Weighted Discounting, Time Inconsistent Stopping and Their Applications

Wei Wei
St Hugh's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

2016/11/28

To my wife,

Yijing Song.

Acknowledgements

First of all, I want to thank my supervisor Prof Xunyu Zhou, who introduced me to conducting research on time inconsistency. I learned a lot from his deep insights into control theory as well as the helpful discussions with him. He also gave me the space to pick up a topic according my own interests. Further, I am also gratefule to him for supporting me in all my academic projects and for his help in both academic and non-academic life.

I thank Prof Hanqing Jin for generously permitting me to present my work in his reading group. The discussions with him and the other members, Siyuan Li, Yunfei Peng and Pengyu Wei, in the reading group sharpened my understanding of mathematical finance.

Without Prof Ebert at Tilburg University, there would be no stochastic dominance and group diversity, which is the fundamental building block of this thesis. I therefore thank Prof Ebert for sharing his economic insights and for further creating the space and atmosphere for our many discussions at Tilburg University. It is the visit to Prof Ebert that opens the world of behavioral economics for me.

I am also deeply indebted to Professor Lishang Jiang, who served as my supervior in my MPhil studies. I thank him for introducing me to conducting academic research. The skills in partial differential equations, which played, is playing and will play a key role in my academic career, are totally attributed him. His scholarship and his integrity will be always carved in my memory.

I want to thank Dr Gechun Liang at King's College London. I have had a very pleasant collaboration with him when he was a researcher at

Oxford University. From him, I learned a lot about backward stochastic differential equations and optimal stopping and the later one has become a major subject in my thesis.

Most importantly, I thank my family for their long-lasting support and love, especially my wife, Yijing Song, my parents, Huilan Zhou and Guosheng Wei. They may not be experts in my field, but without them it is impossible for me to complete the DPhil studies.

Abstract

This thesis studies groups whose members disagree about the method of discounting. In particular, we provide a comprehensive treatment of discount functions that are given by the weighted average of the group members' discount functions. This class of "weighted discount functions" admits a natural notion of group diversity, which has consequences for behavior. We show that more diverse groups discount less heavily and make more patient decisions. Within a real options framework, for example, greater group diversity leads to delayed investment. Then, we show that well-known behavioral discount functions can be written as a weighted discount function. Therefore, all results in this thesis find a correspondence in a single agent setting with non-standard, behavioral time preferences. In particular, we provide an equilibrium stopping result for individuals who are aware of their time-inconsistency, but lack commitment. Finally, we propose an example which shows that the Bellman system for the equilibrium stopping problem may not admit any proper solution and illustrate that the non-existence is caused by group diversity or decreasing impatience.

Contents

List of Figures

Chapter 1

Introduction

How do firms and public entities make investment decisions? The standard approach is that "the firm" or "the policy maker" compares investment opportunities by their net present values (NPVs), i.e., by the sum of their discounted expected cash flows. This thesis relaxes some of the simplifying assumptions of this approach.

First, and most importantly, we assume that investment decisions are taken by a group of people, rather than by a single individual. Even if a single decision maker owns all the decision power, she may still attach some weight to the opinions of others when making her decision. CEOs consult with board members, experts, or friends. In modern societies, the members of a household make important decisions together. And also presidents collect the opinions of party members, other advisers – or their spouses.

The second issue we wish to address in this paper is the "perennial dilemma of what discount rate to use" (Weitzman 2001, abstract). Attempts to resolve the "discount rate dilemma" go back at least to Ramsey (1928). Knowledge of the appropriate discount rate is of great practical interest for policy decisions, in particular for those concerned with the very long run. Examples include measures against climate change, pension reforms, the level of public debt, investment in infrastructure, or investment in education; see Gollier (2014) for a recent discussion of sustainable investment and cost-benefit analysis. Weitzman's survey evidence from more than 2000 economists – including 52 Nobel prize winners – shows a widely dispersed distribution of what discount rate should be used to evaluate investment.[1] Our model incorporates this evidence by allowing for group diversity which is reflected by differences in opinion about the appropriate method of discounting. In addition to group disagreement, we allow for the possibility that each individual group member has some uncertainty about what discount rate to use.

Third, we explore the relationship between the marginal behavior of a time preference and the weighted discounting. We view a discount function as a utility function of time. This perspective allows us to discuss time preferences within a well-developed framework. Our discussion is based on the law of diminishing marginal utility, which is a cornerstone of the utility theory. Besides that an agent is not sure about what discount rate use, we find another source of weighted discounting for a single agent.

Fourth, we propose a game theoretic formulation to a time inconsistent stopping problem, where the time inconsistency comes from non-exponential discounting. Our setting considered in this thesis is assumed to be a continuous time Markovian economy. Recently, there is a growing literature devoted to building a general framework for time inconsistent control problems in continuous time setting (e.g., Ekeland and Lazrak 2006; Bjork and Murgoci 2010). This thesis, in the same line of the above literature, develops a general framework for time inconsistent stopping problems. In this framework, we derive a Bellman system via a recursive

[1] Frederick et al. (2002) review the experimental literature on time preferences and Falk et al. (2016) present survey evidence representing 90% of the world population. Both studies report substantial heterogeneity in discounting behavior.

approach[2]. Then we verify the Bellman system in the sense that the stopping rule obtained from the Bellman system is the Nash equilibrium outcome of an intra-personal game.

Fifth, our model takes into account the value of timing an investment opportunity optimally, as put forward in the real options approach proposed by Brennan and Schwartz (1985) and McDonald and Siegel (1986). For example, in some situations it can be beneficial to wait and see, rather than to decide for or against an investment right away. The real options literature has neither investigated the investment decisions of groups, nor has it given answers to the discount rate dilemma. This paper offers a solution to both issues. Our contribution to the literature on sustainable investment and cost-benefit analysis is, furthermore, to incorporate insights from the real options literature regarding investment timing.

Finally, we give an example which shows that the Bellman system may not admit any proper solution, even if the coefficients of the system are smooth.

The starting point of this paper is to provide an in-depth study of group discount functions (Chapter 2). Intuitively, the discount function of a group is given by a weighted average of the discount functions of its members, where more important or more valued members receive more weight. For sake of concreteness, we first suppose that there are finitely many group members who are all exponential discounters. The group's *weighting distribution* $F(r)$ assigns a weight to each member or, equivalently, to each (positive) discount rate r. Member i ($i = 1, \ldots, n$) favors discount rate r_i and thus her discount function is given by $e^{-r_i t}$. Letting f denote the probability mass function of F, the *group discount function* (also

[2]The recursive approach is suggested by Strotz (1955) originally and used to derive Bellman systems for many other time inconsistent problems. For example, see Basak and Chabakauri (2010) for a mean-variance case.

weighted discount function) is given by

$$h(t) = \sum_{i=1}^{n} f(r_i) e^{-r_i t}. \tag{1.0.1}$$

The idea of averaging discount functions to model social or group preferences is by no means new and goes back to seminal articles such as Marglin (1963) and Feldstein (1964). This thesis, however, does not model the group dynamics (e.g., through negotiation or voting) with the objective to derive the group's weighting distribution F. Neither do we make moral arguments as to what the weights of group members should be. Instead, this paper sets in once the weighting distribution F is determined, i.e., we take it as exogenously given. Therefore, our paper is complementary to the literature on preference aggregation. Our contribution to that literature is a comprehensive analysis of the class of *weighted discount functions* by proving a number of results about their importance, prevalence, and other properties. In particular, we explain how properties of the weighting distribution F translate into properties of the discount function h, and vice versa.

The first contribution of Chapter 2 is to introduce and characterize a notion of *group diversity,* which is naturally suggested by the definition of a weighted discount function. We define a group as being *more diverse* if its weighting distribution is more dispersed than that of another group, where dispersion is measured using a proper notion of stochastic dominance. Intuitively, greater group diversity means that there are stronger differences of opinions among its members about what discount rate to use. We then characterize group diversity (a property of the weighting distribution F) in terms of a property of the weighted discount function h. We prove that a group is more diverse than another group if and only if, at any future time, it discounts less heavily. We further clarify the relationship between group diversity and the important concepts of patience (Quah and Strulovici 2013) and decreasing impatience (Prelec 2004). Moreover, we show that all weighted discount functions exhibit decreasing impatience as defined by Prelec (2004), which corroborates Jackson and Yariv's (2015) important result that a group discount function as above necessitates a present bias.[3]

The second contribution of Chapter 2 is to clarify the importance and prevalence of weighted discount functions. Note that we may also think of the weighted discount function in equation (1.0.1) as belonging to a *single* individual – an exponential discounter with constant impatience – who is uncertain about what discount rate to use. Therefore, a weighted discount function may reflect not only inter-personal (i.e., group) disagreement about what discount rate to use, but also intra-personal disagreement. One might further wonder whether some of the commonly used discount functions can be written as a weighted discount function. That is, for a given, well-known discount function h, does there exist a distribution of opinions F such that the resulting group discount function is given by h? This turns out to be true. As will be explained in Chapter 2, it can even be argued that "all of the commonly used discount functions" are weighted discount functions. From a mathematical result called Bernstein's Lemma, we readily obtain that exponential (Samuelson 1937), hyperbolic (Harvey 1986), proportional (Mazur 1987; Harvey 1995), generalized hyperbolic (Loewenstein and Prelec 1992), and constant sensitivity (Ebert and Prelec 2007) discount functions are weighted discount functions. The extension of quasi-hyperbolic discounting (Phelps and Pollak 1968; Laibson 1997) to continuous time (Harris and Laibson 2013) are likewise weighted discount functions.[4]

[3]Our result generalizes and refines Weitzman's 2001 observation of time-varying discount rates for the special case where opinions F are Gamma-distributed. Jackson and Yariv (2015) show that with any heterogeneity in time preferences, utilitarian aggregation necessitates a present bias. Gollier and Zeckhauser (2005) have also noted that group heterogeneity results in time-varying discount rates and time-inconsistency. We are the first to point out a relationship between group heterogeneity and decreasing impatience as characterized by Prelec (2004).

[4]Bleichrodt et al. (2009) provide an overview of discount functions used in the literature. The

With this in mind, consider a group whose members are weighted discounters (not necessarily exponential) themselves. Note that, the weighted average of weighted discount functions (being a weighted average of exponential discount functions) is, still, a weighted average of exponential discount functions, and thus a weighted discount function. Therefore, for example, the discount function of a group of hyperbolic discounters (and/ or other weighted discounters) is also a weighted discount function. This argument shows that the "enormously simplifying" Assumption 1 in Weitzman (2001, pp. 263-264), that groups consist of experts "thinking in terms exponential discounting" is unnecessary. His approach to resolve the discount rate dilemma applies more generally to groups of non-standard discounters – as long as they belong to our class of weighted discounters.

The third contribution of Chapter 2 is to explore the issue of why a single agent exhibits weighted discounting. The Bernstein's Lemma provides a good opportunity to approach a weighted discount function via the signs of its all order derivatives. In the well-studied utility theory, the signs of the derivatives of a utility function are closed related to the risk attitudes (e.g., Eeckhoudt and Schlesinger (2006); Deck and Schlesinger (2010)). Particularly, Deck and Schlesinger (2010) propose a simple type of lottery that is used to explain the signs of higher order derivatives and higher order risk attitudes. Recently, Ebert (2015a) views the discount functions as the utility functions of random times. In the same line of Deck and Schlesinger (2010), he introduces the concept prudent and temperate discounting and discusses the economic meaning of the signs of the third and fourth order derivatives of a discount function.

Different from the aforementioned literature, which explains the higher order derivatives of a utility or discount function from the perspective the theory of risk attitudes, our method is based on the law of diminishing marginal utility. Our

two new discount functions proposed in that paper are also weighted discount functions with decreasing impatience.

discussion is motivated by the following question.

You will receive two amounts of money: $1000 two days later and $1000 two years later. You now have to postpone one of them for 100 days. Which one do you incline to delay?

Our result shows that the time preference of a single agent can be characterized by a weighted discount function if and only if the agent's time preference shows "diminishing marginal impatience" towards a specific sequence of receipts in which things are only a matter of time.

The fact that weighted discount functions exhibit decreasing impatience constitutes the main challenge in studying a group's investment decision, which we approach in Chapter 3. Except for the trivial case where the group consists of a single member, in which case the weighted discount function is exponential, the decision problem becomes time-inconsistent (Strotz 1955). For that reason, in our study of group investment behavior, we find ourselves facing similar difficulties as the behavioral literature on timing (or stopping) decisions under time-inconsistency. In particular, we study the group's decision problem within the intra-personal game framework (e.g, Phelps and Pollak 1968; Laibson 1997). When making its decision, the group takes into account that the group's future self may not agree with the planned stopping decision. We thus define equilibrium rules of investment timing that are not deviated from by the group's future selves and, given this restriction, maximize the discounted value of investment. The main result of Chapter 3 is a theorem that characterizes the equilibrium stopping behavior under weighted discounting. Equilibrium behavior is obtained as the solution to a system of Bellman equations, which offers intuitive interpretations. Its equations reflect the fact that time-inconsistency is the consequence of inter- or intra-personal uncertainty about the discount rate. The system involves more constraints on the potential investment opportunity set the greater this uncertainty is.

While our results hold for arbitrary payoff or utility functions, and stopping problems that are not specific to investment, in Section 4.1 we focus on the Bellman system of the standard irreversible investment problem (as in, e.g., Dixit and Pindyck 1994). In that case, the investment decision amounts to *when* to invest, namely, when to pay a fixed cost in exchange for the receipt of the present value of some project that evolves according to a geometric Brownian motion. For general weighted discount functions, we show that investment behavior is characterized by a threshold strategy. This means that it is desirable to invest once the project value is at or beyond a certain level, and to wait if it is beneath that value.

We then focus on some special weighted discount functions to obtain various analytical results. Note that, investment behavior under generalized hyperbolic discounting (Loewenstein and Prelec 1992) comprises an example of investment behavior under weighted discounting. This important special case has not been solved in the literature yet, and thus our results also break new ground in the behavioral literature. In particular, we are the first to obtain results on investment behavior under continuously changing levels of impatience, as is implied by generalized hyperbolic discounting, the constant absolute decreasing impatience (CADI), or the constant relative decreasing impatience (CRDI) functions defined in Bleichrodt et al. (2009). The investment threshold for generalized hyperbolic discounting can be derived in closed form. Moreover, for the discount function being the one proposed by Harris and Laibson (2013) we recover the results of Grenadier and Wang (2007) on the investment behavior of a sophisticated entrepreneur. In particular, their results correspond to those obtained for a discrete-uniform weighting distribution F. For the weighted discount function being the exponential one, our Bellman system for the real options case reduces to a single Bellman equation – the one studied extensively in Dixit and Pindyck (1994).

In Section 4.2, we study the comparative statics of investment behavior (i.e., the level of the threshold). Higher investment cost, expected growth of the project value, and project value volatility all lead to a higher threshold (i.e., later investment), which generalizes extant results for exponential discounting to arbitrary weighted discount functions. Moreover, for the patience ranking of discount functions proposed by Quah and Strulovici (2013), we show that more patient agents invest later.

Furthermore, we study the comparative statics of investment behavior with respect to group diversity. The main result is that greater group diversity leads to *delayed* investment. Waiting longer with making a decision when disagreement is strong seems consistent with daily observations. Further, note that delayed investment is characterized by a higher threshold, and thus comes along with more risk-taking – a bird in the hand (the current NPV of the project) is worth *less* than two in the bush (a potentially higher project NPV). The result that greater group diversity leads to more risk-tolerant investment may be surprising, because greater group diversity comes along with stronger decreasing impatience and thus stronger time-inconsistency. Within the intra-personal game framework, stronger time-inconsistency comes along with a more restricted set of feasible investment strategies and thus, usually, with *less* risk-taking. However, the direct effect of greater group diversity leading to less heavy discounting dominates the effect of more time-inconsistency. Overall, more diverse groups are more patient in their investment decisions in the sense that they are willing to wait, looking for a yet more profitable time to invest.

In chapter 5, we show a solvable example, whose time consistent counterpart is used to analyze monopoly or oligopoly investment under uncertainty[5]. In the conventional time consistent case, the standard solving scheme for obtaining a solution to a Bellman system for an optimal stopping problem is the so called

[5]See, for example, Back and Paulsen (2009) and Dixit and Pindyck (1994).

"value matching + smooth pasting". And if the coefficients of the Bellman system are smooth, the solution to the Bellman system exists[6]. However, by the example of this chapter, it is to see that

- the result of "value matching + smooth pasting" may neither solve the Bellman system nor the equilibrium stopping problem;
- the Bellman system may not admit any proper solution.

Finally, we argue that it is heavy group diversity or decreasing impatience that results in these sigular phenomena, which are distinct from the time consistent cases.

Chapter 6 concludes this paper. Additional results are collected in appendices.

[6]See, for example, Friedman (2010).

Chapter 2

Weighted discount functions

In this chapter, we formally define the class of weighted discount functions and characterize our notion of group diversity. We also illustrate the importance and prevalence of weighted discounting.

2.1 General definition of weighted discount functions

The following definition extends equation (1.0.1) to distributions of opinions that may be continuous.

Definition 2.1.1 *Let* $h : [0, \infty) \to (0, 1]$ *be strictly decreasing with* $h(0) = 1$. h *is a weighted discount function if there exists a (cumulative) distribution function* $F(r)$ *concentrated on* $[0, \infty)$ *such that*

$$h(t) = \int_0^\infty e^{-rt} dF(r). \tag{2.1.1}$$

F *is called the* weighting distribution *of* h.

A weighted discount function is thus defined as a weighted average of exponential discount functions with different discount rates r.

2.2 Prevalence and examples of weighted discount functions

As an introductory example, consider the so-called pseudo-exponential discount function. It is given by $h(t) = \delta e^{-rt} + (1-\delta)e^{-(r+\lambda)t}, \lambda > 0, r > 0, 0 < \delta < 1$ (Ekeland and Lazrak 2006; Karp 2007), i.e., it is given by the weighted average of just two opinions. The corresponding weighting distribution F in equation (2.1.1) is a step function and can be written as

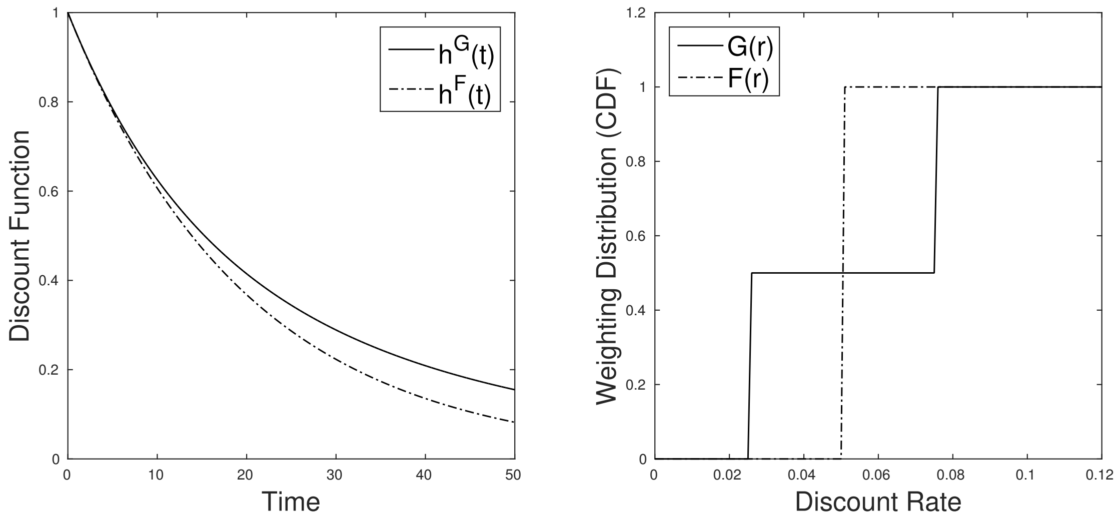
$$F(x) = \begin{cases} 0, & x < r \\ \delta, & r \leq x < r + \lambda \\ 1, & x \geq r + \lambda. \end{cases} \tag{2.2.1}$$

The corresponding probability mass function takes the jump sizes of F, which are δ and $1-\delta$, as the weights of opinion r and $r+\lambda$, respectively. Whenever $\delta = 1$, the pseudo-exponential discount function is reduced to the exponential discount function, $h(t) = e^{-rt}$. The weighting distribution of standard exponential discounting with rate r is thus degenerate, i.e., given by a step function with a single jump of size 1 at r. Figure 2.1 plots a pseudo-exponential and an exponential discount function (left panel) as well as their weighting distributions (right panel).

