$

so we reach the extended HJB system of theorem 1.

The Existence of a Solution to the Backward System

In order to prove existence we first start with an arbitrary incentive compatible control law $\hat{h}$ and consider the second part of the HJB system, the backward equation:

$$\mathcal{A}^{\hat{h}} f^s(t,W) + R^s(t)(a_t^{\hat{h}} - c_t^{\hat{h}}) = 0.$$

For each t, w the backward equation

$$\mathcal{A}^{\hat{h}} f^t(t,W) + R^t(t)(a_t^{\hat{h}} - c_t^{\hat{h}}) = 0$$

is a semilinear parabolic Partial Differential Equation (PDE). Observe that if the PDE had a solution $f^t_{\hat{h}}(t,W)$ then for every s we could consider a version of W^s that reaches W by time t. This would imply we would have $f^s_{\hat{h}}(t,W)$, where we underline the dependence on the control $\hat{h}$. An equivalent representation of the backward system is

$$f^t_{\hat{h}}(t,w) = f^t_{\hat{h}}(T,w_T) - \int_t^T R^t(r)(a_r^{\hat{h}} - c_r^{\hat{h}})dr - \int_t^T Y_r^{\hat{h}} dZ_r.$$

For an adapted process $Y^{\hat{h}}$. A solution to the backward system alone would correspond to a pair of processes $f^t_{\hat{h}}(t,W)$ and $Y_t^{\hat{h}}$. However, notice that for a given arbitrary control $\hat{h}$, the c_t is pinned down by the incentive condition. In particular, observe that $a_t \in [\underline{a}, \bar{a}]$ and due to assumption 3 for any given a_t, ψ_t, W_t there is a unique c_t, denoted by $c_t^{IC}(a_t,\psi_t)$. Moreover given that any control is a mapping $\hat{h}(t,W_t)$ the backward system is accompanied by a forward system

$$dW(t) = \theta_1(t,t,\Gamma_t) + \theta_2(t,t,\Gamma_t) + \frac{1}{2}\left(\frac{a_s^*}{\sigma}\Gamma_s\right)^2 \theta_{\Gamma\Gamma}(t,t,\Gamma_t), \tag{34}$$

where θ denotes the decoupling field of the agent's BSVIE. However, also notice that the solution to the backward system $f^t_{\hat{h}}(t, W)$ and $Y^{\hat{h}}_t$ does not appear in the forward system, hence this is a *decoupled* forward-backward stochastic differential equation (FBSDE). For a textbook treatment of FBSDE's we refer the readers to Ma et al. (1999). Our goal is to ensure that the for any incentive compatible control $\hat{h}$ the FBSDE system is well-posed and therefore has a unique solution, and that solution is also continuous.

Given that the system is decoupled, first let us write down the generator of the backward system. Since we are only looking at Markovian, incentive compatible controls, the generator only depends on the forward part:

$$R^t(s)(a_s^{\hat{h}(W_t)} - c_s^{\hat{h}(W_t)}).$$

From Cvitanic and Zhang (2012), section 9.5, proposition 9.5.2 we know that to have the FBSDE well posed we need to establish that the forward component has a unique solution that satisfies the Markov Property and the backward part has a unique solution. For W_t analogous to Sannikov (2008) under Markovian controls lemma 1 will also have a unique, Markovian solution. Now given that the forward system is Markovian, we need to establish that the backward system is well posed. Now we observe that since the backward system does not have the backward terms in the generator but only the forward ones, the generator $R^t(s)(a_s^{\hat{h}(W_t)} - c_s^{\hat{h}(W_t)})$ is trivially uniformly Lipschitz continuous in $V^{\hat{h}}_t$ and Y_t.[25] Thus by theorem 9.3.5 of Cvitanic and Zhang (2012) the backward system has a unique solution. Now since the backward system has a unique solution and the forward system has a unique Markovian solution by proposition 9.5.2 of Cvitanic and Zhang (2012) the FBSDE is well posed and there is a triple $W^{\hat{h}}$, $f^t_{\hat{h}}(t, W)$ and $Y^{\hat{h}}_t$ for any incentive compatible Markovian control.