The following well-known theorem is attributed to Bernstein (1928), and provides a necessary and sufficient condition for a given discount function to be a weighted discount function.

Theorem 2.2.1 (Bernstein's Theorem) *A discount function h is a weighted*

Figure 2.1: Discount functions and weighting distributions



Notes. The left panel shows a pseudo-exponential discount function h^G with parameters $r = 0.025$ and $\lambda = 0.05$ as well as an exponential discount function h^F with parameter $r = 0.05$. The right panel shows the corresponding weighting distributions G and F. The left panel further illustrates that G has a uniformly larger discount factor than group F, i.e., $h^G(t) \geq h^F(t)$ for all $t \geq 0$. Namely, group G discounts less heavily at any future time t. In section 2.3 below, we show that this is a consequence of group G being more diverse than the single-member group F.

discount function if and only if it is continuous on $[0, \infty)$, *infinitely differentiable on* $(0, \infty)$, *and satisfies* $(-1)^n h^{(n)}(t) \geq 0$, *for all non-negative integers* n *and for all* $t > 0$.

Bernstein's theorem can be used to verify that the discount functions mentioned in the introduction are indeed weighted discount function. Brockett and Golden (1987) refer to the class of increasing utility functions with derivatives that alternate in sign as the class that contains "all commonly used utility functions." In analogy, we may just refer to the class of weighted discount functions as the class that contains "all commonly used discount functions."[1]

[1]Ebert (2015a) characterizes the derivatives of the discount functions through risk taking-

While every weighting distribution F defines a discount function h^F – and while most discount functions are weighted discount functions – both the weighting distribution and the weighted discount function being available in closed form is rather the exception than the rule. In particular, it may not be possible to compute the integral in Definition 2.1.1 explicitly and, likewise, for a given weighted discount function h^F, it may not be possible to recover the weighting distribution F in closed form. The results in this paper, however, do not depend on whether a closed form expression of either the discount function or the weighting distribution is available. For example, the CADI function in Bleichrodt et al. (2009) may not be obtained as the group discount function of any nice distribution function F, but we can still study its implications for investment behavior and draw inference on the diversity of the group it represents. Likewise, the discount function obtained from a Chi-Square distribution may not be available in closed-form, but nevertheless we can study the investment decisions of a group with Chi-square-distributed opinions over discount rates.

There are also some weighting distributions that lead to nice weighted discount functions (and vice versa). We already gave the example of the pseudo-exponential discount function, which corresponds to a binary weighting distribution. This important example is easily generalized to finitely many group members, whose opinions are multinomially distributed. Even though not mentioned by Weitzman (2001), the weighted discount function obtained from the Gamma distribution (with proper re-parametrization as below) coincides with that of Loewenstein and Prelec (1992). In particular, the generalized hyperbolic discount function with

behavior over time risks (the risk of something certain happening sooner or later). Since Brockett and Golden (1987) refer to increasing utility functions with derivatives that alternate in sign as *mixed risk averse,* Ebert (2015a) calls the discount functions that are decreasing with derivatives that alternate in sign – i.e., the weighted discount functions – *anti-mixed risk averse.* Ebert (2015a) does not investigate the weighting representation in Definition 2.1.1, nor does he study dynamic decisions such as stopping.

parameters $\alpha > 0, \beta > 0$ can be written as (the first expression repeats the original definition of Loewenstein and Prelec (1992))

$$h(t) = \frac{1}{(1+\alpha t)^{\frac{\beta}{\alpha}}} = \int_0^\infty e^{-rt} f(r; \frac{\beta}{\alpha}, \alpha) dr \tag{2.2.2}$$

where

$$f(r; k, \theta) = \frac{r^{k-1} e^{-\frac{r}{\theta}}}{\theta^k \Gamma(k)} \tag{2.2.3}$$

denotes the density function of the Gamma distribution with parameters k and θ and where $\Gamma(k)$ denotes the Gamma function evaluated at k, i.e., $\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx$. A simpler example is given by the discount function studied in Mazur (1987) and Harvey (1995), which corresponds to the special case of Loewenstein and Prelec (1992) where $\alpha = \beta$. In that case, equation (2.2.2) becomes

$$h(t) = \frac{1}{1+\alpha t} = \int_0^\infty e^{-rt} \left(\frac{1}{\alpha} e^{-\frac{1}{\alpha} r} \right) dr, \tag{2.2.4}$$

which shows that the weighting distribution of this hyperbolic discount function is given by the familiar exponential distribution with mean α.

As a final example, we propose a new discount function, the *uniform-uncertainty discount function*, that is obtained from a uniform distribution of opinions over an interval $[\underline{r}, \overline{r}]$:

$$h(t) = \begin{cases} 1, \text{ if } t = 0 \\ \frac{e^{-\underline{r}t} - e^{-\overline{r}t}}{t(\overline{r} - \underline{r})}, \text{ if } t > 0 \end{cases} = \int_{\underline{r}}^{\overline{r}} e^{-rt} \frac{1}{\overline{r} - \underline{r}} dr.$$

Before we conclude this section, an economically important observation results from the mathematically simple fact that any weighted average of weighted discount functions is also a weighted discount function. This implies that our results are not restricted to groups of exponential discounters, each of whom is certain about what discount rate to use. Instead, we may consider groups with some or all members being hyperbolic or CADI, because the latter – as just shown – are also weighted discount functions. Likewise, we may consider groups of individual discounters, each of whom is uncertain about the discount rate to use. By iterating the argument it follows that the group discount function of individuals who are (i) weighted discounters and, in addition, (ii) uncertain about which parameters to use, is (still) a weighted discount function. We formalize these arguments and give an example in appendix A.

2.3 Greater group diversity

In this section, we establish a notion of comparative group diversity in time preferences. The main result is that a group has a uniformly greater discount factor than another group (i.e., discounts less heavily at any future time) if and only if it is more diverse. Greater group diversity is captured by a more dispersed weighting distribution F of the corresponding weighted discount function h^F, and will be defined formally using stochastic dominance. To avoid technical difficulties, for our stochastic dominance results we assume that all weighting distributions have their support contained in $[a, b], 0 \leq a \leq b < \infty$. Let $F^0(x) := F(x)$ and define $F^i(x) = \int_a^x F^{i-1}(t)dt$ for $i \geq 1$. For the definition of stochastic dominance of arbitrary order and the subsequent, well-known expected utility characterization theorem, see, e.g., Ingersoll (1987).

Definition 2.3.1 *The distribution F dominates the distribution G in the sense of Nth-order* stochastic dominance*, denoted by $G \preceq_{NSD} F$, if*

(i) $F^{N-1}(x) \leq G^{N-1}(x)$ for all $a \leq x \leq b$,

(ii) $F^i(b) \leq G^i(b)$ *for* $i = 1, 2, \ldots, N-2$.

Theorem 2.3.2 *The following are equivalent:*

(i) $G \preceq_{NSD} F$,

(ii) $\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x)$, *for all function* u *such that sgn* $u^{(n)}(x) = (-1)^{n+1}$ *for* $n = 1, 2, \ldots, N$ *and* $x \in [a, b]$.

Theorem 2.3.2 characterizes stochastic dominance in terms of expected utility preferences. A second-order stochastically dominating distribution is preferred, for example, by all decision makers whose utility functions are increasing and concave. It can be shown that a second-order stochastically dominating distribution must have a higher mean and a lower variance than the distribution that it dominates, with at least one of the two inequalities being strict. From Definition 2.3.1, it is easy to see that $G \preceq_{NSD} F \Rightarrow G \preceq_{(N+1)SD} F$. By Theorem 2.3.2, we can define the "weakest" stochastic dominance order as follows.

Definition 2.3.3 (Infinity-stochastic dominance.) F dominates G via infinity-stochastic dominance *(denoted as* $G \preceq_{\infty SD} F$*), if* $\int_a^b u(x)dF(x) \geq \int_a^b u(x)dG(x)$, *for all function* u *such that sgn* $u^{(n)}(x) = (-1)^{n+1}$ *for any* $n \in \mathbb{N}^+, x \in [a, b]$.

We are now in the position to introduce our notion of greater group diversity.

Definition 2.3.4 (Greater group diversity.) *Let* F *and* G *be weighting distributions of weighted discount functions* h^F *and* h^G*, respectively. We say that* h^G *(*G*) is* more diverse *than* h^F *(*F*) if* $G \preceq_{\infty SD} F$.

At the first sight, the reader may feel that the "direction" of $\preceq_{\infty SD}$ should just be the other way around. The notation we chose is in line with the one that is standard in the stochastic dominance literature, where $G \preceq_{\infty SD} F$ means that F is *preferred over* G because F has less risk (dispersion) in the stochastic dominance sense. We reinterpret this by saying that $G \preceq_{\infty SD} F$ means that group G is *more diverse* than group F.

From the definition of ∞-stochastic dominance it follows directly that if F dominates G via stochastic dominance of some order N, then F dominates G by ∞-stochastic dominance, i.e., G is more diverse than F. In this sense, the notion of greater group diversity is weak. In particular, when we later prove a consequence of greater group diversity, then this consequence will also follow from, say, diversity in the sense of second-order stochastic dominance. Most importantly, though, our weak notion of greater group diversity allows for the following characterization.

Theorem 2.3.5 (group diversity and larger discount factors.) *Suppose F and G are weighting distributions with finite support and the corresponding discount functions are given by h^G and h^F, respectively. Then:*

$$G \preceq_{\infty SD} F \text{ if and only if } h^G(t) \geq h^F(t), \forall t \geq 0. \tag{2.3.1}$$

Proof. Suppose first that $G \preceq_{\infty SD} F$. Note that $u(x) := -e^{-xt}$ defines a mixed risk averse utility function (i.e., is increasing with derivatives that alternate in sign) on $[0, \infty)$ for each fixed $t > 0$. Therefore, Definition 2.3.3 yields that $h^F(t) \leq h^G(t)$, $\forall t \geq 0$.

Now suppose that $h^F(t) \leq h^G(t)$ for all $t \geq 0$. For any mixed risk averse utility function[2] u, the first order derivative of u satisfies the assumptions of Bernstein's Theorem (Theorem 2.2.1). Thus, there exists a distribution function F_u such that

[2]We say a utility function u shows mixed risk averse if $sgn(u^{(n)}) = (-1)^n, \forall n = 1, 2, 3, \ldots$.

$u'(t) = \int_0^\infty e^{-ts} dF_u(s)$. It follows that

$$\begin{aligned}
u(x) &= \int_a^x u'(t)dt + u(a) \\
&= \int_a^x \int_0^\infty e^{-ts} dF_u(s)dt + u(a) \\
&= \int_0^\infty \int_a^x e^{-ts} dt dF_u(s) + u(a) \\
&= \int_0^\infty \frac{1}{s}(e^{-as} - e^{-xs}) dF_u(s) + u(a).
\end{aligned}$$

Hence,

$$\begin{aligned}
\int_a^b u(x)dF(x) &= \int_a^b \int_0^\infty \frac{1}{s}(e^{-as} - e^{-xs}) dF_u(s)dF(x) + u(a) \\
&= \int_0^\infty \frac{1}{s}(e^{-as} - \int_a^b e^{-xs} dF(x)) dF_u(s) + u(a) \\
&= \int_0^\infty \frac{1}{s}(e^{-as} - h^F(s)) dF_u(s) + u(a).
\end{aligned}$$

Comparing with the analogous expression for G, from Theorem 2.3.2 and because $h^F(t) \leq h^G(t)$ for all $t \geq 0$ it follows that $G \preceq_{\infty SD} F$.

This completes the proof. ■

The latter property of a *uniformly larger discount factor* is central to many results of this thesis. A larger discount factor implies less discounting, and "uniform" means that this is the case at *any* future time. Graphically, it means that the discount function of a more diverse group always lies above that of a less diverse group. Figure 2.1 illustrates this result through the fact that the pseudo-exponential discount function h^G lies above the exponential discount function h^F at all times. Indeed, in that example, G is more diverse than F in the sense of second-order stochastic dominance. In particular, the binary distribution G is a mean-preserving spread (Rothschild and Stiglitz 1970) of the degenerate distribution F. The intuition behind the proof for this special case is that exponential functions are convex so that, by Jensen's inequality, averaging two discount factors whose discount *rates* average to 5%, results in a discount factor larger than a single exponential discount factor with rate 5%. For higher-orders of stochastic dominance, the result follows from Bernstein's theorem and the stochastic dominance theorem, Theorem 2.3.2. The result that larger discount factors also necessitate greater group diversity is more complex and proven in Theorem 2.3.5.

Consider Theorem 2.3.5 for the special case where F and G are degenerate so that h^F and h^G are exponential discount functions. The equivalence (2.3.1) then says that group G has a smaller discount rate if and only if it has a larger discount factor. In other words, in the exponential discounting case, comparing discount factors is equivalent to comparing discount rates. For general discount functions, however, comparing discount factors is *not* equivalent to comparing expected discount rates. For weighted discount functions, Theorem 2.3.5 establishes the necessary and sufficient condition for ordering discount factors. Ordering discount factors is equivalent to comparing the weighting distributions in the infinity-stochastic dominance sense. This further motivates the study of group diversity in general and its definition in terms of infinity-stochastic dominance in particular.

Theorem 2.3.5 makes the assumption of finite support as is common in the stochastic dominance literature. While it is merely of a technical nature, the weighting distributions of several important discount functions do have infinite support (e.g., hyperbolic discount functions). For that reason, we put the property of larger discount factors (rather than greater group diversity) in the center of our mathematical analysis. Greater group diversity should be kept in mind, though, whenever we speak of larger discount factors.

2.4 Decreasing impatience and other properties of weighed discount functions

In this section, we relate greater group diversity (equivalently, the property of a uniformly larger discount function) to important and well-known properties of discount functions.

Definition 2.4.1 (Decreasing impatience (Prelec 2004).) *A discount function h satisfies* decreasing impatience (DI) *if Prelec's (2004)* measure of decreasing impatience

$$P(t) = -\frac{(\ln h(t))''}{(\ln h(t))'}$$

is non-negative.

The following result is similar to Proposition 1 in Jackson and Yariv (2015), who show that the discount functions of finite groups are present-biased.

Proposition 2.4.2 (Weighted discount functions imply decreasing impatience.) *All weighted discount functions imply decreasing impatience, provided that the Prelec's measure P exists*[3].

Proof. We have to show that $P(t) = -\frac{(\ln h(t))''}{(\ln h(t))'} \leq 0$. First, note that for a weighted discount function $h(t) = \int_0^\infty e^{-rt} dF(r)$, we have

$$h'(t) = -\int_0^\infty re^{-rt} dF(r)$$

so that

$$(\ln h(t))' = \frac{h'(t)}{h(t)} < 0.$$

[3] It is easy to see that for a weighted discount function h, $P(t), t > 0$ exists if and only if $h(t) \not\equiv 1$.

Therefore, it remains to show that $-(\ln h(t))'' = \left(-\frac{h'(t)}{h(t)}\right)'$ is non-positive. Let ξ denote a random variable with distribution function $F_\xi(x) = \frac{\int_0^x e^{-rt}dF(r)}{\int_0^\infty e^{-rt}dF(r)}$. Then,

$$\begin{aligned}
\left(-\frac{h'(t)}{h(t)}\right)' &= \left(\frac{\int_0^\infty re^{-rt}dF(r)}{\int_0^\infty e^{-rt}dF(r)}\right)' \\
&= -\frac{\int_0^\infty r^2 e^{-rt}dF(r)}{\int_0^\infty e^{-rt}dF(r)} + \left(\frac{\int_0^\infty re^{-rt}dF(r)}{\int_0^\infty e^{-rt}dF(r)}\right)^2 \\
&= -\mathbb{E}[\xi^2] + \mathbb{E}[\xi]^2 = -\mathbf{Var}(\xi) \leq 0,
\end{aligned}$$

which concludes the proof. ■

Proposition 2.4.2 illustrates the main difficulty that awaits us in Chapter 3 when studying the investment decisions of groups. Due to decreasing impatience, the group is in general time-inconsistent. The next result is on weighted discount functions that differ in their degree of decreasing impatience.

Proposition 2.4.3 (Comparative DI and larger discount factors.) *Consider weighted discount functions h^F and h^G with weighting distributions F and G and measures of decreasing impatience P_F and P_G, respectively. Let the expectation of F be no less than that of G. Then,*

$$P_G(t) \geq P_F(t) \text{ for all } t \geq 0 \Longrightarrow h^G(t) \geq h^F(t) \text{ for all } t \geq 0. \tag{2.4.1}$$

Proof. Define auxiliary functions $g_i = -(\ln h^i)'$, $i = F, G$. By the assumption that group F has a higher expectation about the appropriate discount rate, $(h^F)'(0) = -\int_0^\infty rdF(r) \leq -\int_0^\infty rdG(r) = (h^G)'(0)$, so that

$$g_F(0) = -\frac{(h^F)'(0)}{h^F(0)} \geq -\frac{(h^G)'(0)}{h^G(0)} = g_G(0). \tag{2.4.2}$$

Since $P_i = -\frac{g_i'}{g_i} = -(\ln g_i)'$, $i = F, G$, the assumption that F exhibits less DI than

G, i.e., $P_F(t) \leq P_G(t) \forall t \geq 0$, can be restated as

$$(\ln g_F(t))' \geq (\ln g_G(t))'. \tag{2.4.3}$$

Therefore, (2.4.2) and (2.4.3) together imply that $\ln g_F(t) \geq \ln g_G(t), \forall t \geq 0$, which is equivalent to

$$g_F(t) \geq g_G(t), \ \forall t \geq 0, \tag{2.4.4}$$

Since $g_i = -(\ln h^i)'$ and $h^i(0) = 1, i = F, G$, we have $\ln h^F(t) \leq \ln h^G(t)$, from which the result follows. ■

The assumption of F having an equal or larger expectation means that the weighted average of group F's discount *rates* (rather than the group's discount factors) is equal or larger for F. Therefore, group F on average discounts with a larger discount rate. The link to the discount function is as follows. It is easily verified that the negative of the expectation of F can be computed as $(h^F)'(0)$. Therefore, an equal or larger expectation means that h^F initially decreases more strongly. If, in addition, h^F is less convex in the sense that its impatience is less decreasing, then h^F must always lie below h^G. From Theorem 2.3.5, it follows that h^G having more decreasing impatience and G having less or equal expectation implies greater group diversity.