Finding a Fixed Point to The Extended HJB System

From the previous part we know that for incentive compatible control $\hat{h}$ there exists a unique process $V^{\hat{h}} = f^t_{\hat{h}}(t, w)$ that satisfies the backward system. For any incentive compatible control $\tilde{h}$, consider the infinitesimal generator with control $\tilde{h}$ on the process $f^t_{\hat{h}}(t, W)$, denoted by $\mathcal{A}^{\tilde{h}} f^t_{\hat{h}}(t, W)$. Fix a time t and a state W pair and consider the static optimization problem solved for every t and W,

$$\sup_{\tilde{h}} \{(a_t^{\tilde{h}} - c_t^{\tilde{h}}) - \mathcal{A}^{\tilde{h}} f^t_{\hat{h}}(t, W)\} = 0.$$

Essentially, for any given control $\hat{h}$ the sup above generates another control h using the value function generated from the backward system. So, if one were to consider a mapping from controls to controls we are trying to find a fixed point of such a mapping, with additional difficulties arising from the filtered information structure and rather large control space. The difficulty here is that we do not have convexity of the best response correspondence and we do not have quasi-concavity of the utility function on ones strategy.

[25]The backward terms not appearing in generator is not a suprising when the backward system is a sort of "continuation utility", see Duffie and Epstein (1992), El Karoui et al. (1997)

Remember that we had defined the dynamically inconsistent control problems as the limit of a game between time t selves. In case of finitely many such selves the existence of an equilibria can be found by backward induction. However as noted in Wei et al. (2017) taking the formal limit for infinitely many such selves in general is very difficult. Well-posedness is established in Wei et al. (2017) for cases where the diffusion term is not controlled, hence their methods are not applicable. Existence and uniqueness remain very much an open problem in general dynamically inconsistent stochastic optimal control problems as noted by Björk et al. (2017).

In order to tackle this difficulty we define the game between the selves properly and search for existence of equilibria in the corresponding game. Towards that goal we utilize techniques from static games with non-atomic players such as the ones explored in Schmeidler (1973), Mas-Colell (1984), Khan and Sun (2002) and also closely related to mean-field games Guéant et al. (2011) and aggregate games Acemoglu and Jensen (2015). In particular, we will translate our game to the incomplete information game identified in Balder (1991). [26]

Consider the exogenously random parts of the model $\Omega = Z_{[0,T]}$ where Z is the standard Brownian motion that governs the randomness in the M and W. The probability triple is denoted by $(\Omega, \mathcal{P}, \mathcal{F})$. Notice that when defined this way, a realization ω identifies a whole path of the exogenous random factors. W can be identified by ω and given controls. Let $S = A \times C \times \mathbb{R}$ denotes the space of strategies. The player space is identified as $[0, T]$. $\mathcal{P}$ is common knowledge among the players. We characterize players having incomplete observation of a realization as differential information. In particular, the information component of a player t is characterized by a sub $\sigma-$ algebra of $\mathcal{F}$, denoted by $\mathcal{F}_t$, which corresponds to the natural filtration since our players' observation is differing according to time. Let $\mathfrak{M}$ denote the space of all measures on S. Furthermore, let $\mathfrak{S}$ denote the set of all measurable decision rules δ, $\delta : \Omega \to \mathfrak{M}$. Hence, players are allowed to randomize.

We equip $\mathfrak{S}$ with the weak topology, which is the weakest topology where functions identified below are continuous on $\mathfrak{S}$:

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

Where $L^1(\Omega, \mathcal{P}, \mathcal{F})$ denote the space of $\mathcal{P}$ integrable functions and $C_B(A)$ denote the space of bounded and continuous functions, we also assume with the usual L^1 norm on $(\Omega, \mathcal{P}, \mathcal{F})$. With this notation a pure strategy profile is a distribution over $\mathfrak{S}$, we denote the set of all probability distributions over $\mathfrak{S}$ as $M(\mathfrak{S})$.

Finally for each t let $\mathfrak{S}_t \subset \mathfrak{S}$ denote the set of all $\mathcal{F}_t$ measurable decision rules δ, $\delta : \Omega \to \mathfrak{M}$. Now from here we 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.[27] Due to lemma 2 of Balder (1991) D is $\mathcal{F} \times \mathcal{B}(\mathfrak{S})$ measurable

[26]See also, Jovanovic and Rosenthal (1988) and Balder (2002).

[27] Balder (1991) explores more general sub σ-algebras in order to introduce the pointwise convergence topology which is due to Cotter (1986) for measurability of conditional expectations.

and $\mathfrak{S}_t$ is a compact subset of $\mathfrak{S}$ for every $\mathcal{F}_t \in \{\mathcal{F}_t\}_{t\in[0,T]}$. [28] The restriction of the σ-algebra $\mathcal{F}$ to D is the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$.

First we observe that any control h defined in the first part satisfies $h \in M(\mathfrak{S})$ with the additional property that each $h_t \in \mathfrak{S}_t$. Observe that for any given t, ω and $\mathcal{F}_t$ using the backward system for any control we can define:

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

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

$$u_t(\delta, h|\mathcal{F}_t) = \mathbb{E}\left[(a(\delta) - c(\delta)) - \mathcal{A}^{(a(\delta),c(\delta))} f_h^t(t, W) | \mathcal{F}_t \right].$$

First we make the observation that u_t is continuous in both its arguments and is measurable by definition. Notice that with the randomness baked in to the $f^t(t,x)$ part, a player is identified by the characteristics $(t, \mathcal{F}_t)$, however we have not restricted δ to be $\mathcal{F}_t$ measurable.

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

Now, we can identify a *characteristic - strategy*(CS) distribution, which is a distribution over $[0,T] \times D$ which specifies the way possible characteristic-strategy combinations are distributed in the game. Letting λ be CS distribution the respective marginals of the distribution $\lambda_{|[0,T]\times\mathcal{F}}$ coincides with a distribution over characteristics and $\lambda_{|\mathfrak{S}}$ coincides with a distribution over strategies.

Definition 8 *A CS distribution λ is an equilibrium if*

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

where λ_U^T denote the uniform distribution over $[0,T]$.

First, let us observe that in the game between type t selves the distribution over $[0,T] \times \mathcal{F}$ is identified by the filtered probability space generated by $Z_{[0,T]}$. Now, we observe that $\mathfrak{S}$ is defined as above complete and separable, $\mathfrak{S}_t$ is compact as noted above, $U(t, \mathcal{F}_t, \cdot, \cdot)$ is continuous in both arguments and $U(\cdot, \cdot, \cdot, \nu)$ for $\nu \in \mathfrak{S}$ is measurable, hence by Theorem 1 of Balder (1991), there exists an equilibrium distribution. Hence there exists a control law h^* such that for every t, W

$$(a^{h^*}, c^{h^*}) \in \arg\sup_{\tilde{h}} \{a_t^{\tilde{h}} - c_t^{\tilde{h}} - \mathcal{A}^{\tilde{h}} f_{h^*}^t(t, W)\} = 0.$$

This implies there exist a solution to the extended HJB system.

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

Verification Theorem

In this section we are going to prove the verification theorem. Assume that the functions $V(t, W)$, $f^t(t, W)$ and $\hat{h}(t, W)$ have the following properties.

Theorem 4 *Assume the following*

- $V(t, W)$, $f^t(t, W)$ *solves the HJB equation*
- $V(t, W)$ *and* $f^t(t, W)$ *are smooth.*[29]
- $\hat{h}$ *is an admissible control and argmax of the* V *equation*

Proof of Verification

Step 1 Probabilistic interpretation for f
Apply Ito's formula to $f^s(t, W)$ by extended HJB system conditions we know that it has a drift of $R(s, r)(a_s - c_s)$ and use the boundary condition at time T, to write as:

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

Step 2 We are going to show that $V(t, W) = J(t, W, h^*)$ for all (t, W).
Define function H as $H(t, W, t, h) := R(t, t)(a_t - c_t)$. From the extended HJB system

$$\mathcal{A}^h V(t, W) + \mathcal{A}^h f^t(t, W) - \mathcal{A}^h f(t, W, t) + H(t, W, t, h) = 0.$$

Again from the extended HJB system

$$H(t, W, h^*) + \mathcal{A}^{h^*} f^t(t, W) = 0.$$

Then we end up with

$$\mathcal{A}^h V(t, W) = \mathcal{A}^h f(t, W, t),$$

for all t and W. Since V is smooth we can apply Itos lemma,

$$\mathbb{E} V(T, W^h_T) = V(t, W) + \mathbb{E} \left[\int_t^T \mathcal{A}^h V(s, w_s) ds \right],$$

so we can rewrite as

$$\mathbb{E} V(T, W^h_T) = V(t, W) + \mathbb{E} \left[\int_t^T \mathcal{A}^h f(s, w, s) ds \right].$$

Applying the same reasoning $f^t(t, W)$ by the boundary conditions for V and f implies

$$V(t, W) = J(t, W, h^*).$$

[29]It is enough to be in C^2 respect to x and C^1 respect to t.