Finally, we show that greater patience, as defined by Quah and Strulovici (2013), also implies greater group diversity.[4]

Definition 2.4.4 (Comparative patience (Quah and Strulovici 2013).) *The discount function h^G exhibits more patience than another discount function*

[4]Definition 2.4.4 and the subsequent Proposition 2.4.5 also apply to discount functions that are not necessarily weighted discount functions.

h^F (denoted by $h^G \succeq_P h^F$) if

$$\frac{h^G(t)}{h^F(t)} \text{ is increasing in } t.$$

Proposition 2.4.5 (More patient groups have larger discount factors.) *Consider weighting distributions F and G and denote the corresponding weighted discount functions with h^F and h^G, respectively. Then we have*

$$h^G \succeq_P h^F \Rightarrow h^G(t) \geq h^F(t) \text{ for all } t \geq 0.$$

Proof. Since $h^F \preceq_P h^G$ implies that $\frac{h^G}{h^F}$ is increasing, the fact that $\frac{h^G(0)}{h^F(0)} = 1$ yields the claim. ■

2.5 Weighted discounting and diminishing marginal impatience

In this section, we illustrate why a single agent's time preference exibits weighted discounting. To describe the convexity of a discount function we introduce the "diminishing marginal impatience" and then extend this concept to higher order derivatives. Our discussion is based on the follwoing question.

You will receive two amounts of money: \$1000 two days later and \$1000 two years later. You now have to postpone one of them for 100 days. Which one do you incline to delay?

2.5.1 Decreasing discount functions

Most of existing discount functions are decreasing. This feature shows that delay is undesirable when an agent faces a positive future payoff.

Definition 2.5.1 *An agent shows a regular time preference if and only if her discount function h is decreasing, i.e.,*

$$h(t_1) \geq h(t_2), \tag{2.5.1}$$

whenever $t_1 \leq t_2$.

Particularly, for a smooth function h, it is equivalent to $h' \leq 0$.

2.5.2 Diminishing marginal impatience

Let us go back to the question proposed at the beginning of this section. We here put it in a more general environment.

You will receive two amounts of money: x dollars (\$1000) l (two) days later and x dollars (\$1000) $l + l_1$ (two) years later. You now have to postpone one of them for l_2 (100) days. Which one do you incline to delay?

If the agent's preference is measured by the present value, then she will add the extra delay on the second one if and only if

$$h(l + l_2) + h(l + l_1) \leq h(l) + h(l + l_1 + l_2). \tag{2.5.2}$$

Rearrange (2.5.2), then we have

$$h(l + l_2) - h(l + l_1 + l_2) \leq h(l) - h(l + l_1).$$

Suppose that the agent has a regular time preference, then the above inequality shows that the dissatisfaction that the extra delay brings decreases with the current waiting time. This feature is quite similar to the law of diminishing marginal utility, which is described by the concavity of the utility function. Motivated by this, we introduce the following definition.

Definition 2.5.2 *We say an agent's time preference follows the law of diminishing marginal impatience, if and only if her discount function h is convex.*

Particularly, for a smooth function h, it is equivalent to $h'' \geq 0$.

An oft-studied theme in the research of discounting is decreasing impatience which is described by the decrease of the discount rate $-\frac{h'}{h}$. Notice that $\frac{h'}{h}$ can be written as $(\log h)'$. If an agent's time preference is modelled by $\log h$, then the decreasing impatience is equivalent to diminishing marginal impatience.

The following proposition and example illustrate the relationship between diminishing marginal impatience and decreasing impatience.

Proposition 2.5.3 (DI implies diminishing marginal impatience.) *A discount function follows the law of diminishing marginal impatience if it implies decreasing impatience.*

Proof. Suppose the discount function is h, then the result follows the fact that $(-\frac{h'}{h})' \leq 0$ implies $h'' \geq 0$. ■

However, it is not difficult to find an example showing that the opposite direction is not true.

Example 2.5.4 *Suppose the discount function $h(t) = e^{-t^2-3t}$, then h shows the diminishing marginal impatience but not the decreasing impatience.*

2.5.3 Extension to higher order derivatives

From the monotonic decrease of a discount function (Definition 2.5.1), we know that delay is "bad". Hence, from (2.5.2), we observe that an agent whose time preference follows the law of diminishing marginal impatience tends to postpone the bad thing. Based on this observation, we will extend the law of diminishing marginal impatience and discuss the signs of higher order derivatives. Before that, for sake of convenience, we introduce some notations.

Let $\mathcal{A}$ be a family of finite sets whose elements are non-negative real numbers. We intend to let any $A \in \mathcal{A}$ define a stream of waiting time. Hence we permit that the elements taking the identical value can appear repeatedly in a set. For example, we do not view $\{a, a\}$ as a singleton but a set with two elements.

For any set $A \in \mathcal{A}$ and a non-negative real number d, we define the "delay" by the operator $\oplus$ as follows.

$$A \oplus d \equiv \{l + d\}_{l \in A}$$

For any $A, B \in \mathcal{A}$, $A \sqcup B$ means two steams of payoffs that will be received. One is at the time in A, the other at the time in B, i.e.,

$$A \sqcup B = \{\{l\}_{l \in A}, \{l\}_{l \in B}\}.$$

For a given discount function h, let m^h be a function mapping from $\mathcal{A}$ to $\mathbb{R}^+ \cup \{0\}$,

$$m^h(A) := \sum_{l \in A} h(l),$$

From the definition of m^h, it is to see that for any $x > 0, A \in \mathcal{A}$, $xm^h(A)$ describes the present value of a stream of payoffs valued at x at each time $l \in A$. In light of this, we define the preference order as follows.

Definition 2.5.5 *We say $A \in \mathcal{A}$ is preferred to $B \in \mathcal{A}$ (denoted by $A \succeq B$) under the discount function h if and only if $m^h(A) \geq m^h(B)$.*

Define, for all $d > 0, l \geq 0$, $\Delta_d h(l) := h(l + d) - h(l)$ and by induction, $\Delta_d^0 h(l) := h(l), \Delta_d^1 h(l) := \Delta_d h(l), \Delta_d^n h(l) := \Delta_d \Delta_d^{n-1} h(l)$.

In order to explain higher order derivatives of a discount function, we now introduce a sequence of time sets.

Consider $A_1 = \{l\}$, $B_1 = \{l + d\}$, where $l, d \in \mathbb{R}^+ \cup \{0\}$. Since $m^h(A_1) - m^h(B_1) = h(l) - h(l + d) = -\Delta_d^1 h(l)$, then we have that a discount function h

exhibits a regular time preference if and only if $A_1 \succeq B_1$.

Consider $A_2 = A_1 \sqcup (B_1 \oplus d), B_2 = (A_1 \oplus d) \sqcup B_1$. By simple algebra, we have $m^h(A_2) - m^h(B_2) = h(l+2d) - 2h(l+d) + h(l) = \Delta_d^2 h(l)$. Then a discount function h follows the law of diminishing marginal impatience if and only if $A_2 \succeq B_2$.

As stated at the beginning of this part, the construction of A_2 (B_2) exhibits the bahavior of an agent who delays "the bad (good) thing". Next we continue to construct A_3 and B_3 based on this rule. For an agent following the law of diminishing marginal impatience, A_2 is good and B_2 is bad, hence we construct A_3 and B_3 as follows.

Consider $A_3 = A_2 \sqcup (B_2 \oplus d), B_3 = (A_2 \oplus d) \sqcup B_2$. It is to see that $m^h(A_3) - m^h(B_3) = -(h(l+3d) - 3h(l+2d) + 3h(l+d) - h(l)) = -\Delta_d^3 h(l)$. If an agent tends to postpone the bad thing rather than the good one, then she would prefer A_3 to B_3. i.e., $m^h(A_3) - m^h(B_3) = -\Delta_d^3 h(l) \geq 0$. For a smooth function h, this is equivalent to $h''' \leq 0$.

By induction, for any $l \geq 0, d > 0, n > 1$, we define $A_n = A_{n-1} \sqcup (B_{n-1} \oplus d), B_n = (A_{n-1} \oplus d) \sqcup B_{n-1}$. The following lemma shows the relationship between A_n, B_n and the nth order difference.

Lemma 2.5.6

$$m^h(A_n) - m^h(B_n) = (-1)^n \Delta_d^n h(l),$$

for any $n \geq 1, l \geq 0, d > 0$.

Hence, $A_n \succeq B_n$ *under* h *if and only if* $(-1)^n \Delta_d^n h(l) \geq 0$.

Proof. We prove this lemma by induction.

By the discussion above this lemma, Lemma 2.5.6 holds for $n = 1$.

Suppose that when $n = k$, $m^h(A_k) - m^h(B_k) = (-1)^k \Delta_d^k h(l)$.

When $n = k + 1$,

$$m^h(A_{k+1}) = m^h(A_k \sqcup (B_k \oplus d)) = m^h(A_k) + m^h(B_k \oplus d),$$

$$m^h(B_{k+1}) = m^h(B_k \sqcup (A_k \oplus d)) = m^h(B_k) + m^h(A_k \oplus d)).$$

Hence,

$$\begin{aligned} m^h(A_{k+1}) - m^h(B_{k+1}) &= m^h(A_k) - m^h(B_k) - (m^h(A_k \oplus d) - m^h(B_k \oplus d)) \\ &= (-1)^k \Delta_d^k h(l) - (-1)^k \Delta_d^k h(l+d) \\ &= (-1)^k \Delta_d^k (h(l) - h(l+d)) \\ &= (-1)^k \Delta_d^k (-\Delta_d h(l)) \\ &= (-1)^{k+1} \Delta_d^{k+1} h(l). \end{aligned}$$

This completes the proof. ■

The following theorem tells that an agent's time preference can be described by a weighted discount function if and only if she has a regular time preference and prefers delaying the bad thing to the good one towards the time sets A_n, B_n, $n \geq 2$.

Theorem 2.5.7 *A discount function h is a weighted discount function if and only if $A_n \succeq B_n$ under h, for any $n \geq 1, l \geq 0, d > 0$.*

Proof. This is an immediate result of Lemma 2.5.6, Theorem 2.2.1 and Caballé and Pomansky (1996), Corollary 2.3. ■

While we believe that our results on weighted discount functions are interesting in their own right, we will illustrate their usefulness explicitly when studying the comparative statics of investment behavior in Section 4.2 with respect to group properties such as patience or diversity.

Chapter 3

General framework

In this chapter, we focus on the game theoretic framework for a stopping problem with weighted discounting. First, to obtain a formal definition of an equilibrium stopping rule, we embed our stopping problem into a control framework by formulating the stopping rule as a binary control. Next, within the game theoretic framework for time inconsistent control problems (e.g., Bjork and Murgoci (2010)), we give a definition of an equilibrium stopping rule. Then, following Strotz (1955), Caplin and Leathy (2006) and many other since, we derive a Bellman system by a recursive approach for a finite horizon case and illustrate why the weighted representation of a discount function plays a key role in taming continuously changing time preferences for a stopping problem. Finally, we turn to the infinite horizon case, in which we explain the Bellman system in terms of a decision made by a group. The main result in this chapter is that the stopping rule obtained by solving the Bellman system is indeed the equilibrium stopping rule defined in the game theoretic framework.

3.1 The dynamics

Consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which supports a standard Brownian motion $(W_t)_{t\geq 0}$ with its natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$ after augmentation.

Suppose the underlying process $X = \{X_t\}_{t\in[0,T]}$ follows

$$\frac{dX_t}{X_t} = b(X_t)dt + \sigma(X_t)dW_t. \tag{3.1.1}$$

where the mean return b and volatility $\sigma(> 0)$ are bounded and continuous functions. X may describe the wealth of an agent gambling in casino, the quality of a current job offer, the project value in a real options setting, or the value of a stock or of the underlying of some derivative.

To guuarantee that the differential equation (3.1.1) admits a unique strong solution, we assume that the functions $xb(x)$ and $x\sigma(x)$ satisfy the global Lipschitz condition. Moreover, by the standard discussion in stochastic differential equations (e.g. Karatzas and Shreve 2006, p.306), for $X_0 = x$ and any $m \geq 1$, we have the following estimate,

$$\mathbb{E}[\max_{0\leq s\leq t} |X_s|^m] \leq C(1 + |x|^m)e^{Ct}, \quad 0 \leq t \leq T, \tag{3.1.2}$$

where C is a positive constant depending on m, T and the bound of b and σ.

3.2 Stopping rules

In time consistent stopping problems, a stopping rule is often formulated as a stopping time by which an agent can decide when to stop. This stopping time tells the agent whether to stop not only at the current decision date, but in the following periods as well.

However, in time inconsistent cases, since the preferences change over time, the agent is only able to commit her current behaviour under her current preference. This implies that the agent is only capable of making the decision of whether to stop instead of when to stop. Therefore, it seems more natural to introduce a definition, which is based on whether to stop, to describe a stopping rule.

Definition 3.2.1 *A* stopping rule *is a measureable function of time and the process value* $u : ([0, T] \times \mathbb{R}^+, \mathcal{B}[0, T] \times \mathcal{B}(\mathbb{R}^+)) \to (\{0, 1\}, 2^{\{0,1\}})$ *where* 0 *indicates "continue" and* 1 *indicates "stop". Each stopping rule* u *defines a sequence of* stopping times $\tau_u^t, t \in [0, T]$ *via*

$$\tau_u^t = \inf\{s \geq t, u(s, X_s) = 1\}. \tag{3.2.1}$$

The so called "stopping rule" is reflected by the definition of stopped process $\{X_{s\wedge\tau_u^t}\}_{s\in[t,T]}$, and the superscript t means that the agent will use the stpping rule u under the assumption that she has never stopped before the calendar time t.

For sake of convenience, similar to the time consistent case, we introduce the concept of "stopping region" for a stopping rule u

$$\mathcal{S}_u = \{(t, x) \in [0, T] \times \mathbb{R}^+ : u(t, x) = 1\}.$$

The complement set $\mathcal{C}_u$ of $\mathcal{S}_u$ is the continuation region

$$\mathcal{C}_u = \{(t, x) \in [0, T] \times \mathbb{R}^+ : u(t, x) = 0\}.$$

Then the stopping rule u can be shown as follows.

$$u = \begin{cases} 1 & \text{if } (t, x) \in \mathcal{S}_u, \\ 0 & \text{if } (t, x) \in \mathcal{C}_u. \end{cases} \tag{3.2.2}$$

The corresponding stopping time at time t is a hitting time given by

$$\tau_u^t = \inf\{s \geq t : (s, X_s) \in \mathcal{S}_u\}. \tag{3.2.3}$$

3.3 Game theoretic formulation

Let the continuous functions c and G from $\mathbb{R}^+$ into $\mathbb{R}$ denote the running profit and the exercise payoff respectively. At any point $(t, x) \in [0, T] \times \mathbb{R}^+$, we consider the following performance functional

$$J(t, x; u) = \mathbb{E}[\int_t^{T\wedge\tau_u^t} h(s-t)c(X_s)ds + h(T \wedge \tau_u^t - t)G(X_{T\wedge\tau_u^t})|X_t = x], \tag{3.3.1}$$

In the rest of the paper, we suppose that the running profit c and the exercise payoff G have the polynomial growth, i.e., there exist constants $C > 0, m \geq 1$, such that $|c(x)| + |G(x)| \leq C(1 + |x|^m)$.

In time consistent cases in which $h(t) = e^{-rt}$ with $r > 0$, the agent would like to search for an optimal stopping rule, which can be formulated as the following optimization problem,

$$\max_u J(t, x; u). \tag{3.3.2}$$

However, for a general discount function h, the time consistency of the optimal stopping rule may be lost, since the dynamic programming principle does not apply any more. This implies that the optimal stopping rule is changing as time goes by. Hence, the concept "optimal stopping rule" may not make a lot of economic sense, since the agent is looking for a strategy that she will not really follow in the following periods.

Given the above argument, instead of the "optimal stopping rule", we are going to seek an "equilibrium stopping rule" by the game theoretic approach[1]. For our stopping problem, given a stopping rule u and a self $t \geq 0$, we can define the following stopping rule to be used from t on:

$$u^{t,\epsilon,a}(s,x) = \begin{cases} u(s,x) & \text{if } s \in [t+\epsilon, T], \\ a & \text{if } s \in [t, t+\epsilon), \\ 0 & \text{otherwise.} \end{cases} \tag{3.3.3}$$

where $\epsilon > 0$ and $a \in \{0,1\}$ are fixed. This stopping rule $u^{t,\epsilon,a}$ coincides with the self t's original stopping rule u except for the (short) time interval $[t, t+\epsilon)$. On that interval, $u^{t,\epsilon,a}$ is either constantly 0 or constantly 1. We call $u^{t,\epsilon,a}$ the (ϵ, a)-*deviation* from u at time t.

Definition 3.3.1 (Equilibrium stopping rule) *Let $\hat{u}$ be a given stopping rule. Then we say $\hat{u}$ is an equilibrium stopping rule if for any $(t,x) \in [0,T] \times \mathbb{R}^+, a \in \{0,1\}$, we have*

$$\liminf_{\epsilon \to 0} \frac{J(t,x;\hat{u}) - J(t,x;u^{t,\epsilon,a})}{\epsilon} \geq 0, \tag{3.3.4}$$

where $u^{t,\epsilon,a}$ is the (ϵ, a) deviation from $\hat{u}$ at time t, i.e.,

$$u^{t,\epsilon,a}(s,x) = \begin{cases} \hat{u}(s,x) & \textit{if } s \in [t+\epsilon, T], \\ a & \textit{if } s \in [t, t+\epsilon), \\ 0 & \textit{otherwise.} \end{cases} \tag{3.3.5}$$

with $a \in \{0,1\}, \epsilon > 0$.

[1] This framework has been already used to model many time inconsistent control problems; see Ekeland and Pirvu (2008); Björk et al. (2014).

3.4 Bellman system

To obtain a Bellman system for the stopping problem defined by Definition 3.3.1, similar to the time consistent case, we introduce the definition of a value function.

Definition 3.4.1

$$V(t,x) := J(t,x;\hat{u}), \tag{3.4.1}$$

where $\hat{u}$ is the equilibrium stopping rule defined by Definition 3.3.1.

Now we give a heuristic derivation of the Bellman system which is comprised of a terminal condition, a differential equation and an auxiliary function. Let us start with the terminal condition.

- At terminal time T, suppose the underlying process $X_T = x$, then the agent obtains the payoff $\max\{G(x), 0\}$ and the equilibrium stopping rule $\hat{u}$ is given by

$$\hat{u}_T = \begin{cases} 0 & \text{if } G(x) < 0, \\ 1 & \text{otherwise.} \end{cases}$$

.

Next, we derive the differential equations that the value function $V(t,x)$ follows. At the point $(t,x) \in [0,T) \times \mathbb{R}^+$, the agent has two options: stop or continue.