Step 3 Optimality of h^*
Next step is to show that h^* is an equilibrium control law. Suppose agents use an arbitrary control law $\hat{h}$ over period lenght $\Delta > 0$. Let, $J(t, W, \hat{h}_\Delta)$ the pay-off principal under such strategy should be the case

$$\liminf_{\Delta \to 0} \frac{J(t, W, h) - J(t, W, \hat{h}_\Delta)}{\Delta} \geq 0. \tag{35}$$

Since $J(t + \Delta, W_{t+\Delta}, h_\Delta) = V(t + \Delta, W_{t+\Delta})$, we can rewrite the equation as follows:

$$\begin{aligned} \mathrm{E}^h_{t,W} V(t + \Delta, W_{t+\Delta}) = V(t, W_t) - \mathrm{E}^{h_\Delta}_{t,W} \left[\int_t^{t+\Delta} R^t(s)(a_s - c_s) ds \right] \\ + \mathrm{E}^h_{t,W} f(t + \Delta, W^h_{t+\Delta}, t + \Delta) - \mathrm{E}^h_{t,W} f(t + \Delta, W^h_{t+\Delta}, t). \end{aligned}$$

Then we can write the difference

$$\begin{aligned} \mathrm{E}^h_{t,W} V(t + \Delta, W_{t+\Delta}) = V(t, W_t) - \mathrm{E}^{h_\Delta}_{t,W} \left[\int_t^{t+\Delta} R^t(s) \left(a_s - c_s\right) ds \right] \\ + \mathrm{E}^h_{t,W} f(t + \Delta, W^h_{t+\Delta}, t + \Delta) - \mathrm{E}^h_{t,W} f(t + \Delta, W^h_{t+\Delta}, t). \end{aligned}$$

after simplifications,

$$\begin{aligned} & - \mathrm{E}^h_{t,W} \left[\int_t^{t+\Delta} R^t(s) u_p(a_s - c_s ds) \right] + \mathrm{E}^{\hat{h}_\Delta}_{t,W} \left[\int_t^{t+\Delta} R^t(s)(a_s - c_s) ds \right] \\ & + \mathrm{E}^h_{t,W} f(t + \Delta, W^h_{t+\Delta}, t + \Delta) - \mathrm{E}^h_{t,W} f(t + \Delta, W^h_{t+\Delta}, t) \\ & + \mathrm{E}^{\hat{h}}_{t,W} f(t + \Delta, W^h_{t+\Delta}, t + \Delta) - \mathrm{E}^{\hat{h}}_{t,W} f(t + \Delta, W^h_{t+\Delta}, t) \end{aligned}$$

as $\Delta \to 0$ a and subtract $f(t, W, t)$ to the right hand side rearrange then by 11 its greater or equal to zero.

A.4 Proof of Lemma 2

To proceed with a solution, first we are going to assume that the agent has access to private savings and the contract has to be terminated at a deterministic deadline T.[30] The no-savings condition of the agent implies, by Lemma 3 of He (2011) in the optimal contract

$$u(c_t, a_t) = \gamma W_t,$$

therefore,

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

[30]Savings assumption simplifies the contract but it is not necessary.

In the optimal contract by Martingale representation theorem (for instance see Sannikov (2008)) we can write agent's continuation utility as follows:

$$dW_t = -\gamma\eta W_t \sigma \psi_t dZ_t, \tag{36}$$

moreover in this case agent's incentive compatibility condition becomes

$$\psi_t = a_t. \tag{37}$$

Then by Ito's Rule we can calculate the evolution of $\ln(W)$ as follows

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

A.5 Discounting Functions Dependent Only on Time Difference

In the case of discounting bi-functions that only depend on the time difference, that is $R(s-t)$ instead of $R(t,s)$ we can simplfy the HJB equation. Observe that once the discounting function does not change we can do differentiation under the integral sign and have:

$$\begin{aligned} \mathcal{A}^h f(t,W,t) &= \int_t^T \left(R(s-t)\mathcal{A}^h\mathbb{E}^h_{t,w}\left[a_s - c_s\right] - R'(s-t)\mathbb{E}^h_{t,w}\left[a_s - c_s\right)]\right) ds, \\ &- R(T-t)\mathcal{A}^h\mathbb{E}^h_{t,w}\left[c_T\right] + R'(T-t)\mathbb{E}^h_{t,w}\left[c_T\right] \end{aligned}$$

and

$$\mathcal{A}^h f^t(t,W) = \int_t^T R(s-t)\mathcal{A}^h\mathbb{E}^h_{t,w}\left[(a_s - c_s)\right] ds + R(T-t)\mathbb{E}^h_{t,w}\left[\mathcal{A}^h c_T\right].$$

Therefore, HJB system can be written as

$$\mathcal{A}^h V + (a_t - c_t) + \int_t^T R'(s-t)\mathbb{E}^h_{t,w}\left[(a_s - c_s)\right] ds - R'(T-t)\mathbb{E}^h_{t,w}\left[c_T\right] = 0. \tag{38}$$

A.6 Proofs for Quasi-Hyperbolic Case

In this case, we use simplied HJB system as in equation 38.