- If the agent stops at t, i.e., she uses the stopping rule $u^{t,dt,1}$,

  then the corresponding performance functional $J(t,x;u^{t,dt,1})$ defined by (3.3.1) is

$$J(t,x;u^{t,dt,1}) = G(x).$$

- If the agent decides to continue on the infinitesimal interval $[t, t+dt)$, since her stopping rule has been set up since $t+dt$, what she actually follows at time t is the stopping rule $u^{t,dt,0}$. Then her performance functional in this case is $J(t,x;u^{t,dt,0})$.

$$\begin{aligned}
J(t,x;u^{t,dt,0}) &= \mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^{t+dt}} h(s-t)c(X_s)ds + h(T\wedge\tau_{\hat{u}}^{t+dt}-t)G(X_{T\wedge\tau_{\hat{u}}^{t+dt}})|X_t=x] \\
&= \mathbb{E}[\int_t^{t+dt} h(s-t)c(X_s)|X_t=x] \\
&+ \mathbb{E}[\int_{t+dt}^{T\wedge\tau_{\hat{u}}^{t+dt}} h(s-(t+dt))c(X_s)ds|X_t=x] \\
&+ \mathbb{E}[\int_{t+dt}^{T\wedge\tau_{\hat{u}}^{t+dt}} (h(s-t)-h(s-(t+dt)))c(X_s)ds|X_t=x] \\
&+ \mathbb{E}[h(T\wedge\tau_{\hat{u}}^{t+dt}-(t+dt))G(X_{T\wedge\tau_{\hat{u}}^{t+dt}})|X_t=x] \\
&+ \mathbb{E}[(h(T\wedge\tau_{\hat{u}}^{t+dt}-t)-h(T\wedge\tau_{\hat{u}}^{t+dt}-(t+dt)))G(X_{T\wedge\tau_{\hat{u}}^{t+dt}})|X_t=x] \\
&= \mathbb{E}[\int_t^{t+dt} h(s-t)c(X_s)|X_t=x] \\
&+ \mathbb{E}[\int_{t+dt}^{T\wedge\tau_{\hat{u}}^{t+dt}} (h(s-t)-h(s-(t+dt)))c(X_s)ds|X_t=x] \\
&+ \mathbb{E}[(h(T\wedge\tau_{\hat{u}}^{t+dt}-t)-h(T\wedge\tau_{\hat{u}}^{t+dt}-(t+dt)))G(X_{T\wedge\tau_{\hat{u}}^{t+dt}})|X_t=x] \\
&+ \mathbb{E}[V(t+dt,X_{t+dt})|X_t=x].
\end{aligned}$$

Therefore,

$$\begin{aligned}
V(t,x) = \max\{&\mathbb{E}[\int_t^{t+dt} h(s-t)c(X_s)|X_t=x] \\
&+ \mathbb{E}[Dh(T\wedge\tau_{\hat{u}}^{t+dt}-(t+dt))G(X_{T\wedge\tau_{\hat{u}}^{t+dt}})|X_t=x] \\
&+ \mathbb{E}[\int_{t+dt}^{T\wedge\tau_{\hat{u}}^{t+dt}} (Dh(s-(t+dt)))c(X_s)ds|X_t=x] \\
&+ \mathbb{E}[V(t+dt,X_{t+dt})|X_t=x], G(x)\}.
\end{aligned}$$

where $Dh(t) = h(t+dt)-h(t), \forall t\in[0,T)$.

Moving $V(t,x)$ to the right-hand side of the above equality, and dividing it by dt, we have

$$
\begin{aligned}
0 = \max\{ & \frac{\mathbb{E}[\int_t^{t+dt} h(s-t)c(X_s)|X_t = x]}{dt} \\
+ & \frac{\mathbb{E}[Dh(T \wedge \tau_{\hat{u}}^{t+dt} - (t+dt))G(X_{T\wedge\tau_{\hat{u}}^{t+dt}})|X_t = x]}{dt} \\
+ & \frac{\mathbb{E}[\int_{t+dt}^{T\wedge\tau_{\hat{u}}^{t+dt}} (Dh(s-(t+dt)))c(X_s)ds|X_t = x]}{dt} \\
+ & \frac{\mathbb{E}[V(t+dt, X_{t+dt})|X_t = x] - V(t,x)}{dt}, G(x) - V(t,x)\}
\end{aligned}
$$

Here, we use the fact that

$$
\max\{ad, b\} = 0 \Leftrightarrow \max\{a, b\} = 0, \forall d > 0.
$$

Note that for the weighted discount function $h(t) = \int_0^\infty e^{-rt}dF(r)$, the first order derivative $h'(t) = -\int_0^\infty re^{-rt}dF(r)$, then we obtain the Bellman system as follows.

Definition 3.4.2 *Given the performance functional (3.3.1), the Bellman equation for the value function V in $[0,T) \times \mathbb{R}^+$ is given by*

$$
\max\{(\mathcal{M}V)_t(x) + c(x) - \int_0^\infty rw(t,x;r)dF(r), G(x) - V(t,x)\} = 0, \qquad (3.4.2)
$$

with the terminal condition

$$
V(T,x) = \max\{G(x), 0\}, \qquad (3.4.3)
$$

for $\forall x \in \mathbb{R}^+$.

Here, $\mathcal{M}$ is the infinitesimal generator for the underlying process X given by

$$
\mathcal{M} = \frac{\partial}{\partial t} + \frac{1}{2}\sigma(x)^2x^2\frac{\partial^2}{\partial x^2} + b(x)x\frac{\partial}{\partial x},
$$

and the auxiliary function w is defined as

$$w(t,x;r) = \mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t} e^{-r(s-t)}c(X_s)ds + e^{-r(T\wedge\tau_{\hat{u}}^t - t)}G(X_{T\wedge\tau_{\hat{u}}^t})|X_t = x]. \tag{3.4.4}$$

The equilibrium stopping rule $\hat{u}$ is given by

$$\hat{u} = \begin{cases} 1 & \text{if } V(t,x) = G(x), \\ 0 & \text{otherwise.} \end{cases} \tag{3.4.5}$$

In particular, the Bellman equation (3.4.2), the terminal condition (3.4.3), the auxiliary function w and the equilibrium stopping rule $\hat{u}$ together comprise the Bellman system. The triple $(V, w, \hat{u})$ is called the solution to the Bellman system.

We now state the verification theorem. Since the value function of many interesting stopping problems does not satisfy the $C^{1,2}$ regularity, we consider our problem in a larger function space $W^{1,2}_{loc}([0,T) \times R^+)$, in which the Ito formula still applies[2].

Theorem 3.4.3 (Verification Theorem.) *Assume that the triple $(V, w, \hat{u})$ solves the Bellman system defined by Definition 3.4.2. Assume that $V \in W^{1,2}_{loc}([0,T) \times \mathbb{R}^+) \cap C([0,T] \times \mathbb{R}^+)$ and $|V(t,x)| \le k_1(1+|x|^m)$, for some constants $k_1 > 0, m \ge 1$. Then $\hat{u}$ is an equilibrium stopping rule defined by Definition 3.3.1 and $V(t,x) = \int_0^\infty w(t,x;r)dF(r)$.*

Proof. First, we show that V is the value of a performance functional valued at $\hat{u}$.

At any point $(t,x) \in \mathcal{S}_{\hat{u}}, V(t,x) = G(x) = J(t,x;\hat{u})$.

At any point $(t,x) \in \mathcal{C}_{\hat{u}}$, consider the function w. We show that w has the

[2]See, for example, Krylov (2008), Chapter 2.

polynomial growth. Note that c, G have the polynomial growth, then we have

$$\begin{aligned}
w(t,x;r) &\leq \mathbb{E}\int_t^T |c(X_s)|ds|X_t = x] + \mathbb{E}[|G(X_{\tau_{\hat{u}}\wedge T})||X_t = x] \\
&\leq \mathbb{E}\int_t^T |C(1+|X_s|^m)|ds|X_t = x] + \mathbb{E}[C(1+|X_{\tau_{\hat{u}}\wedge T}|^m)|X_t = x] \\
&\leq \mathbb{E}\int_t^T |C(1+\max_{t\leq s\leq T}|X_s|^m)|ds|X_t = x] + \mathbb{E}[C(1+\max_{t\leq s\leq T}|X_s|^m)|X_t = x]
\end{aligned}$$

Then the estimate for $\mathbb{E}[\max_{t\leq s\leq T}|X_s|^m]$ in (3.1.2) implies that w has the polynomial growth, i.e., there exist $m \geq 1$, $C_1 > 0$ dependent of m, T and the bound of b, σ, such that $w(t,x;r) \leq C_1(1+|x|^m)$.

We now show that Dynkin's formula can be used on w. Define the stopping time $\theta_n = T \wedge \tau_{\hat{u}}^t \wedge \inf\{s > t : |X_s - x| \geq n\}$, then by Ito formula, we have

$$\mathbb{E}[w(\theta_n, X_{\theta_n}; r)|X_t = x] = w(t,x) + \mathbb{E}[\int_t^{\theta_n} (\mathcal{M}w)_s(X_s)ds|X_t = x].$$

By Feynman-Kac formula, before the stopping time θ_n, the function w follows

$$(\mathcal{M}w)_t(x) - rw + c(x) = 0, \tag{3.4.6}$$

then we have

$$\mathbb{E}[w(\theta_n, X_{\theta_n}; r)|X_t = x] = w(t,x) + \mathbb{E}[\int_t^{\theta_n} rw(s, X_s; r) - c(X_s)ds|X_t = x].$$

Since w, c have the polynomial growth, then together with the estimate for the $\mathbb{E}[\max_{t\leq s\leq T}|X_s|^m]$ in (3.1.2), we can pass the limit in the expectation. Let $n \to \infty$, we have

$$\mathbb{E}[w(T\wedge \tau_{\hat{u}}^t, X_{T\wedge\tau_{\hat{u}}^t}; r)|X_t = x] = w(t,x) + \mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t} rw(s, X_s; r) - c(X_s)ds|X_t = x].$$

Similar to the discussion for w, we also have that Dynkin's formula can be used on V, which yields that

$$\begin{aligned}\mathbb{E}[V(T\wedge\tau_{\hat{u}}^t, X_{T\wedge\tau_{\hat{u}}^t})|X_t=x] &= V(t,x)+\mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t}(\mathcal{M}V)_s(X_s)ds|X_t=x]\\ &= V(t,x)+\mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t}(-c(X_s)+\int_0^\infty rw(s,X_s;r)dF(r))ds|X_t=x]\end{aligned}$$

We now consider the expectation,

$$\mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t}(\int_0^\infty rw(s,X_s;r)dF(r))ds|X_t=x] = \int_0^\infty \mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t} rw(s,X_s;r)ds|X_t=x]dF(r)$$

By Feynman-Kac formula, before the stopping time $\tau_{\hat{u}}^t$, the function w follows the equation (3.4.6). Then we have

$$\begin{aligned}\mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t} rw(s,X_s;r)ds|X_t=x] &= \mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t}(\mathcal{M}w)_s(X_s)+c(X_s))ds|X_t=x]\\ &= \mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t} c(X_s)ds|X_t=x]-w(t,x;r)\\ &+\mathbb{E}[w(T\wedge\tau_{\hat{u}}^t, X_{T\wedge\tau_{\hat{u}}^t};r)|X_t=x]\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{E}[V(T\wedge\tau_{\hat{u}}^t, X_{T\wedge\tau_{\hat{u}}^t})|X_t=x] = V(t,x)&+\int_0^\infty \mathbb{E}[w(T\wedge\tau_{\hat{u}}^t, X_{T\wedge\tau_{\hat{u}}^t};r)|X_t=x]dF(r)\\ &-\int_0^\infty w(t,x;r)dF(r).\end{aligned}$$

By the definition of V and w, it is to see that $V(T\wedge\tau_{\hat{u}}^t, X_{T\wedge\tau_{\hat{u}}^t}) = w(T\wedge\tau_{\hat{u}}^t, X_{T\wedge\tau_{\hat{u}}^t};r)$.

It follows that

$$
\begin{aligned}
V(t,x) &= \int_0^\infty w(t,x;r)dF(r) \\
&= \int_0^\infty \mathbb{E}[\int_t^{T\wedge\tau_u^t} e^{-r(s-t)}c(X_s)ds + e^{-r(T\wedge\tau_{\hat{u}}^t-t)}G(X_{T\wedge\tau_{\hat{u}}^t})|X_t = x]dF(r) \\
&= \mathbb{E}[\int_t^{T\wedge\tau_u^t}\int_0^\infty e^{-r(s-t)}dF(r)c(X_s)ds + \int_0^\infty e^{-r(T\wedge\tau_{\hat{u}}^t-t)}dF(r)G(X_{T\wedge\tau_{\hat{u}}^t})|X_t = x] \\
&= J(t,x;\hat{u}).
\end{aligned}
$$

Second, we prove that the stopping rule $\hat{u}$ satisfies Definition 3.3.1.

For the stpping rule $u^{t,\epsilon,a}$, if $a=1$, then $J(t,x;u^{t,\epsilon,a}) = G(x)$. Thanks to the Bellman system, we have $V(t,x) \geq G(x)$. This yields (3.3.1).

If $a=0$, similar to the heuristic derivation of the Bellman system, we have that

$$
\begin{aligned}
J(t,x;u^{t,\epsilon,0}) &= \mathbb{E}[\int_t^{t+\epsilon} h(s-t)c(X_s)|X_t=x] \\
&+ \mathbb{E}[\int_{t+\epsilon}^{T\wedge\tau_{\hat{u}}^{t+\epsilon}} (h(s-t)-h(s-(t+\epsilon)))c(X_s)ds|X_t=x] \\
&+ \mathbb{E}[(h(T\wedge\tau_{\hat{u}}^{t+\epsilon}-t)-h(T\wedge\tau_{\hat{u}}^{t+\epsilon}-(t+\epsilon)))G(X_{T\wedge\tau_{\hat{u}}^{t+\epsilon}})|X_t=x] \\
&+ \mathbb{E}[V(t+\epsilon,X_{t+\epsilon})|X_t=x] \\
&= \mathbb{E}[\int_t^{t+\epsilon} h(s-t)c(X_s)|X_t=x] \\
&+ \mathbb{E}[\int_{t+\epsilon}^{T\wedge\tau_{\hat{u}}^{t+\epsilon}} (\int_0^\infty e^{-r(s-(t+\epsilon))}(e^{-\epsilon r}-1)dF(r)c(X_s)ds|X_t=x] \\
&+ \mathbb{E}[\int_0^\infty e^{-r(T\wedge\tau_{\hat{u}}^{t+\epsilon}-(t+\epsilon))}(e^{-\epsilon r}-1)dF(r)G(X_{T\wedge\tau_{\hat{u}}^{t+\epsilon}})|X_t=x] \\
&+ \mathbb{E}[V(t+\epsilon,X_{t+\epsilon})|X_t=x] \\
&= \mathbb{E}[\int_t^{t+\epsilon} h(s-t)c(X_s)|X_t=x] \\
&+ \mathbb{E}[(\int_0^\infty (e^{-\epsilon r}-1)w(t+\epsilon,X_{t+\epsilon};r)dF(r)|X_t=x] \\
&+ \mathbb{E}[\int_t^{t+\epsilon} (\mathcal{M}V)_s(X_s)ds|X_t=x]+V(t,x).
\end{aligned}
$$

Then by the Bellman system, we have that

$$
\begin{aligned}
\frac{J(t,x;u^{t,\epsilon,0})-J(t,x;\hat{u})}{\epsilon} &\geq \frac{1}{\epsilon}(\mathbb{E}[\int_t^{t+\epsilon} h(s-t)c(X_s)|X_t=x] \\
&+ \mathbb{E}[\int_0^\infty (e^{-\epsilon r}-1)w(t+\epsilon,X_{t+\epsilon};r)dF(r)|X_t=x] \\
&+ \mathbb{E}[\int_t^{t+\epsilon}(-c(X_s)+\int_0^\infty rw(s,X_s;r)dF(r))ds|X_t=x]) \\
&= \mathbb{E}[\frac{\int_t^{t+\epsilon}(h(s-t)-1)}{\epsilon}c(X_s)|X_t=x] \\
&+ \mathbb{E}[\int_0^\infty \frac{e^{-\epsilon r}-1}{\epsilon}w(s,X_s;r)dF(r)|X_t=x] \\
&+ \frac{\mathbb{E}[\int_t^{t+\epsilon}(\int_0^\infty rw(s,X_s;r)dF(r))ds|X_t=x]}{\epsilon}.
\end{aligned}
$$

Due to the growth condition on c, G, all variables in the expectation can be dominated by $K_1(1+\max_{0\leq s\leq T}|X_s|^m), m \geq 1, K_1 > 0$ dependent of T, m and the bound of b, σ. Hence, by the estimate (3.1.2), we can pass the limit into the expectation by the dominated convergence theorem. Then we have

$$\liminf_{\epsilon\to 0} \frac{J(t,x;\hat{u}) - J(t,x;u^{t,\epsilon,0})}{\epsilon} \geq 0$$

This completes the proof. ■

3.5 Discussions

The term $-\int_0^\infty rw(t,x;r)dF(r)$ distinguishes the Bellman equation (3.4.2) from the conventiional one. We now define $-\int_0^\infty rw(t,x;r)dF(r)$ by $f(t,x)$. By simple calculation, we have that

$$f(t,x) = \mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t} h'(s-t)c(X_s)ds + h'(T\wedge\tau_{\hat{u}}^t - t)G(X_{T\wedge\tau_{\hat{u}}^t})|X_t = x].$$

It is to see that f plays the same role as the discounting term in the conventional Bellman equation for a time consistent stopping problem. Due to the presence of the non-constant discount rate $-\frac{h'(T\wedge\tau_{\hat{u}}^t - t)}{h(T\wedge\tau_{\hat{u}}^t - t)}$, we probably can not find any explicit relation between f and the value function V. This feature of f, to a large extent, reflects the essence of time inconsistency, that is, the future preferences vary over time and thus can not be described by the current value function.

In some papers about time inconsistent control (e.g., Bjork and Murgoci (2010)), the "dynamics" of f is described as a continuum of differential equations. Turning to our stopping problem, using the similar idea, we introduce the function

$$f^t(s,x) = \mathbb{E}[\int_s^{T\wedge\tau_{\hat{u}}^s} h'(v-t)c(X_v)dv + h'(T\wedge\tau_{\hat{u}}^s - t)G(X_{T\wedge\tau_{\hat{u}}^s})|X_s = x].$$

At $\forall t \leq s$, the function f^t can be viewed as a function of (s, x). Hence the auxiliary function $f(t, x)$ can be obtained by solving the following dynamic systems until $s = t$, for $\forall t \in [0, T)$,

$$(\mathcal{M}f^t)_s(x) + h'(s-t)c(x) = 0, (s, x) \in \mathcal{C}_{\hat{u}}, \tag{3.5.1}$$

with the boundary and terminal conditions given respectively by

$$f^t(s, x)|_{\mathcal{S}_{\hat{u}}} = h'(s-t)G(x)|_{\mathcal{S}_{\hat{u}}},$$

$$f^t(T, x) = h'(T-t)G(x).$$

In fact, the differential equation (3.5.1) can not be viewed as the dynamics of f, since it can not describe $f(t, x)$ whenever the agent deviates from time t. In other words, if we had to interpret (3.5.1) as a dynamic system of f, then this dynamic system would be one that can not go forward dynamically. This fact, in turn, features continuously changing time preferences, that is, the preferences vary over time and thus can not be described by a single dynamic system.