Proposition 5 *Under assumptions 4, 5 the value function of the principal satisfies the following HJB system:*

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

subject to (15) and incentive compatibility condition (17) and boundary condition

$$V(T, W) = B(T, W_T) \text{ for all } W \tag{39}$$

where $B(t, W)$ is defined by (6).

We are going to guess the principal's value function has the following functional form

$$F(t, W) = A(t)\ln(-W) + B(t)$$

with the boundary condition $A(T) = \frac{1}{\eta}$. Given our guess,

$$\dot{F}_t = \dot{A}_t \ln(-W) + \dot{B}(t), F_W = \frac{A(t)}{W}, F_{WW} = -\frac{A(t)}{W^2}.$$

By plugging our functional form guess to extended HJB equation we reach

$$1 - a_t - a_t\eta^2\gamma^2 A(t) = 0 \Rightarrow a_t = \frac{1}{1 + \eta^2\gamma^2\sigma^2 A(t)}.$$

To get the $\dot{A}(t)$, we can collect the log terms

$$\dot{A}_t = \frac{1}{\eta}\left(1 + \int_t^T R'(s-t)\, ds + R'(T-t)\right)$$

or equivalently

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

plug in the quasi hyperbolic discounting function

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

also note that this case boundary condition becomes $A(T) = \frac{1}{\eta}$. Depending on the λ action can be non-monotone function over time. Especially if λ is very high. If we are assume exponential discounting with discount rate r, optimal contract becomes

$$\dot{A}_t = \frac{1}{\eta}(1-r)\, e^{-r(T-t)},$$

and

$$A(t) = \frac{2r - 1 + (1-r)e^{-r(T-t)}}{\eta r}.$$

Observe that depending on the $r, \dot{A}_t$ is either always positive or negative, however, it can not be non-monotone. Given A_t

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

we can solve for a_t.

Proof of Proposition 2

Part 1) It is easy to see that by inspection if $\beta = 1$, $\dot{A}(t)$ equals to $\frac{1}{\eta}(\rho+\lambda-1)$. Similarly if$\beta = 0$ $\dot{A}(t)$ equals to $\frac{1}{\eta}(\rho-1)$ which is equivalent to time-consistent solutions.

Part 2) Fix any $t < T$, then look at the limit

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

Which corresponds to solution in Holmstrom and Milgrom (1987) with discount rate equal to $\frac{\rho(\lambda+\rho)}{\beta\lambda+\rho}$.

Parts 3) Recall the closed form solution for $a(t)$:

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

Then it is easy to see that dynamics of $a(t)$ 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

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

It is easy to see that there is a unique point in which sign of the derivative can change (if it ever happens). In particular for t small enough the first term dominates, but for t large enough with $(1-(\rho-\lambda))(1-\beta)$ being negative, the second term will dominate to change the sign of the derivative.

A.7 Proofs for Anticipatory Utility Case

Observe that under assumption 6 we have double integral, but using tower property of condition expectations we can change the order of integration. In particular observe that at any periods t, s $s \geq t$ and for any period $\tau \geq t$, the term $\mathrm{E}((dM_\tau - c_\tau|\mathcal{F}_s)|\mathcal{F}_t)) = \mathrm{E}(dM_\tau - c_\tau|\mathcal{F}_t)$. Then from the anticipatory part we can collect all the terms that have $\mathrm{E}(dM_\tau - c_\tau|\mathcal{F}_t)$ as follows:

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

In addition we have the standard discounted part with $e^{-\rho}$ yielding the discounting function:

$$R^t(s) = R(s-t) = \frac{e^{-\rho(\tau-t)}-e^{-\zeta(\tau-t)}}{\zeta-\rho}$$

Proposition 6 *Under assumptions 4, 5 the value function of the principal satisfies the following HJB system:*

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

subject to (15) and incentive compatibility condition (17) and boundary condition

$$V(T, w) = B(T, W_T) \text{ for all } w \tag{40}$$

where $B(t, W)$ is defined by (6).