A fundamental difference between the auxiliary functions w and f is that the dynamics of w can be described by a single differential equation but that is not the case for f. Particularly, for any r, the auxiliary function $w(t, x; r)$ can be obtained by the following differential equation

$$(\mathcal{M}w)_t(x) - rw + c(x) = 0, (t, x) \in \mathcal{C}_{\hat{u}}, \tag{3.5.2}$$

with the boundary and terminal conditions respectively given by

$$w(t, x; r) = G(x)|_{\mathcal{S}_{\hat{u}}}, \tag{3.5.3}$$

$$w(T, x; r) = \max\{G(x), 0\}. \tag{3.5.4}$$

The differential equation (3.5.2), the Bellman equation (3.4.2) and Theorem 3.4.3 together allow us to view the time inconsistent stopping problem from a different perspective. That is, the stopping decision of an agent with weighted discounting seems to be made by a committee, in which each member is time consistent; and any deviation from the stopping rule in a infinitesimal time interval would leave the choices of the agent worse off, see the Bellman equation (3.4.2). We will continue our discussion in this line in next section.

We close the discussion by the following example, which shows that when the time preference is modeled by an exponential function, the Bellman system defined by Definition 3.4.2 goes back to the conventional one.

Example 3.5.1 *If $h(t) = e^{-rt}$ with $r > 0$, then*

$$f(t, x) = -r\mathbb{E}[\int_t^{T\wedge\tau_{\hat{u}}^t} e^{-r(s-t)}c(X_s)ds + e^{-r(T\wedge\tau_{\hat{u}}^t - t)}G(X_{T\wedge\tau_{\hat{u}}^t})|X_t = x] = -rV(t, x),$$

and the Bellman equation (3.4.2) becomes

$$\max\{(\mathcal{M}V)_t(x) - rV(t, x) + c(x), G(x) - V(t, x)\} = 0, \tag{3.5.5}$$

which is the classical Bellman equation for pricing an American derivative with exponential discounting.

3.6 An extension: the infinite horizon case

Many problems in financial economics are modelled by stopping problems with infinite horizon. In this section, we extend our equilibrium stopping model to the infinite horizon case. Here, we focus on the stationary stopping rule u, which is defined as follows.

Definition 3.6.1 *A* stationary stopping rule *is a measureable function of the process value* $u : (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+)) \to (\{0,1\}, 2^{\{0,1\}})$ *where* 0 *indicates "continue" and* 1 *indicates "stop". Each stopping rule* u *defines a sequence of* stopping times $\tau_u^t, t \in [0, \infty)$ *via*

$$\tau_u^t = \inf\{s \geq t, u(X_s) = 1\}, \tag{3.6.1}$$

Similar to the finite horizon case, we introduce the concept of "stopping region"

$$\mathcal{S}_u = \{x \in \mathbb{R}^+ : u(x) = 1\}. \tag{3.6.2}$$

The complement set $\mathcal{C}$ of $\mathcal{S}$ is the continuation region

$$\mathcal{C}_u = \{x \in \mathbb{R}^+ : u(x) = 0\}. \tag{3.6.3}$$

Then the stopping rule u can be shown as follows.

$$u = \begin{cases} 1 & \text{if } x \in \mathcal{S}_u, \\ 0 & \text{if } x \in \mathcal{C}_u. \end{cases}$$

The corresponding stopping time at time t is a hitting time given by

$$\tau_u = \inf\{s \geq 0 : X_s \in \mathcal{S}_u\}.$$

For any stopping rule, we introduce the performance functional[3],

$$J(t,x;u) = \mathbb{E}[\int_t^{\tau_u^t} h(s-t)c(X_s)ds + h(\tau_u^t - t)G(X_{\tau_u^t})|X_t = x], \tag{3.6.4}$$

The stationary underlying process yields a time homogeneous performance functional, i.e.,

$$J(t,x;u) = J(0,x,u) = \mathbb{E}[\int_0^{\tau_u^0} h(s-t)c(X_s)ds + h(\tau_u^0)G(X_{\tau_u^0})|X_0 = x]. \tag{3.6.5}$$

(3.6.5) implies that each self of the agent encounters the same stopping problem. This allows us to consider the optimal stopping problem only at time 0. Hence, for sake of convenience, we abbreviate $J(0,x;u)$ and τ_u^0 to $J(x;u)$ and τ_u respectively.

We now introduce an equilibrium stopping rule for the infinite horizon case as follows.

Definition 3.6.2 *The stopping rule $\hat{u}$ is an equilibrium stopping rule if*

$$\liminf_{\epsilon\to 0} \frac{J(x;\hat{u}) - J(x;u^{\epsilon,a})}{\epsilon} \geq 0 \tag{3.6.6}$$

where $u^{\epsilon,a}$ is defined as follows,

$$u^{\epsilon,a}(s,x) = \begin{cases} \hat{u}(x) & \text{if } s \in [\epsilon,\infty), \\ a & \text{if } s \in [0,\epsilon). \end{cases} \tag{3.6.7}$$

with $a \in \{0,1\}, \epsilon > 0$.

Next, we will introduce the Bellman system for the value function $V := J(x;\hat{u})$.

[3]Here we assume that $\lim_{t\to\infty} h(t)G(X_t) = \lim_{t\to\infty} h(t)c(X_t) = 0$ almost surely. This condition says that the discounted payoff eventually vanishes as time goes by and ensures that waiting forever cannot be optimal behavior. A stronger condition, which is specific to exponential discounting and a geometric Brownian motion setting (e.g., Dixit and Pindyck 1994), is that the drift of the process is smaller than the discount rate.

Similar to the finite horizon case in last section, we define the Bellman system as follows.

Definition 3.6.3 *Given the performance functional (3.6.5) with weighted discount function $h(t) = \int_0^\infty e^{-rt}dF(r)$, where the function F is a distribution function of a non-negative random variable, the Bellman equation for the value function V in $\mathbb{R}^+$ is given by*

$$\max\{(\mathcal{A}V)(x) + c(x) - \int_0^\infty rw(x;r)dF(r) + c(x), G(x) - V(x)\} = 0, \qquad (3.6.8)$$

Here, $\mathcal{A}$ is the infinitesimal generator for the underlying process X given by

$$\mathcal{A} = \frac{1}{2}\sigma(x)^2 x^2 \frac{\partial^2}{\partial x^2} + b(x)x\frac{\partial}{\partial x},$$

and the auxiliary function w is defined as

$$w(x;r) = \mathbb{E}[\int_0^{\tau_{\hat{u}}} e^{-rs}c(X_s)ds + e^{-r\tau_{\hat{u}}}G(X_{\tau_{\hat{u}}})|X_0 = x]. \qquad (3.6.9)$$

The equilibrium stopping rule $\hat{u}$ is given by

$$\hat{u} = \begin{cases} 1 & \text{if } V(x) = G(x), \\ 0 & \text{otherwise.} \end{cases} \qquad (3.6.10)$$

In particular, the Bellman equation (3.6.8), the auxiliary function w and the equilibrium stopping rule $\hat{u}$ together comprise the Bellman system. The triple $(V, w, \hat{u})$ is called the solution to the Bellman system.

Parallel to the finite horizon case, we show that the Bellman system defined by Definition 3.6.3 characterizes the equilibrium stopping rule as follows.

Theorem 3.6.4 (Verification theorem: the stationary case.) *Consider a* *stationary stopping rule $\hat{u}$, functions $w(x;r)$ and $V(x) = \int_0^\infty w(x;r)dF(r)$. Assume that $V \in W^2_{loc}(\mathbb{R}^+)$ and $|V(x)| \leq k_2(1+|x|^m)$, for some constants $k_2 > 0, m \geq 1$. If the triple $(V, w, \hat{u})$ solves the Bellman system defined by Definition 3.6.3. then $\hat{u}$ is an equilibrium stopping rule defined by Definition 3.6.2 and V is the corresponding value function.*

Proof. The proof is essentially the same as the proof of Theorem 3.4.3, we thus omit it. ■

In the remainder of this section, we explain the intuition behind the Bellman system and how the assumption of weighted discounting makes its way into the result. Theorem 3.6.4 tells us how we can obtain an equilibrium stopping rule $\hat{u}$ or, equivalently, the values of $x \in \mathcal{S}_{\hat{u}}$ where the agent will stop and the values $x \in \mathcal{C}_{\hat{u}}$ where she will continue. The function $w(x;r) = \mathbb{E}[\int_0^{\tau_{\hat{u}}} e^{-rs}c(X_s)ds + e^{-r\tau_{\hat{u}}}G(X_{\tau_{\hat{u}}})]$ depends on the equilibrium stopping rule and describes group member r's expected discounted payoff in equilibrium when the current value of the process is x. If the group consists of just one member with discount rate r, then $V(x) = w(x;r)$, i.e., the value function V is given by that member's expected discounted payoff and equation (3.6.8) becomes the well-known Bellman equation (e.g., Dixit and Pindyck 1994)

$$\max\{\mathcal{A}V(x) - rV + c(x), G(x) - V(x)\} = 0.$$

Note that the above equation is *independent* of $\hat{u}$ and can be solved once the model primitives are given. Then, $\hat{u}$ is obtained immediately from equation (3.6.10).[4]

Now consider the case of several group members. Along the equilibrium stop-

[4] In other words, in that case equation (3.6.10) is decoupled from – and thus not a part of – the Bellman system. Therefore, we are back to the classical case of a single Bellman equation.

ping rule $\hat{u}$, the value function $V(x)$ can be written as

$$\begin{aligned} V(x) &= \mathbb{E}[\int_0^{\tau_{\hat{u}}} h(s)c(X_s)ds + h(\tau_{\hat{u}})G(X_{\tau_{\hat{u}}})|X_0 = x] \\ &= \mathbb{E}[\int_0^{\tau_{\hat{u}}} \int_0^{\infty} e^{-rs}dF(r)c(X_s)ds + \int_0^{\infty} e^{-r\tau_{\hat{u}}}dF(r)G(X_{\tau_{\hat{u}}})|X_0 = x] \\ &= \int_0^{\infty} \mathbb{E}[\int_0^{\tau_{\hat{u}}} e^{-rs}c(X_s)ds + e^{-r\tau_{\hat{u}}}G(X_{\tau_{\hat{u}}})|X_0 = x]dF(r) \\ &= \int_0^{\infty} w(x;r)dF(r). \end{aligned} \tag{3.6.11}$$

This calculation shows that the weighted form of the discount function carries over to the value function. Moreover, it clarifies the motivation behind defining the function $w(x;r)$. To derive the Bellman equation (3.6.8), we first write down the differential equation describing w and, then, exploit the relationship between V and w derived above. More specifically, because w features standard exponential discounting it satisfies

$$\mathcal{A}w(x;r) - rw(x;r) + c(x) = 0, x \in \mathcal{C}_{\hat{u}}$$

with the boundary condition $w(x;r)|_{\mathcal{S}_{\hat{u}}} = G(x)|_{\mathcal{S}_{\hat{u}}}$. The latter is the value matching condition. Note that $w(x;r)$ does *not* satisfy a smooth-pasting condition. This is because group member r's discounted expected payoff is not maximized in equilibrium, but only the value function itself is. Since V is the weighted average of the w, by integrating the differential equation of w against F, V must satisfy

$$\begin{aligned} &\int_0^{\infty} \left(\mathcal{A}w(x;r) - rw(x;r)\right) dF(r) + c(x) = 0, x \in \mathcal{C}_{\hat{u}} \\ \Longleftrightarrow &\left(\mathcal{A}\int_0^{\infty} w(x;r)dF(r)\right) - \int_0^{\infty} rw(x;r)dF(r) + c(x) = 0, x \in \mathcal{C}_{\hat{u}} \\ \Longleftrightarrow &\mathcal{A}V(x) - \int_0^{\infty} rw(x;r)dF(r) + c(x) = 0, x \in \mathcal{C}_{\hat{u}}. \end{aligned}$$

Equation (3.6.8) is then obtained by comparing the value of continuation and stopping.

We close this section by considering the example of the pseudo-exponential discount function. In that case, the group consists of two group members with discount rates r and $r + \lambda > r$ whose weights in the decision process are δ and $1-\delta$, respectively; see equation (2.2.1). In that case, we can recover the equilibrium stopping rule from the following four equations:

$$
\begin{aligned}
\max\{\mathcal{A}V(x)- (\delta r w(x;r) + (1-\delta)(r+\lambda)w(x;r+\lambda)) + c(x), G(x) - V(x)\} &= 0,\\
w(x;r) = \mathbb{E}[\int_0^{\tau_{\hat{u}}} e^{-rs} c(X_s)ds + e^{-r\tau_{\hat{u}}} G(X_{\tau_{\hat{u}}})], &\\
w(x;r+\lambda) = \mathbb{E}[\int_0^{\tau_{\hat{u}}} e^{-rs} c(X_s)ds + e^{-(r+\lambda)\tau_{\hat{u}}} G(X_{\tau_{\hat{u}}})], &\\
\hat{u}(x) = \begin{cases} 1 & \text{if } \delta w(x;r) + (1-\delta)w(x;r+\lambda) = G(x), \\ 0 & \text{otherwise,} \end{cases} &
\end{aligned}
$$

We note that, in this case, the last equation is indeed coupled with the other ones.

Chapter 4

Dynamic investment and comparative statics

4.1 Dynamic investment under weighted discounting

This section applies our general optimal stopping result for weighted discount functions to a standard irreversible investment problem as in Brennan and Schwartz (1985), McDonald and Siegel (1986), or Dixit and Pindyck (1994). We derive the optimal investment strategies for several well-known discount functions explicitly.

4.1.1 Economic setup

Consider the opportunity to invest in a project. The payoff process X of the underlying project follows a geometric Brownian motion,

$$\frac{dX_t}{X_t} = bdt + \sigma dW_t. \tag{4.1.1}$$

Investment in X can be made at any time t at cost I. The performance functional of the agent, whose time preferences are described by a weighted discount function h – being obtained from some weighting distribution F – is given by

$$J(x;u) = \mathbb{E}\left[\int_0^\infty e^{-r\tau_u} dF(r)(X_{\tau_u} - I)\right].$$

To ensure the well-posedness of the problem, let

$$b < \inf\{r \in [0,\infty) : F(r) > 0\}. \tag{4.1.2}$$

When F is a step function jumping at $r = r_0$ (the exponential discounting case), condition (4.1.2) is reduced to the standard condition $b < r_0$ (e.g., Dixit and Pindyck 1994, p.141).[1]

4.1.2 Investment behavior under weighted discounting

In this subsection, we solve the Bellman system described in Definition 3.6.3 explicitly when X follows a geometric Brownian Motion (so that $\mathcal{A} = \frac{1}{2}\sigma^2 x^2 \frac{d^2}{dx^2} + bx\frac{d}{dx}$) and when $G(x) = x - I$, for arbitrary weighted discount functions.

The solving scheme is based on the following economic intuition. First, in equilibrium each member of the group uses the same stopping rule. This yields the value matching condition on $w(x;r)$ for all $r \geq 0$. Mathematically, denote the triggering boundary above which investment is optimal by x_*. As usual (e.g., Dixit and Pindyck 1994, p.141), for $x < x_*$ it holds that

$$\frac{1}{2}\sigma^2 x^2 w_{xx}(x;r) + bxw_x(x;r) - rw(x;r) = 0, \tag{4.1.3}$$

[1]Economically, the more general condition (4.1.2) ensures that each member in the group has a non-exploded performance functional, which, in turn, ensures that the performance functional of the group as a whole does note explode. In Appendix B, we provide a formal result on the sufficiency and necessity of the well-posedness condition (4.1.2).

with boundary conditions $w(0;r) = 0$ and $w(x_*;r) = x_* - I$. However, the smooth pasting condition is not imposed on w but on the value function V; see below. The solution to the ordinary differential equation (4.1.3) is given by

$$w(x;r) = \left(\frac{x}{x_*}\right)^{\mu(r)} (x_* - I),\ x < x_*$$

where $\mu(r)$ is the positive square root of the fundamental quadratic (e.g., Dixit and Pindyck 1994, p.142)

$$\frac{1}{2}\sigma^2\mu^2 + (b - \frac{1}{2}\sigma^2)\mu - r = 0,\ \text{i.e.,}\ \mu(r) = \frac{-(b - \frac{1}{2}\sigma^2) + \sqrt{(b - \frac{1}{2}\sigma^2)^2 + 2\sigma^2 r}}{\sigma^2}. \tag{4.1.4}$$

Second, the equilibrium stopping rule maximizes the weighted average of the $w(x;r)$, i.e.,

$$V(x) = \int_0^\infty \left(\frac{x}{x_*}\right)^{\mu(r)} dF(r)(x_* - I). \tag{4.1.5}$$

Therefore, the smooth-pasting condition is given by $\int_0^\infty w_x(x_*;r)dF(r) = 1$. This condition is used to identify the triggering threshold x_*, i.e.,

$$\int_0^\infty \left(\frac{x}{x_*}\right)^{\mu(r)-1} \frac{\mu(r)}{x_*} dF(r)(x_* - I)|_{x=x_*} = 1.$$

It follows that

$$x_* = \frac{A}{A-1}I, \tag{4.1.6}$$

where

$$A = \int_0^\infty \mu(r)dF(r). \tag{4.1.7}$$

Proposition 4.1.1 (Investment behavior under weighted discounting.) *Suppose* *that A defined by (4.1.7) is finite*[2]*. With the notation above, the triple*

$$(V^e(x), w^e(x;r), \hat{u}(x)) = \left\{ \begin{array}{ll} (V(x), w(x;r), 0) & \text{if } x < x_*, \\ (x-I, x-I, 1) & \text{otherwise} \end{array} \right\}$$

solves the Bellman system (3.6.8)–(3.6.10) with its boundary condition. Hence, $\hat{u}$ is an equilibrium stopping rule and V^e is the corresponding value function.

Proof. We prove that the triple $(V^e(x), w^e(x;r), \hat{u})$ defined in Proposition 4.1.1 indeed solves the Bellman system.

On the continuation region $(0, x_*)$, since $V(x) = \int_0^\infty w(x;r)dF(r)$ and since $w(x;r)$ follows the ordinary differential equation (4.1.3), we have that

$$\begin{aligned} &\int_0^\infty \left[\frac{1}{2}\sigma^2 x^2 w_{xx}(x;r) + bxw_x(x;r)\right] dF(r) = \int_0^\infty rw(x;r)dF(r) \\ \Longleftrightarrow &\frac{1}{2}\sigma^2 x^2 V_{xx}(x) + bxV_x(x) - \int_0^\infty rw(x;r)dF(r) = 0. \end{aligned}$$

To obtain the Bellman equation (4.1.3), it remains to check that

$$V(x) - (x-I) > 0 \tag{4.1.8}$$

on the continuation region $(0, x_*)$ as well as that

$$\frac{1}{2}\sigma^2 x^2 (x-I)_{xx} + bx(x-I)_x - \int_0^\infty r(x-I)dF(r) \le 0,$$

or, equivalently,

$$\int_0^\infty (r-b)xdF(r) \ge \int_0^\infty rIdF(r) \tag{4.1.9}$$

on the stopping region $[x_*, \infty)$. As regards inequality (4.1.8), due to condition (4.1.2), we have that $\mu(r) > 1$ and thus $V(x)$ is a convex function. Then, the

[2] A simple condition that guarantees this assumption is $-h'(0) < \infty$.

linearity of the payoff function, together with the value matching and smooth pasting conditions, yields inequality (4.1.8).