We are going to guess the principal's value function has the following functional form

$$F(t, W) = A(t)\ln(-W) + B(t)$$

with the boundary condition $A(T) = \frac{1}{\eta}$. Given our guess,

$$\dot{F}_t = \dot{A}_t \ln(-W) + \dot{B}(t), F_W = \frac{A(t)}{W}, F_{WW} = -\frac{A(t)}{W^2}.$$

By plugging our functional form guess to extended HJB equation we reach

$$1 - a_t - a_t\eta^2\gamma^2 A(t) = 0 \Rightarrow a_t = \frac{1}{1 + \eta^2\gamma^2\sigma^2 A(t)}.$$

To get the $\dot{A}(t)$, we can collect the log terms

$$\dot{A}_t = \frac{1}{\eta}\left(1 + \int_t^T R'(s-t)\, ds + R'(T-t)\right)$$

or equivalently

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

Plugging in $R(T-t)$ and $R'(T-t)$ we have

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

Using the boundary condition $A(T) = 1/\eta$ and integrating

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

The optimal action is equal to

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

Proof of Proposition 4

Part 1) Fix any $t < T$, then look at the limit,

$$\lim_{\zeta\to\infty} a(t) = \frac{1}{1+\eta\gamma^2\sigma^2\frac{2\rho-1+(1-\rho)e^{-\rho(T-t)}}{\rho}}$$

which corresponds to solution in Holmstrom and Milgrom (1987) with discount rate equal to ρ.
Part 2) As $T \to \infty$, we can solve for r in the equation

$$\frac{2r-1}{r} = \frac{2\rho\zeta-1-\zeta}{\rho\zeta} \Rightarrow r = \frac{\rho\zeta}{1+\zeta}$$

Part 3) Recall that the closed form of a_t is given by

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

Then it is easy to see that the dynamics of a_t is 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 have

$$\begin{aligned}
(1-\rho)e^{-\rho(T-t)} - (1-\zeta)e^{-\zeta(T-t)} &= (\zeta-\rho)(\rho-1)e^{-\rho(T-t)} \\
(1+\zeta-\rho)(1-\rho)e^{-\rho(T-t)} &= (1-\zeta)e^{-\zeta(T-t)} \\
1 &= \frac{\zeta-1}{\rho-1}\frac{1}{1+\zeta-\rho}e^{-(\zeta-\rho)(T-t)}.
\end{aligned}$$

Now since we assume $\zeta > \rho > 1$ we have $\frac{\zeta-1}{\rho-1}\frac{1}{1+\zeta-\rho} > 1$ on the other hand we also have $1 \geq e^{-(\zeta-\rho)(T-t)} \geq e^{-(\zeta-\rho)T} > 0$. Observe that if T is small enough then the derivative never changes sign, but for any ζ, ρ there exists T large enough such that there exists a $t^\star < T$ where the derivative changes sign at $t^\star$.

B Online Appendix (Online Publication Only)

In this section we show that there exists an equivalent time consistent problem for the principal with a different objective function.

Proposition 7 *For every dynamically inconsistent principal in this model, there is a time consistent principal with a different objective function*

Suppose for now assume that there exist an optimal contract with equilibrium control law h, then using the h we can construct the following function K as follows:

$$K(t, W, h) = R(t, t)(a_t^* - c_t^*) - \mathcal{A}^h f(t, W, t) + \mathcal{A}^h f^t(t, W)$$

Given the above K, maximizing the following problem becomes standard time consistent optimal control problem.

$$\mathrm{E}_{t,W} \int_t^T K(s, W, h_s) ds$$

Major caveat of this approach is that once needed to use the equilibrium h to construct the time-consistent optimal control problem.

B.1 Optimal Contract Without Hidden Savings

Suppose agent does not have an hidden savings account.

$$dW_t = -\left(\frac{\partial r^t(t)}{\partial t} W_t + u(c_t, a_t)\right) dt + \psi_t \left(dM_t - a_t dt\right)$$

In this case incentive compatibility implies

$$\max_{a_t} a_t \psi_t \frac{\partial r^t(t)}{\partial t} W_t + u(c_t, a, t)$$

Therefore first order condition implies

$$u(a_t, c_t) = \frac{\partial r^t(t)}{\partial t} W_t \frac{\psi_t}{a_t}$$

Note this implies dW_t is a still a geometric Brownian motion

$$dW_t = \frac{\partial r^t(t)}{\partial t} W_t \left(1 - \frac{\psi_t}{a_t}\right) dt - \frac{\partial r^t(t)}{\partial t} \eta W_t \sigma \psi_t dZ_t$$

In this case

$$\mathbb{E}\left[\ln(-W_t)\right] = \ln(-W_0) - \int_0^t \frac{\partial r^s(s)}{\partial s} \left(1 - \frac{\psi_s}{a_s}\right) ds + \frac{1}{2} \int_0^t \eta^2 (\frac{\partial r^s(s)}{\partial s})^2 \sigma^2 \psi_s^2 ds.$$

Therefore

$$c_t = \frac{1}{2}a_t^2 - \frac{\ln\left(\frac{\partial r^t(t)}{\partial t}\right)}{\eta} - \frac{1}{\eta}\ln(-W_t) - \frac{1}{\eta}\ln\left(\frac{\psi_t}{a_t}\right)$$

$$\sup_{a_t,\psi_t} F_t + a_t - \left[\frac{1}{2}a_t^2 - \frac{\ln\left(\frac{\partial r^t(t)}{\partial t}\right)}{\eta} - \frac{1}{\eta}\ln(-W_t) - \frac{1}{\eta}\ln\left(\frac{\psi_t}{a_t}\right)\right] + \frac{1}{2}\left(\frac{\partial r^t(t)}{\partial t}\eta W\psi\sigma\right)^2 F_{WW}$$
$$+\int_t^T R'(s-t)\left(a_s - \left[\frac{1}{2}a_s^2 - \frac{\ln\left(\frac{\partial r^t(t)}{\partial t}\right)}{\eta} - \frac{1}{\eta}\ln(-W_s) - \frac{1}{\eta}\ln\left(\frac{\psi_s}{a_s}\right)\right]\right) ds$$
$$F_W\frac{\partial r^t(t)}{\partial t}W_t\left(1-\frac{\psi_t}{a_t}\right) + \frac{\ln(-W_T)}{\eta}R'(T-t) = 0.$$

We are going to guess the principal's value function has the following functional form

$$F(t,W) = A(t)\ln(-W) + B(t)$$

Given our guess,

$$\dot{F}_t = \dot{A}_t\ln(-W) + \dot{B}(t), F_W = \frac{A(t)}{W}, F_{WW} = -\frac{A(t)}{W^2}.$$

First order conditions becomes (respect to a_t)

$$1 - a_t - \frac{1}{\eta}\frac{1}{a_t} + \frac{\partial r^t(t)}{\partial t}\frac{\psi_t A(t)}{a_t^2} = 0$$

and respect to ψ_t

$$-\frac{1}{\eta}\frac{1}{\psi_t} + \psi_t(\frac{\partial r^t(t)}{\partial t})^2\eta^2\sigma^2 A(t) - \frac{\partial r^t(t)}{\partial t}\frac{A(t)}{a_t} = 0$$

Again we can get the $A(t)$ matching the ln terms

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

Now assume that principal is time consistent with the discount factor r. So $A(t)$ reduces to the following

$$A(t) = \frac{2r - 1 + (1-r)e^{-r(T-t)}}{\eta r}$$

In that case we can solve for a and ψ, we have two equations and two unknowns. There are no closed form solutions, but equations are easy to solve numerically. Since, $\frac{\partial r^t(t)}{\partial t} = -(\gamma + (1-\beta)\lambda)$. If agent is dynamically inconsistent problem is equivalent to a time-consistent problem in which agent has a discount factor $\gamma + (1-\beta)\lambda$.

References

Abreu, D., D. Pearce, and E. Stacchetti (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica: Journal of the Econometric Society*, 1041–1063.

Acemoglu, D. and M. K. Jensen (2015). Robust comparative statics in large dynamic economies. *Journal of Political Economy 123*(3), 587–640.

Aghion, P., M. Dewatripont, and P. Rey (1994). Renegotiation design with unverifiable information.

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychological bulletin 82*(4), 463.

Amador, M., I. Werning, and G.-M. Angeletos (2006). Commitment vs. flexibility. *Econometrica 74*(2), 365–396.

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

Balder, E. J. (2002). A unifying pair of cournot–nash equilibrium existence results. *Journal of Economic Theory 102*(2), 437–470.

Basak, S. and G. Chabakauri (2010). Dynamic mean-variance asset allocation. *Review of Financial Studies 23*(8), 2970–3016.

Bernheim, B. D. and D. Ray (1986). On the existence of markov-consistent plans under production uncertainty. *The Review of Economic Studies 53*(5), 877–882.

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.

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.

Bond, P. and G. Sigurdsson (2017). Commitment contracts. *Review of Economic Studies 85*(1), 194–222.

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 (2016). Dynamic savings choices with disagreements. Technical report, National Bureau of Economic Research.

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.

Caplin, A. and J. Leahy (2006). The recursive approach to time inconsistency. *Journal of Economic Theory 131*(1), 134–156.

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

Cotter, K. D. (1986). Similarity of information and behavior with a pointwise convergence topology. *Journal of Mathematical Economics 15*(1), 25–38.

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

Dewatripont, M. (1988). Commitment through renegotiation-proof contracts with third parties. *Review of Economic Studies 55*(3), 377–390.