As regards inequality (4.1.9), since $b < \inf\{r \in [0, \infty) : F(r) > 0\}$, the function $\int_0^\infty (r - b) x dF(r)$ is increasing in x. Therefore, we only need to prove inequality (4.1.9) at $x = x_*$. If $b \leq 0$, the fact that $\mu(r) > 1$ yields the inequality immediately. Hence, for the remainder of the proof we can assume that $b > 0$.

At $x = x_*$, inequality (4.1.9) becomes

$$\int_0^\infty (r - b) \frac{\int_0^\infty \mu(r) dF(r)}{\int_0^\infty \mu(r) dF(r) - 1} dF(r) \geq \int_0^\infty r dF(r).$$

As μ is a concave function and because $\mu(r) > 1$ for all $r > b$, we have

$$\int_0^\infty (r - b) \frac{\int_0^\infty \mu(r) dF(r)}{\int_0^\infty \mu(r) dF(r) - 1} dF(r) \geq (r_m - b) \frac{\mu(r_m)}{\mu(r_m) - 1}, \tag{4.1.10}$$

where $r_m = \int_0^\infty r dF(r)$. Finally, it remains to prove that

$$(r_m - b) \frac{\mu(r_m)}{\mu(r_m) - 1} \geq r_m, \tag{4.1.11}$$

which is equavalent to

$$r_m - b\mu(r_m) \geq 0.$$

For any $r \geq b$, define $f(r) = r - b\mu(r)$. A simple calculation yields that

$$\begin{aligned} f'(r) &= 1 - \frac{b}{\sqrt{(b - \frac{1}{2}\sigma^2)^2 + 2\sigma^2 r}} \\ &\geq 1 - \frac{b}{b + \frac{1}{2}\sigma^2} \geq 0. \end{aligned}$$

Note that $f(b) = 0$. Therefore, $f(r) \geq 0$, which yields inequality (4.1.11). This completes the proof. ■

Let us compare the solution to the standard result under exponential discounting, which is easily recovered from our more general result. In that case, $h(t) = e^{-r_0 t}$ and the weighted average is trivially $A = \mu(r_0)$, and the triggering boundary is given by the well-known

$$x_* = \frac{\mu(r_0)}{\mu(r_0) - 1} I.$$

A group with with a non-degenerate distribution of opinions $F(r)$ takes a weighted average of the $\mu(r)$ as the coefficient A. The coefficient A determines investment behavior through the triggering boundary x_*. Thus, we note that not only the value function V^e in the solution $(V^e, w^e, \hat{u})$ inherits the weighted representation of the discount function. In the real options application studied here, also the important coefficient A takes a weighted form.

We close this section by presenting analytical solutions for real-option investment under hyperbolic and pseudo-exponential discounting. For the generalized hyperbolic discount function from equation (2.2.2), the triggering boundary x_* is given by

$$x_* = \frac{\displaystyle\int_0^\infty \mu(r) \frac{r^{\frac{\beta}{\alpha}-1} e^{-\frac{r}{\alpha}}}{\alpha^{\frac{\beta}{\alpha}} \Gamma(\frac{\beta}{\alpha})} dr}{\displaystyle\int_0^\infty \mu(r) \frac{r^{\frac{\beta}{\alpha}-1} e^{-\frac{r}{\alpha}}}{\alpha^{\frac{\beta}{\alpha}} \Gamma(\frac{\beta}{\alpha})} dr - 1} I.$$

To the best of our knowledge, our paper is the first to solve a stopping problem with continuously changing time preferences, as is the case for hyperbolic discounting. Most of the literature, instead, has employed continuous-time versions of quasi-hyperbolic discounting. These are also covered by our weighted framework, as discussed next.

For the pseudo-exponential discount function $h(t) = \delta e^{rt} + (1 - \delta) e^{(r+\lambda)t}, \lambda >$

$0, r > 0, 0 < \delta < 1$, the triggering boundary x_* is given by

$$x_* = \frac{\delta\mu(r) + (1-\delta)\mu(\lambda + r)}{\delta\mu(r) + (1-\delta)\mu(\lambda + r) - 1} I.$$

The representation of the triggering boundary x_* for pseudo-exponential discounting is the same as the one obtained by Grenadier and Wang (2007) who considered an investment timing problem in the context of stochastic quasi-hyperbolic discounting.[3] Harris and Laibson (2013) noted that the pseudo-exponential discount function is the expectation of the stochastic quasi-hyperbolic discount function. They argue that the two types of discount functions lead to the same equilibrium in a consumption-savings model. The above example on pseudo-exponential discounting shows that this intuition carries through to an investment timing problem.

4.2 Comparative statics of investment behavior

In this section, we study the investment behavior under weighted discounting in more detail. We generalize some well-known comparative static results to arbitrary weighted discount functions. Then, we obtain new results by turning to comparative statics with respect to decreasing impatience or group diversity, for example.

[3] Grenadier and Wang (2007) employ a continuous-time version of quasi-hyperbolic discounting. In this model, the quasi-hyperbolic discount function takes a stochastic form, i.e., at time t the agent's self n applies the discount factor $D_n(t,s)$ given by $D_n(t,s) = \left\{ \begin{array}{ll} e^{-r(s-t)} & \text{if } s \in [t_n, t_{n+1}), \\ \delta e^{-r(s-t)} & \text{if } s \in [t_{n+1}, \infty). \end{array} \right\}$
where $\{t_n\}_{n\geq 1}$ is a sequence of arrival times which follow a Poisson process with intensity λ and which are independent of the payoff process X.

4.2.1 The impact of the underlying process and entry cost on the triggering boundary x_*

The triggering boundary x_* in equation (4.1.6) depends on the project return rate and volatility, the entry cost, and the discount function. In particular, the influence of these economic factors is independent from one another and determined by the function $\mu(r)$ in equation (4.1.4), the constant I, and the weighting distribution F, respectively. This observation allows us to study the impact of these factors on the investment decision separately. The following proposition lists several properties of the triggering boundary x_* with respect to the entry costs and payoff process.

Proposition 4.2.1 (Comparative statics for entry cost and project dynamics.) *Suppose that the factor A defined by equation (4.1.7) is finite. Then,*

(i) x_* is greater than the entry cost I.

(ii) x_* increases with the entry cost I.

(iii) x_* increases with the return rate b and volatility σ of the project dynamics.

Proof. Statements (i) and (ii) are trivial, and hence we focus on statement (iii). Since x_* is decreasing with respect to A which is the weighted average of μ, we only need to check the monotonicity of μ with respect to b and σ^2.

For any fixed $r > b$, we redefine the function $\mu(r)$ defined by equation (4.1.4) as the function of b and σ^2 by $\nu(b, \sigma^2)$. By simple calculation, we have

$$\frac{\partial \nu}{\partial b} = \frac{1}{\sigma^2}\left(-1 + \frac{b - \frac{1}{2}\sigma^2}{\sqrt{(b - \frac{1}{2}\sigma^2)^2 + 2\sigma^2 r}}\right) \leq 0 \text{ and}$$

$$\frac{\partial \nu}{\partial \sigma^2} = \frac{1}{\sigma^4\sqrt{(b - \frac{1}{2}\sigma^2)^2 + 2\sigma^2 r}}\left(b\sqrt{(b - \frac{1}{2}\sigma^2)^2 + 2\sigma^2 r} + \frac{1}{2}\sigma^2 b - b^2 - \sigma^2 r\right).$$

If $b \leq 0$, it is easy to see that $\frac{\partial \nu}{\partial \sigma^2} \leq 0$. If $b > 0$, since $b > r$, after some algebra we have

$$\begin{aligned} b\sqrt{(b-\frac{1}{2}\sigma^2)^2 + 2\sigma^2 r} + \frac{1}{2}\sigma^2 b - b^2 - \sigma^2 r &\leq b(b+\frac{1}{2}\sigma^2) + \frac{1}{2}\sigma^2 b - b^2 - \sigma^2 r \\ &= \sigma^2(b-r) < 0. \end{aligned}$$

This yields that $\frac{\partial \nu}{\partial \sigma^2} < 0$ and completes the proof. ■

4.2.2 The impact of the discount factor on the triggering boundary x_*

We first prove the following result, which combines nicely with the general results obtained for weighted discount functions in Section 2. The result is strong in the sense that its assumption is the weakest property that we defined in Section 2.

Proposition 4.2.2 (Larger discount factors imply later investment.) *Consider weighted discount functions h^F and h^G with weighting distributions F and G. Then,*

$$h^G(t) \geq h^F(t) \text{ for all } t \geq 0 \;\; \Rightarrow x_*^G \geq x_*^F,$$

where $x_*^G(x_*^F)$ is the triggering boundary defined by (4.1.6) with weighting distribution $h^G(h^F)$.

Proof. We first prove the following lemma, which provides a semi-analytic representation of the triggering boundary x_* in terms of the weighted discount function h^F (rather than in terms of its weighting distribution F, as does equation (4.1.6)).

Lemma 4.2.3 *Consider a weighted discount function h^F with weighting distribu-*

tion F, which yields the triggering boundary x_* in equation (4.1.6). Then

$$x_* = \frac{\max\{\frac{-2(b-\frac{1}{2}\sigma^2)}{\sigma^2}, 0\} + B}{\max\{\frac{-2(b-\frac{1}{2}\sigma^2)}{\sigma^2}, 0\} + B - 1} I, \tag{4.2.1}$$

where

$$B = \frac{1}{\sqrt{2\pi}\sigma}\int_0^\infty t^{-\frac{3}{2}} e^{-\frac{(b-\frac{1}{2}\sigma^2)^2}{2\sigma^2}t}(1-h^F(t))dt. \tag{4.2.2}$$

After some algebra, the factor A defined by equation (4.1.7) can be written as

$$A = \frac{-(b-\frac{1}{2}\sigma^2)}{\sigma^2} + \frac{\sqrt{2}}{\sigma}\int_0^\infty \sqrt{r+C}dF(r), \tag{4.2.3}$$

where $C = \frac{(b-\frac{1}{2}\sigma^2)^2}{2\sigma^2}$. Note that $\sqrt{r+C}$ can be written as

$$\sqrt{r+C} = \frac{1}{2\sqrt{C}}\int_0^r \frac{1}{\sqrt{1+\frac{1}{C}s}}ds + \sqrt{C}. \tag{4.2.4}$$

Moreover,

$$\frac{1}{\sqrt{1+\frac{1}{C}s}} = \int_0^\infty e^{-st} f(t; \frac{1}{2}, \frac{1}{C})dt, \tag{4.2.5}$$

where $f(t; \frac{1}{2}, \frac{1}{C})$ is the density function of the Gamma distribution with shape parameter $\frac{1}{2}$ and scale parameter $\frac{1}{C}$, i.e.,

$$f(t; \frac{1}{2}, \frac{1}{C}) = \frac{1}{\sqrt{\pi}} C^{\frac{1}{2}} t^{-\frac{1}{2}} e^{-Ct}.$$

Next, we plug equation (4.2.5) into equation (4.2.4) to obtain

$$\begin{aligned}
\sqrt{r+C} &= \frac{1}{2\sqrt{\pi}}\int_0^r \int_0^\infty t^{-\frac{1}{2}} e^{-Ct} e^{-st} dtds + \sqrt{C} \\
&= \frac{1}{2\sqrt{\pi}}\int_0^\infty t^{-\frac{1}{2}} e^{-Ct} \int_0^r e^{-st} dsdt + \sqrt{C} \\
&= \frac{1}{2\sqrt{\pi}}\int_0^\infty t^{-\frac{3}{2}} e^{-Ct}(1-e^{-rt})dt + \sqrt{C},
\end{aligned}$$

and hence

$$\begin{aligned}\int_0^\infty \sqrt{C+r}dF(r) &= \int_0^\infty \frac{1}{2\sqrt{\pi}}\int_0^\infty t^{-\frac{3}{2}}e^{-Ct}(1-e^{-rt})dtdF(r)+\sqrt{C}\\ &= \frac{1}{2\sqrt{\pi}}\int_0^\infty t^{-\frac{3}{2}}e^{-Ct}(1-\int_0^\infty e^{-rt}dF(r))dt+\sqrt{C}\\ &= \frac{1}{2\sqrt{\pi}}\int_0^\infty t^{-\frac{3}{2}}e^{-Ct}(1-h^F(t))dt+\sqrt{C}.\end{aligned}$$

Finally, we plug the last equality into equation (4.2.3) to obtain

$$\begin{aligned}A &= \frac{-(b-\frac{1}{2}\sigma^2)}{\sigma^2}+\frac{|b-\frac{1}{2}\sigma^2|}{\sigma^2}+\frac{1}{\sqrt{2\pi}\sigma}\int_0^\infty t^{-\frac{3}{2}}e^{-\frac{(b-\frac{1}{2}\sigma^2)^2}{2\sigma^2}t}(1-h^F(t))dt\\ &= \max\{\frac{-2(b-\frac{1}{2}\sigma^2)}{\sigma^2},0\}+\frac{1}{\sqrt{2\pi}\sigma}\int_0^\infty t^{-\frac{3}{2}}e^{-\frac{(b-\frac{1}{2}\sigma^2)^2}{2\sigma^2}t}(1-h^F(t))dt,\end{aligned}$$

which completes the proof of the lemma.

We now turn to the proof of Proposition 4.2.2. From equation (4.2.2) in Lemma 4.2.3 it follows, because $h^F \geq h^G$, that $B_F \leq B_G$ with $B_i, i = F, G$ defined in equation (4.2.2). Then, from equation (4.2.1) in Lemma 4.2.3 it follows that the triggering boundary is decreasing in the factor B, which closes the proof.

■

Note that the triggering boundary x_* in equation (4.1.6) is expressed in terms of the weighting distribution F – and not in terms of the corresponding weighted discount function h^F. This fact is a natural consequence of our idea to study the investment behavior represented by the weighting distribution F, which is also crucial for the proofs of this paper. As discussed, even though the relationship between F and the group discount function h^F is one-to-one, the actual correspondence for each and every case can be quite complicated. Resolving this complication comprises the difficulty in the proof of Proposition 4.2.2 above. The proof's main idea is to rewrite the intuitive expression for the triggering threshold, equation (4.1.6), through a more complicated and less intuitive expression that, however, involves the original discount function.

In combination with the results on weighted discount functions obtained in Section 2, Proposition 4.2.2 readily yields a number of economically interesting results that we will state explicitly and investigate in some more detail in the following. First, Proposition 2.4.3 and Proposition 4.2.2 yield the following result.

Proposition 4.2.4 (Decreasing impatience and later investment.) *Consider weighted discount function h^F and h^G with weighting distributions F and G and measures of decreasing impatience P_F and P_G, respectively. Let the expectation of F be no less than that of G and $P_G(t) \geq P_F(t)$ $\forall t \geq 0$. Then $x_*^G \geq x_*^F$, where $x_*^G(x_*^F)$ is the triggering boundary defined by (4.1.6) with discount function $h^G(h^F)$.*

Proof. As discussed in the main text, the result is an immediate consequence of Proposition 2.4.3 and Proposition 4.2.2. ■

Second, Proposition 2.4.5 implies that more patience implies a larger group discount factor. Therefore, from Proposition 4.2.2 we obtain the following result.

Proposition 4.2.5 (More patient groups invest later.) *For weighted discount functions h^F and h^G with weighting distributions F and G we have*

$$h^G \succeq_P h^G \Longrightarrow x_*^G \geq x_*^F,$$

where $x_*^G(x_*^F)$ is the triggering boundary defined by (4.1.6) with discount function $h^G(h^F)$.

Proof. As discussed in the main text, Proposition 2.4.5 yields that more patience implies a larger group discount factor. The claim thus follows from Proposition 4.2.2. ■

We will study the impact of group diversity on investment behavior in some more detail.

4.2.3 The impact of group diversity on the triggering boundary x_*

The following result clarifies the impact of the group decomposition (i.e., of the weighting distribution F) on the investment decision. In particular, it shows that greater group diversity, as defined in Section 2 via stochastic dominance deteriorations, leads to later investment.

Proposition 4.2.6 (More diverse groups invest later.) *Suppose F and G are weighting distributions with finite support. If $G \preceq_{\infty SD} F$, then $x_*^G \geq x_*^F$, where $x_*^G(x_*^F)$ is the triggering boundary defined by (4.1.6) with the weighting distribution G (F).*

Proof. The result follows from Theorem 2.3.5 and Proposition 4.2.2, but here we give a direct proof which offers more economic intuition. Let X and Y be two random variables with distributions F^X and F^Y, respectively. Then, the triggering boundaries can be written as

$$x_*^X = \frac{\mathbb{E}[\mu(X)]}{\mathbb{E}[\mu(X)]-1} I \quad \text{and} \quad x_*^Y = \frac{\mathbb{E}[\mu(Y)]}{\mathbb{E}[\mu(Y)]-1} I,$$

where μ is defined by equation (4.1.4). Note that μ satisfies sgn $\mu^{(n)} = (-1)^{n+1}$ for all $n \in \mathbb{N}^+$, i.e. $\mu(r)$ can be interpreted as a mixed risk averse "utility function over discount rates." Then, due to $F^Y \preceq_{\infty SD} F^X$, we have $\mathbb{E}[\mu(X)] \geq \mathbb{E}[\mu(Y)]$. This yields the conclusion and completes the proof. ■

Note that under greater group diversity it is more difficult to find a consensus investment strategy. In other words, greater group diversity comes along with more restrictions on the equilibrium stopping rule. Therefore, as explained in the introduction, one might have expected that greater group diversity leads to less risk-taking and thus earlier stopping. However, the fact that greater group diversity leads to a uniformly larger discount factor (Theorem 2.3.5) dominates this effect, illustrating once more that Proposition 4.2.2 is a strong result. Since the proof of Proposition 4.2.2 is technical and unintuitive, in the appendix we give another, direct proof of Proposition 4.2.6 that illustrates how greater group diversity defined via stochastic dominance connects to the expression of the triggering threshold (4.1.6).

Next, we illustrate the quantitative impact of the group diversity on the triggering boundaries. To this means, let us consider once more pseudo-exponential discounting. The decision-making group consists of two members with discount rates r and $r+\lambda > r$ and whose weights in the decision process are δ and $1-\delta$, respectively; see equation (2.2.1). The fact that the weighting distribution is binary allows for a simple and intuitive characterization of stochastic dominance relationships through statistical moments; see Chiu (2010) and Ebert (2015b) for details. In particular, the weighting distribution with parameters δ, r, and $r+\lambda$ can be re-parametrized in terms of its mean M, standard deviation SD, and standardized skewness S. Moreover, it can be shown that a binary distribution first- (second-, third-) order dominates another binary distribution with lower mean (greater variance, lower skewness), given that the other two moments are the same. In the following, we can thus conduct comparative statics with respect to the moments of the weighting distribution, rather than having to deal with the iterated integrals in the original definition of stochastic dominance.

Figure 4.1 plots the value function of the real-option investment problem under pseudo-exponential discounting with four different weighting distributions which differ in their first three moments. Table 4.1 shows the parameters for each of the four scenarios. The weighting distribution of the Benchmark scenario first-order stochastically dominates that of the FSD scenario, because it has a larger expected discount rate, while standard deviation and skewness of the discount rate distribution are the same. Figure 4.1 shows that the FSD value function

Table 4.1: Parameters used for the scenarios in Figure 4.1

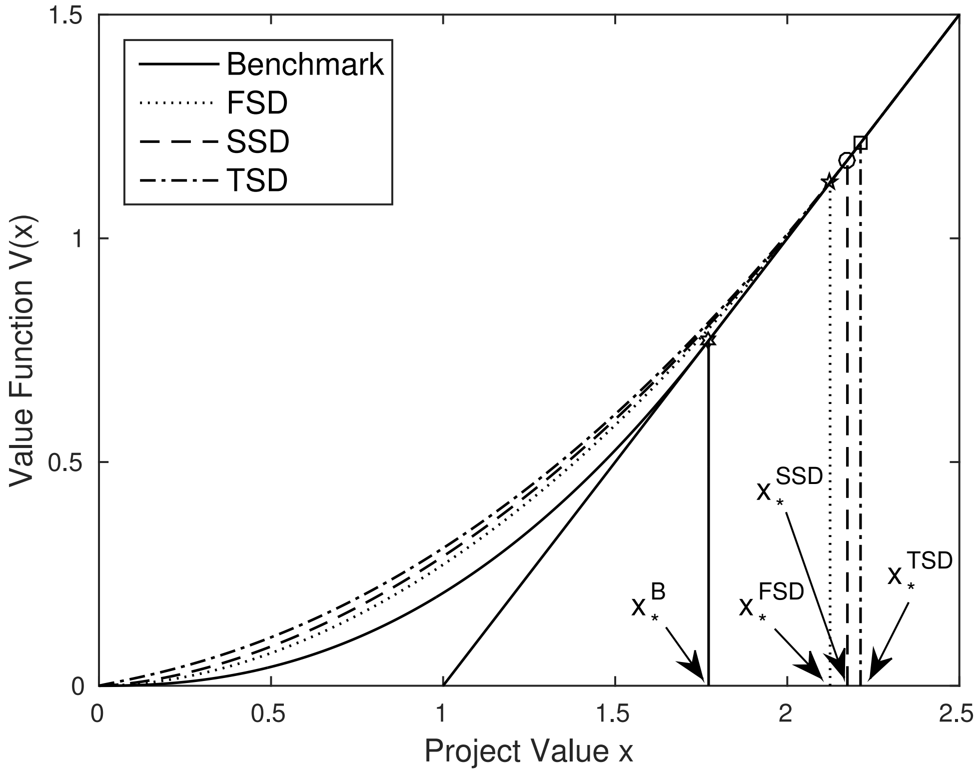
	r	λ	δ	M	SD	S
Benchmark	0.0825	0	1	0.0825	0	0
FSD	0.0525	0	1	0.0525	0	0
SSD	0.03	0.045	0.5	0.0525	0.0225	0
TSD	0.0075	0.0562	0.2	0.0525	0.0225	-1.5000

Notes. The table defines the four scenarios denoted by $Benchmark, FSD, SSD$, and TSD studied in Figure 4.1. Each scenario is characterized by a different parametrization of the weighting distribution of a pseudo-exponential discount function as defined in equation (2.2.1). The parameters of the weighting distribution are shown in the first three columns, while the remaining three columns show its mean M, standard deviation SD, and standardized skewness S. The scenarios are chosen such that the weighting distribution in scenario $Benchmark$ is less first-order diverse than scenario FSD; which in turn is less second-order diverse than scenario SSD; which in turn is less third-order diverse than scenario TSD.

is larger than the Benchmark value function, which implies a larger investment threshold: $x_*^{FSD} > x_*^B$. It is not surprising, of course, that a lower discount rate leads to later investment, but note that this result can be proven (and interpreted) in terms of our result that greater group diversity leads to later investment: In the terminology of stochastic dominance, the FSD weighting distribution is first-order "riskier", i.e., the group in the FSD scenario is first-order more diverse.

Next, the weighting distribution of the FSD scenario second-order stochastically dominates that of the SSD scenario. Equivalently, the SSD weighting distribution is second-order more diverse than that of the FSD scenario. This is because the SSD weighting distribution has a greater standard deviation than the weighting distribution of the FSD scenario, but identical mean and skewness. Greater second-order diversity also leads to later investment: $x_*^{SSD} > x_*^{FSD}$. Finally, the figure illustrates that a weighting distribution with lower skewness (increased third-order diversity) results in later investment: $x_*^{SSD} > x_*^{FSD}$. This means that if one

Figure 4.1: The impact of group diversity on investment behavior



Notes. The figure illustrates that greater group diversity leads to a more risky investment decision (i.e., to stopping later, at a higher threshold level). x_*^i denotes the threshold for scenario i ($i = B, FSD, SSD, TSD$) which is defined in Table 4.1). The figure also plots the value function for each scenario, whose intersection with the payoff schedule determines the threshold. $x_*^B < x_*^{FSD}$ means that we have later stopping under exponential discounting for smaller discount rates (which comprise a first-order stochastic dominance deterioration). $x_*^{FSD} < x_*^{SSD}$ means that a group of two individuals who disagree about the discount rate stops later than a single individual whose discount rate corresponds to the group average. $x_*^{SSD} < x_*^{TSD}$ illustrates that a group who is more diverse in the sense that there is one member with low weight in the decision, but an extremely low opinion about the appropriate discount rate, stops later than a group whose disagreement is more symmetric. The project drift and volatility assumed are $b = 0.01$ and $\sigma = 0.2$, respectively. The project entry cost is set to $I = 1$.

of the group members has small weight δ, but a strong opinion in favor of a very small discount rate, this results in later investment compared to when the weights on group members are more symmetric – even if the group mean and standard deviation of the discount rate are the same.

4.2.4 The impact of discount function parameters on the triggering boundary x_*

Finally, we explain how one can study the impact of discount function parameters on investment behavior. For sake of concreteness, we do this with reference to generalized hyperbolic discounting; see equation (2.2.2). From Prelec (2004, p.524), we know that the parameter α captures decreasing impatience. To the best of our knowledge, a clear economic interpretation of the parameter β has not yet been proposed[4], but we give one here. β is equal to the expectation of the generalized hyperbolic weighting distribution (the Gamma distribution with parameters as in equation (2.2.2)). That is, β identifies the group's weighted mean opinion about the appropriate discount *rate*. This once more illustrates the added value that may arise from interpreting discounting behavior through *both* the weighted discount function and its weighting distribution, as proposed in this paper.

Proposition 4.2.7 (Comparative statics for generalized hyperbolic discounting.) *Consider the generalized hyperbolic discount function* $h(t; \alpha, \beta) = \frac{1}{(1+\alpha t)^{\frac{\beta}{\alpha}}}$ *with* $\alpha, \beta > 0$. *Then the triggering boundary* x_* *defined by equation (4.1.6)*

(i) increases with the degree of decreasing impatience, α.

(ii) decreases with the group weighted mean opinion about the discount rate, β.

[4]Loewenstein and Prelec (1992) merely use the parameter for functional flexibility. Prelec (2004) notes that β controls time preference, but this is only true once α has been fixed. In other words, many different values of β can imply same time preference, and vice versa.

Proof. Noting that

$$h(t) = e^{-\int_0^t \frac{\beta}{1+\alpha s} ds},$$

it is easy to see that for $\beta \leq \beta_1$ and $\alpha \leq \alpha_1$ the functions

$$R_1(t; \alpha, \alpha_1, \beta) := \frac{h(t; \alpha_1, \beta)}{h(t; \alpha, \beta)} \text{ and } R_2(t; \alpha, \beta_1, \beta) := \frac{h(t; \alpha, \beta)}{h(t; \alpha, \beta_1)}$$

are increasing in t. This yields

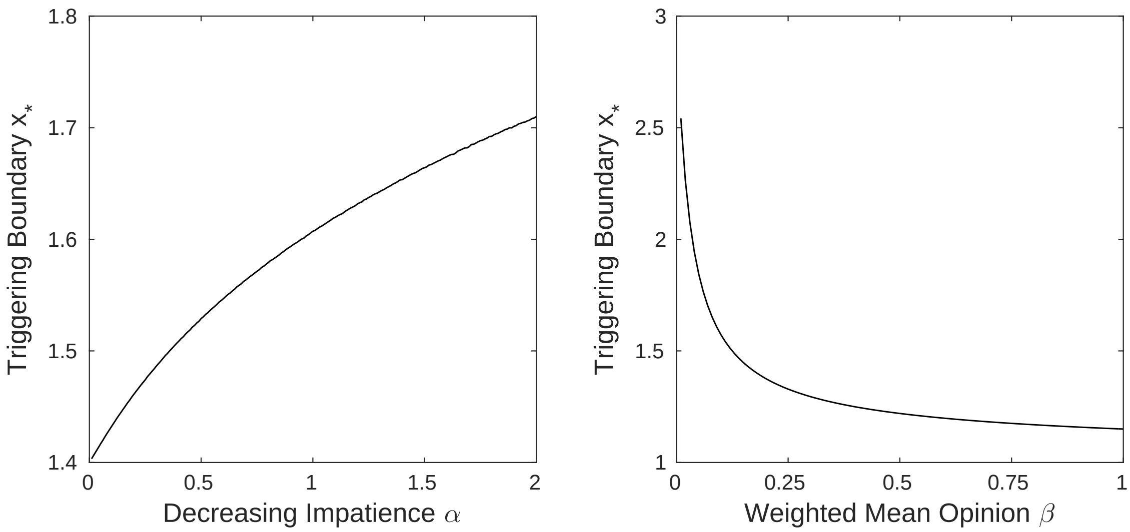
$$h(t; \alpha, \beta_1) \preceq_P h(t; \alpha, \beta) \preceq_P h(t; \alpha_1, \beta), \text{ for } \beta \leq \beta_1, \alpha \leq \alpha_1.$$

Then, the properties of the triggering boundaries (statements (i) and (ii)) follow from Proposition 4.2.5. ■

Figure 4.2 illustrates the quantitative impact of the parameters α and β on the triggering boundary x_*.

We see that greater decreasing impatience leads to later investment (left panel of Figure 4.2). Since the weighting distribution mean is fixed to $\beta = 0.14$, Proposition 2.4.3 applies and says that the discount factor, at any date t, is increasing with greater decreasing impatience α. Proposition 4.2.2 then says that larger discount factors lead to later stopping (larger x_*). As regards the parameter β, we see that a greater mean of the weighting distribution leads to earlier stopping (lower x_*). This latter observation is intuitively consistent with our result that less group diversity (and recall that a first-order stochastic dominance "risk" decrease is an instance of less group diversity) leads to earlier stopping.

Figure 4.2: The impact of decreasing impatience, α, and the weighted mean of the group discount rate, β, on the triggering boundary under generalized hyperbolic discounting (Loewenstein and Prelec 1992)



Notes. The figure plots the variation of the triggering boundary with respect to the degree of decreasing patience α (left panel) and the weighted mean of the group discount rate β (right panel). Both graphs use the parameters $b = -0.01, \sigma = 0.2, I = 1$. In the left graph, $\beta = 0.14$. In the right graph, $\alpha = 0.22$. The parameter range well covers empirical estimates (e.g., Abdellaoui et al. 2010; Attema et al. forthcoming).

Chapter 5

An existence (non-existence) result

In this chapter, we show a solvable example, whose time consistent case is used to analyze monopoly or oligopoly investment under uncertainty[1].

5.1 The model

Let $X = \{X_t\}_{t \geq 0}$ denote the payoff value process and suppose that its dynamics is given by

$$dX_t = \sigma X_t dW_t, X_0 = x \tag{5.1.1}$$

For a stopping rule u, consider the following cost functional,

$$J(x; u) = \mathbb{E}[\int_0^{\tau_u} h(t) X_t dt + h(\tau_u) K], \tag{5.1.2}$$

[1]See, for example, Back and Paulsen (2009) and Dixit and Pindyck (1994).

where h is a weighted discount function defined by Definition 2.1.1 and K is a positive constant.

Due to the time inconsistency incurred by non-exponential discounting, an agent who aims to obtain an optimal solution will turn to pursuing an equilibrium stopping rule given by

Definition 5.1.1 *The stopping rule $\hat{u}$ is an equilibrium stopping rule if*

$$\limsup_{\epsilon \to 0} \frac{J(x;\hat{u}) - J(x;u^{\epsilon,a})}{\epsilon} \leq 0 \tag{5.1.3}$$

where $u^{\epsilon,a}$ is defined as follows,

$$u^{\epsilon,a}(s,x) = \begin{cases} \hat{u}(x) & \text{if } s \in [\epsilon, \infty), \\ a & \text{if } s \in [0, \epsilon). \end{cases} \tag{5.1.4}$$

with $a \in \{0,1\}, \epsilon > 0$.

Similarly, we characterize the equilibrium stopping rule as follows.

Proposition 5.1.2 (Equilibrium Characterization) *Consider the cost functional (5.1.2) with weighted discount function $h(t) = \int_0^\infty e^{-rt} dF(r)$, a stopping rule $\hat{u}$, and functions $w(x;r) = \mathbb{E}[\int_0^{\tau_{\hat{u}}} e^{-rt} X_t dt + e^{-r\tau_{\hat{u}}} K]$ and $V(x) = \int_0^\infty w(x;r)dF(r)$. Suppose that $V \in W^2_{loc}(\mathbb{R}^+)$, has polynomial growth and $(V, w, \hat{u})$ solves*

$$\min\{\frac{1}{2}\sigma^2 x^2 V_{xx} + x - \int_0^\infty r w(x;r) dF(r), K - V\} = 0, \tag{5.1.5}$$

$$\hat{u}(x) = \begin{cases} 1 & \text{if } V = K, \\ 0 & \text{otherwise}, \end{cases} \tag{5.1.6}$$

Then, $\hat{u}$ is an equilibrium stopping rule and the value function of the problem is given by $V(x)$, i.e., $V(x) = J(x;\hat{u})$.

Proof. The proof is essentially the same as the proof of Theorem 3.4.3, we thus omit it. ■

We aim to find an equilibrium stopping rule by solving the Bellman system in Theorem 5.1.2. The solving scheme is based on the so called "value matching + smooth pasting", which is a standard procedure for obtaining a solution to a Bellman system or an optimal stopping problem. By this example, we aim to show that

- the result of "value matching + smooth pasting" may neither solve the Bellman system nor the equilibrium stopping problem;
- the Bellman system may not admit any proper solution.

5.2 An explicit solution

We now look for a solution to the Bellman system by "value matching + smooth pasting".

Suppose that the continuation region $\mathcal{C} = \{x \in \mathbb{R}^+ : x < x_*\}$, we now identify the triggering boundary x_* as follows.

First, for any r, w follows

$$\frac{1}{2}\sigma^2 x^2 w_{xx} - rw + x = 0, x \in \mathcal{C},$$

with the value matching conditions,

$$w(x_*; r) = K,$$
$$w(0; r) = 0.$$

Then it follows that

$$w(x;r) = (K - \frac{x_*}{r})(\frac{x}{x_*})^{\alpha(r)} + \frac{x}{r},$$

where

$$\alpha(r) = \frac{\frac{1}{2}\sigma^2 + \sqrt{\frac{1}{4}\sigma^4 + 2\sigma^2 r}}{\sigma^2}. \tag{5.2.1}$$

Since V is the weighted average of w, then we have

$$\begin{aligned} V(x) &= \int_0^\infty w(x;r)dF(r) \\ &= \int_0^\infty (K - \frac{x_*}{r})(\frac{x}{x_*})^{\alpha(r)} dF(r) + \int_0^\infty \frac{x}{r} dF(r). \end{aligned}$$

Therefore, the smooth pasting condition $V_x(x_*) = 0$ yields that

$$\int_0^\infty (K - \frac{x_*}{r})\alpha(r)\frac{1}{x_*} dF(r) + \int_0^\infty \frac{1}{r} dF(r) = 0,$$

then we have

$$x_* = \frac{\displaystyle\int_0^\infty \alpha(r)dF(r)}{\displaystyle\int_0^\infty \frac{\alpha(r) - 1}{r} dF(r)} K. \tag{5.2.2}$$

To make the above integral finite and interchange of Interchange of differentiation and integration valid, we impose the following assumption in the rest of this chapter.

Assumption 5.2.1

$$\max\{\int_0^\infty \frac{1}{r} dF(r), \int_0^\infty r dF(r)\} < \infty,$$

Proposition 5.2.1 *With the notation above, the triple*

$$(V^e(x), w^e(x;r), \hat{u}(x)) = \left\{ \begin{array}{rr} (V(x), w(x;r), 0) & \text{if } x < x_*, \\ (K, K, 1) & \text{otherwise} \end{array} \right\}$$

solves the Bellman system in Theorem 5.1.2 if and only if

$$\int_0^\infty \alpha(r)dF(r) \geq \int_0^\infty rdF(r) \int_0^\infty \frac{\alpha(r)-1}{r} dF(r). \tag{5.2.3}$$

Proof. We begin with showing the sufficiency.

To show the triple $(V^\epsilon(x), w^\epsilon(x;r), \hat{u}(x))$ solves the Bellman system, it remains to verify

$$\frac{1}{2}\sigma^2 x^2 V_{xx} - \int_0^\infty rw(x;r)dF(r) + x \geq 0, x \geq x_* \tag{5.2.4}$$

and

$$V(x) < K, x < x_*, \tag{5.2.5}$$

Let us first verify the inequality (5.2.4).

Since $V(x) = w(x;r) = K, x \geq x_*$, then (5.2.4) becomes

$$x \geq \int_0^\infty KrdF(r), x \geq x_*.$$

Now we only need to show $x_* \geq \int_0^\infty KrdF(r)$, which follows from the condition (5.2.3) immediately.

For (5.2.5), let us consider V_{xx} on $(0, x_*)$. Note that

$$V_{xx}(x) < 0 \Leftrightarrow \int_0^\infty (K - \frac{x_*}{r})(\frac{x}{x_*})^{\alpha(r)} \alpha(r)(\alpha(r)-1)dF(r) < 0,$$

$$\alpha(r)(\alpha(r)-1)=\frac{2r}{\sigma^2}.$$

Then we have that

$$V_{xx}(x)<0 \Leftrightarrow \int_0^\infty (rK-x_*)(\frac{x}{x_*})^{\alpha(r)}dF(r)<0.$$

Note that both $\alpha(r)$ and $rK-x_*$ are increasing functions of r and $x<x_*$, then the rearangement inequality (e.g., Hardy et al. 1952; Denuit and Dhaene 2003) yields that

$$\int_0^\infty (rK-x_*)(\frac{x}{x_*})^{\alpha(r)}dF(r)<\int_0^\infty (rK-x_*)dF(r)\int_0^\infty (\frac{x}{x_*})^{\alpha(r)}dF(r)$$

Pluging the representation (5.2.2) of x_* into above inequality, we have that $V_{xx}<0$ on $(0,x_*)$ by condition (5.2.3).

Due to the smooth pasting condition, we have that V is increasing. This fact, combined with the value matching condition yields that $V<K, x<x_*$.

We now turn to the necessary condition.

If condition (5.2.3) does not hold, then by simple calculation, we have that $V_{xx}(x_*-)>0$. Then the smooth pasting condition $V_x(x_*)=0$ and the value matching condition $V(x_*)=K$ yield that $V>K$ in a left neighborhood of x_*, which shows that the triple $(V^\epsilon, w^\epsilon, \hat{u})$ does not solve the Bellman system.

This completes the proof. ■

Proposition 5.2.2 *Suppose that the X follows the dynamics given by (5.1.1) and the cost functional is given by (5.1.2), then $\hat{u}$ defined in Proposition 5.2.1 is an equilibrium stopping rule if and only if (5.2.3) holds.*

Proof. The sufficiency is an immediate result of Theorem 5.1.2.