Duffie, D. and L. G. Epstein (1992). Stochastic differential utility. *Econometrica: Journal of the Econometric Society*, 353–394.

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

Ekeland, I. and T. A. Pirvu (2008). Investment and consumption without commitment. *Mathematics and Financial Economics 2*(1), 57–86.

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

Eliaz, K. and R. Spiegler (2006). Contracting with diversely naive agents. *Review of Economic Studies 73*(3), 689–714.

Fischer, S. (1980). Dynamic inconsistency, cooperation and the benevolent dissembling government. *Journal of Economic Dynamics and Control 2*, 93–107.

Freeman, R. B. (1999). The economics of crime. *Handbook of labor economics 3*, 3529–3571.

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

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

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

Gottlieb, D. and X. Zhang (2018). Long-term contracting with time-inconsistent agents. Working paper.

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 (2019). Investment timing and incentive costs. *Review of Financial Studies*. forthcoming.

Guéant, O., J.-M. Lasry, and P.-L. Lions (2011). Mean field games and applications. In *Paris-Princeton lectures on mathematical finance 2010*, pp. 205–266. Springer.

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

Hamaguchi, Y. (2019). Small-time solvability of a flow of forward-backward stochastic differential equations. *arXiv preprint arXiv:1902.11178*.

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, 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.

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

Iijima, R. and A. Kasahara (2016). Gradual adjustment and equilibrium uniqueness under noisy monitoring. Working paper.

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.

Jovanovic, B. and R. W. Rosenthal (1988). Anonymous sequential games. *Journal of Mathematical Economics 17*(1), 77–87.

Khan, M. A. and Y. Sun (2002). Non-cooperative games with many players. *Handbook of game theory with economic applications 3*, 1761–1808.

Kocherlakota, N. R. (1996). Reconsideration-proofness: A refinement for infinite horizon time inconsistency. *Games and Economic Behavior 15*(1), 33–54.

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.

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

Lindensjö, K. (2016). Time-inconsistent stochastic control: solving the extended hjb system is a necessary condition for regular equilibria. *arXiv preprint arXiv:1611.02902*.

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.

Ma, J., J.-M. Morel, and J. Yong (1999). *Forward-backward stochastic differential equations and their applications*. Number 1702. Springer Science & Business Media.

Marinovic, I. and F. Varas (2018). CEO horizon, optimal pay duration, and the escalation of short-termism. *Journal of Finance*. forthcoming.

Mas-Colell, A. (1984). On a theorem of schmeidler. *Journal of Mathematical Economics 13*(3), 201–206.

Mian, A. and J. A. Santos (2018). Liquidity risk and maturity management over the credit cycle. *Journal of Financial Economics 127*(2), 264–284.

Miller, D. A., T. E. Olsen, and J. Watson (2018). Relational contracting, negotiation, and external enforcement. *NHH Dept. of Business and Management Science Discussion Paper* (2018/8).

Obara, I. and J. Park (2017). Repeated games with general discounting. *Journal of Economic Theory 172*, 348–375.

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

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

Pan, J., C. S. Webb, and H. Zank (2015). An extension of quasi-hyperbolic discounting to continuous time. *Games and Economic Behavior 89*, 43–55.

Peleg, B. and M. E. Yaari (1973). On the existence of a consistent course of action when tastes are changing. *The Review of Economic Studies 40*(3), 391–401.

Peng, S. (1990). A general stochastic maximum principle for optimal control problems. *SIAM Journal on control and optimization 28*(4), 966–979.

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

Polyanin, A. D. and V. E. Nazaikinskii (2015). *Handbook of linear partial differential equations for engineers and scientists*. Chapman and hall/crc.

Quah, J. K.-H. and B. Strulovici (2013). Discounting, values, and decisions. *Journal of Political Economy 121*(5), 896–939.

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

Sadzik, T. and E. Stacchetti (2015). Agency models with frequent actions. *Econometrica 83*(1), 193–237.

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

Schmeidler, D. (1973). Equilibrium points of nonatomic games. *Journal of statistical Physics 7*(4), 295–300.

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

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

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters 8*(3), 201–207.

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 Z. Wu (2015). Time-inconsistent optimal control problem with random coefficients and stochastic equilibrium hjb equation. *Mathematical Control & Related Fields 5*(3), 651–678.

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

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

Xu, Q. (2018). Kicking maturity down the road: early refinancing and maturity management in the corporate bond market. *Review of Financial Studies 31*(8), 3061–3097.

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

Yong, J. (2008). Well-posedness and regularity of backward stochastic volterra integral equations. *Probability Theory and Related Fields 142*(1-2), 21–77.

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.