For the necessary condition, from the proof of Proposition 5.2.1, on one hand, we have that there exists $x \in \mathcal{C}$, such that $J(x;\hat{u}) > K$ if condition (5.2.3) does not hold. On the other hand, we have

$$\limsup_{\epsilon \to 0} \frac{J(x;\hat{u}) - J(x;u^{\epsilon,1})}{\epsilon} = \limsup_{\epsilon \to 0} \frac{J(x;\hat{u}) - K}{\epsilon} = \infty$$

where $u^{\epsilon,1}$ is defined in Definition 5.1.4.

This is a contradiction and completes the proof. ■

We close this section by presenting a closed form solution in the generalized hyperbolic case.

Proposition 5.2.3 *Suppose that the discount function h is a generalized hyperbolic discount function, that is, $h(t) = \frac{1}{(1+\gamma t)^{\frac{\beta}{\gamma}}}, \gamma > 0, \beta > 0$ can be written as*

$$h(t) = \int_0^\infty e^{-rt} \frac{r^{\frac{\beta}{\gamma}-1} e^{-\frac{r}{\gamma}}}{\gamma^{\frac{\beta}{\gamma}} \Gamma(\frac{\beta}{\gamma})} dr.$$

where $\Gamma(k)$ is the Gamma function evaluated at k, i.e., $\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx$.

If $\beta \le \frac{\sigma^2}{2}$, then the triggering boundary x_* is given by

$$x_* = \frac{\displaystyle\int_0^\infty \alpha(r) \frac{r^{\frac{\beta}{\gamma}-1} e^{-\frac{r}{\gamma}}}{\gamma^{\frac{\beta}{\gamma}} \Gamma(\frac{\beta}{\gamma})} dr}{\displaystyle\int_0^\infty \frac{\alpha(r)-1}{r} \frac{r^{\frac{\beta}{\gamma}-1} e^{-\frac{r}{\gamma}}}{\gamma^{\frac{\beta}{\gamma}} \Gamma(\frac{\beta}{\gamma})} dr} K.$$

Proof. It suffices to verify the condition (5.2.3).

Note that $\alpha(r) - 1 = -\frac{1}{2} + \frac{\sqrt{\frac{1}{4}\sigma^4 + 2\sigma^2 r}}{\sigma^2}$ is a concave function, then we have

$$\alpha(r) - 1 \le (\alpha(r) - 1)'|_{r=0} r + \alpha(0) - 1 = \frac{2}{\sigma^2} r.$$

Moreover, it is to see that

$$\int_0^\infty r dF(r) = \beta,$$

and

$$\alpha(r) \geq 1.$$

Therefore,

$$\int_0^\infty \frac{\alpha(r)-1}{r} dF(r) \int_0^\infty r dF(r) \leq \beta \frac{2}{\sigma^2} \leq 1 \leq \int_0^\infty \alpha(r) dF(r).$$

This completes the proof. ■

5.3 A non-existence result

In this section, for sake of convenience, we suppose that any weighting distribution F has its support in $[a, c], 0 < a \leq c < \infty$. The main result is that the Bellman system does not admit any proper solution whenever the condition (5.2.3) does not hold.

Lemma 5.3.1 *For any stopping rule u, the cost functional J has linear growth, i.e.,*

$$0 \leq J(x; u) \leq \frac{x}{a} + K. \tag{5.3.1}$$

Proof. This is a result of the following calculation,

$$
\begin{aligned}
0 \le J(x;u) &= \mathbb{E}[\int_0^{\tau_u} h(t)X_t dt + h(\tau_u)K | X_0 = x] \\
&\le \int_0^\infty \mathbb{E}[\int_0^\infty e^{-rt} X_t dt + K | X_0 = x] dF(r) \\
&= \int_0^\infty \frac{x}{r} dF(r) + K \le \frac{x}{a} + K.
\end{aligned}
$$

■

In Theorem 5.1.2, to use Ito formula, we choose $W^2_{loc}(\mathbb{R}^+)$ as our working space. The following proposition shows that whenever the condition (5.2.3) does not hold, the Bellman system does not admit a solution in $W^2_{loc}(\mathbb{R}^+)$ even though the barrier function, running cost and coefficients of the differential equations are smooth.

Proposition 5.3.2 *Consider the following Bellman system*

$$
\min\{\frac{1}{2}\sigma^2 x^2 V_{xx} + x - \int_0^\infty r w(x;r) dF(r), K - V\} = 0, \tag{5.3.2}
$$

$$
\hat{u} = \begin{cases} 1 & \text{if } V(x) = K, \\ 0 & \text{otherwise.} \end{cases} \tag{5.3.3}
$$

$$
w(x;r) = \mathbb{E}[\int_0^{\tau_{\hat{u}}} e^{-rs} X_s ds + e^{-r\tau_{\hat{u}}} K | X_0 = x]. \tag{5.3.4}
$$

where the underlying process X is defined by (5.1.1) and $K > 0$.

If the condition (5.2.3) does not hold, i.e.,

$$
\int_0^\infty \alpha(r) dF(r) < \int_0^\infty r dF(r) \int_0^\infty \frac{\alpha(r) - 1}{r} dF(r),
$$

where $\alpha(r)$ is defined by (5.2.1), then the Bellman system does not admit any solution $(V, w, \hat{u})$ such that $V(x) = \int_0^\infty w(x;r) dF(r)$ and $V \in W^2_{loc}(\mathbb{R}^+)$.

Proof. Suppsoe that $V(x) = \int_0^\infty w(x;r) dF(r)$ we show that $V \notin C^1(\mathbb{R}^+) \supset$

$W^2_{loc}(\mathbb{R}^+)$.

First, we show that there does not exist an open interval $(s_1, s_2) \subset \mathcal{S}_{\hat{u}}$, $s_1 < s_2 < aK$. If not, then $V = w = K$ on (s_1, s_2) and $\frac{1}{2}\sigma^2 x^2 V_{xx} + x - \int_0^\infty rw(x;r)dF(r) < 0$. This is a contradiction to the Bellam equation (5.3.2).

Second, we show that there does not exist an open inverval $(c_1, c_2) \subset \mathcal{C}_{\hat{u}}$, where $c_1 < c_2 < aK$ and $c_1, c_2 \in \mathcal{S}_{\hat{u}}$. If not, then for any $r \in [a, c]$, $w(x;r)$ satisfies the following differential equation

$$\frac{1}{2}\sigma^2 x^2 w_{xx} - rw + x = 0,$$

with the boundary conditions

$$w(c_1; r) = w(c_2; r) = K$$

Therefore $w_{xx}(c_2-) = \frac{2}{\sigma^2 x^2}(rK - c_2) \geq \frac{2}{\sigma^2 c_2^2}(aK - c_2) > 0$ and hence $V_{xx}(c_2-) = \int_0^\infty w_{xx}(c_2-; r)dF(r) > 0$.

On the other hand, from the Bellman equation (5.3.2), we have that $V \leq K$. Then c_2 is a local maximal point for the function V, which implies that $V_{xx}(c_2-) \leq 0$. This is a contradiction.

Based on the above argument, we have that $(0, aK) \subset \mathcal{C}_{\hat{u}}$. Define $d = \inf\{x \in \mathbb{R}^+ : x \in \mathcal{S}_{\hat{u}}\}$. If $d < \infty$, then we can repeat the solving scheme in Section 5.2 and obtain a solution which is the same as $(V^e, w^e, \hat{u})$ on $(0, d)$ in Proposition 5.2.1. However, since condition (5.2.3) does not hold, The solution does not solve the Bellman system.

If $d = \infty$, then

$$\begin{aligned} V(x) &= J(x; \hat{u}) \\ &= \int_0^\infty \mathbb{E}[\int_0^\infty e^{-rt} X_t dt | X_0 = x] dF(r) \\ &= \int_0^\infty \frac{1}{r} dF(r) x. \end{aligned}$$

which is contradict to $V \leq K$.

This completes the proof.

■

In the end of this section, we present an example showing that the non-existence of the solution in $W^2_{loc}(\mathbb{R}^+)$ to the Bellman system is possible.

Proposition 5.3.3 *Suppose the discount function* $h = \delta e^{-rt} + e^{-(r+\lambda)t}, 0 < \delta < 1, r > 0, \lambda > 0$. *Then for any fixed* r, δ, *when* λ *is sufficiently large, the condition (5.2.3) does not hold, i.e.,*

$$\int_0^\infty \alpha(r) dF(r) < \int_0^\infty r dF(r) \int_0^\infty \frac{\alpha(r) - 1}{r} dF(r).$$

where $\alpha(r)$ *is defined by (5.2.1) and* F *defined by (2.2.1).*

Proof. By simple calculation, it is to see that

$$\int_0^\infty \alpha(r) dF(r) = \delta\alpha(r) + (1 - \delta)\alpha(r + \lambda),$$

$$\int_0^\infty r dF(r) \int_0^\infty \frac{\alpha(r) - 1}{r} dF(r) > (1 - \delta)(r + \lambda)\delta(\frac{\alpha(r) - 1}{r}).$$

Then the result follows from the growth rate of $\alpha(r + \lambda)$. ■

5.4 Group diversity, decreasing impatience and non-existence

In this section, we focus on the impact of time preferences on stopping decisions and non-existence of the solution to the Bellman system. The feature that the discount function appears in the explicit solution as the form of the weighting distribution rather than its original form, although reflecting, to some extent, the importance of introducing the concept of weighted discounting, may cause some technical inconvenience when we explore the issue of how discounting affects stopping decisions.

For example, if we write the hyperbolic discount function as its original form $h(t) = \frac{1}{(1+\alpha t)^{\frac{\beta}{\alpha}}} = \exp(-\int_0^t \frac{\beta}{1+\alpha s} ds)$, it is easy to know the function h increases with α and decreases with β. However, if h is written as the LaPlace transform of a Gamma distribution, i.e., $h(t) = \int_0^\infty e^{-rt} \frac{r^{\frac{\beta}{\alpha}-1} e^{-\frac{r}{\alpha}}}{\alpha^{\frac{\beta}{\alpha}} \Gamma(\frac{\beta}{\alpha})} dr$, the monotonicity of h with its parameters α, β probably becomes challenging.

To address this problem, we develop another representation of the triggering boundary, in which the original form of the discount function reappears.

Lemma 5.4.1 *For $\forall r > 0, C > 0$, we have*

$$\frac{\sqrt{C+r}}{r} = \int_0^\infty e^{-sr} f(s;C) ds, \tag{5.4.1}$$

where

$$f(s;C) = \frac{1}{\sqrt{\pi s}} (e^{-Cs} + \sqrt{\pi C s}\, \mathit{erf}(\sqrt{Cs})), \tag{5.4.2}$$

and erf is the error function, i.e., $\mathit{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

Proof. It is to see that

$$\frac{\sqrt{C+r}}{r} = \frac{C+r}{r\sqrt{C+r}} = \frac{C}{r\sqrt{C+r}} + \frac{1}{\sqrt{C+r}}$$

From the standard textbook about LaPlace transform[2], we have that

$$\frac{1}{\sqrt{C+r}} = \int_0^\infty \frac{e^{-Cs}}{\sqrt{\pi s}} e^{-sr} ds,$$

$$\frac{C}{r\sqrt{C+r}} = \sqrt{C} \int_0^\infty erf(\sqrt{Cs}) e^{-sr} ds,$$

which yields the result. ■

Lemma 5.4.2 *For the function $\alpha(r), r \geq 0$ defined by (5.2.1) and the distribution function F, we have that*

$$\int_0^\infty \frac{\alpha(r)-1}{r} dF(r) = \int_0^\infty h^F(s)\left(\frac{\sqrt{2}}{\sigma} f(s; \frac{1}{8}\sigma^2) - \frac{1}{2}\right) ds,$$

where f is defined by (5.4.2) and h^F is the discount function induced by F.

Proof. By the definition of $\alpha(r)$, we have that

$$\frac{\alpha(r)-1}{r} = \frac{\frac{\sqrt{2}}{\sigma}\sqrt{r + \frac{1}{8}\sigma^2} - \frac{1}{2}}{r}$$

Note that $\frac{1}{r} = \int_0^\infty e^{-rs} ds$, then Lemma 5.4.1 yields that

$$\frac{\alpha(r)-1}{r} = \frac{\sqrt{2}}{\sigma} \int_0^\infty e^{-sr} f(s; \frac{1}{8}\sigma^2) ds - \int_0^\infty \frac{1}{2} e^{-rs} ds.$$

[2]See, for example, Bateman (1954).

Therefore, by integration by parts, we have that

$$\begin{aligned}
\int_0^\infty \frac{\alpha(r)-1}{r} dF(r) &= \int_0^\infty (\frac{\sqrt{2}}{\sigma}\int_0^\infty e^{-sr} f(s;\frac{1}{8}\sigma^2)ds - \int_0^\infty \frac{1}{2}e^{-rs}ds)dF(r) \\
&= \int_0^\infty (\frac{\sqrt{2}}{\sigma}\int_0^\infty e^{-sr} dF(r) f(s;\frac{1}{8}\sigma^2) - \frac{1}{2}\int_0^\infty e^{-rs}dF(r))ds \\
&= \int_0^\infty (\frac{\sqrt{2}}{\sigma} h^F(s) f(s;\frac{1}{8}\sigma^2) - \frac{1}{2}h^F(s))ds.
\end{aligned}$$

This completes the proof. ■

Lemma 5.4.3 *For the function $\alpha(r), r \geq 0$ defined by (5.2.1), consider discount functions h^F and h^G with weighting distributions F and G. Then*

$$\int_0^\infty \frac{\alpha(r)-1}{r} dF(r) \leq \int_0^\infty \frac{\alpha(r)-1}{r} dG(r),$$

if $h^F(t) \geq h^G(t), \forall t \geq 0$.

Proof. Due to Lemma 5.4.2, it suffices to show that

$$\frac{\sqrt{2}}{\sigma} f(s;\frac{1}{8}\sigma^2) - \frac{1}{2} \geq 0,$$

where f is defined by (5.4.2).

By simple calculation, it is easy to see that

$$\lim_{s\to\infty} \frac{\sqrt{2}}{\sigma} f(s;\frac{1}{8}\sigma^2) - \frac{1}{2} = 0.$$

We now show that f is decreasing with s. In fact, for any $C > 0$,

$$f_s(s;C) = -e^{-Cs}\frac{1}{2s\sqrt{\pi s}} < 0.$$

This yields the conclusion and completes the proof. ■

Proposition 5.4.4 (Larger discount factors imply later stopping.) *Consider discount functions h^F and h^G with weighting distributions F and G. Suppose that the corresponding triggering boundaries defined by (5.2.2) are $x_*^i, i = F, G$ respectively. Then*

$$x_*^G \geq x_*^F,$$

if $h^F(t) \leq h^G(t)$.

Proof. This is an immediate result of Lemma 4.2.3 and Lemma 5.4.3. ∎

Combined with the results on weighted discount functions obtained in Section 2, Proposition 5.4.4 readily yields a number of economically interesting results that we will state in the following.

Proposition 5.4.5 (Decreasing impatience and later stopping.) *Consider weighted discount function h^F and h^G with weighting distributions F and G and measures of decreasing impatience P_F and P_G, respectively. Let the expectation of F be no less than that of G and $P_G(t) \geq P_F(t)$ $\forall t \geq 0$. Then $x_*^G \geq x_*^F$, where $x_*^G(x_*^F)$ is the triggering boundary defined by (5.2.2) with discount function $h^G(h^F)$.*

Proof. The result is an immediate consequence of Proposition 2.4.3 and Proposition 5.4.4. ∎

Proposition 5.4.6 (More patient groups stop later.) *For weighted discount functions h^F and h^G with weighting distributions F and G we have*

$$h^G \succeq_P h^G \Longrightarrow x_*^G \geq x_*^F,$$

where $x_*^G(x_*^F)$ is the triggering boundary defined by (5.2.2) with discount function $h^G(h^F)$.

Proof. Proposition 2.4.5 yields that more patience implies a larger group discount factor. The claim thus follows from Proposition 5.4.4. ■

Proposition 5.4.7 (More diverse groups stop later.) *Suppose F and G are weighting distributions with finite support. If $G \preceq_{\infty SD} F$, then $x_*^G \geq x_*^F$, where $x_*^G(x_*^F)$ is the triggering boundary defined by (4.1.6) with the weighting distribution G (F).*

Proof. The result follows from Theorem 2.3.5 and Proposition 5.4.4. ■

The above propositions illustrate the relationship between decreasing impatience, group diversity and stopping decisions. Next we show three propositions showing that group diversity or decreasing impatience results in the non-existence of the solution to the Bellman system.

Proposition 5.4.8 (Greater discount factors and non-existence.) *Consider discount functions h^F and h^G with weighting distributions F and G. Suppose that $\int_0^\infty r dF(r) = \int_0^\infty r dG(r)$ and $h^F(t) \leq h^G(t), \forall t \geq 0$. If condition (5.2.3) holds for $G(h^G)$, then it holds for $F(h^F)$.*

Proof. This is an immediate result of Lemma 4.2.3 and Lemma 5.4.3. ■

Proposition 5.4.9 (Decreasing impatience and non-existence.) *Consider weighted discount function h^F and h^G with weighting distributions F and G and measures of decreasing impatience P_F and P_G, respectively. Let the expectation of F be no less than that of G and $P_G(t) \geq P_F(t)$ $\forall t \geq 0$. Then if condition (5.2.3) holds for $G(h^G)$, then it holds for $F(h^F)$.*

Proof. This is an immediate result of Proposition 2.4.3 and Proposition 5.4.8. ■

Proposition 5.4.10 (Group diversity and non-existence.) *Suppose F and G are weighting distributions with finite support. If $G \preceq_{\infty SD} F$, then if condition (5.2.3) holds for G, then it holds for F.*

Proof. The result follows from Theorem 2.3.5 and Proposition 5.4.8. ∎

Proposition 5.4.9 and Proposition 5.4.10 show that more decreasing impatience or more diversity of a group would lead to more risk of non-existence of the solution to the Bellman system. On one hand, for a single agent exibiting more decreasing patience, the current selve of the agent ecounters more constraint on her decisions. On the other hand, for a group with more dispersed opinions, the consensus is more difficult to reach. The two propositions tell that if the constraint on a current or the disagreement among group members is too heavy, the Bellman system and the so called "value matching" + "smooth pasting" fail to solve the equilibrium stopping problems.

Chapter 6

Conclusion

We have introduced and analyzed a new class of discount functions, the class of *weighted discount functions,* which includes all of the most commonly used discount functions. A weighted discount function corresponds to the weighted average of exponential discount factors with different discount rates. It may thus represent the discount function of a group whose members have different opinions about the appropriate discount rate. Likewise, weighted discount functions may reflect the uncertainty of a single decision maker about what discount rate to use. Following Weitzman (2001), weighted discounting comprises a natural way to cope with the discount rate dilemma.

We have shown that weighted discount functions offer a natural notion of group diversity in time preferences. We have shown that more diverse groups discount less heavily than less diverse groups. The reverse is also true. If a weighted discount function uniformly dominates another discount function, its weighting distribution must be more dispersed.

Group diversity has consequences for investment behavior. Since weighted discount functions feature decreasing impatience, they induce time-inconsistency. We provide a general approach to optimal stopping under weighted discounting within the intra-personal game framework. We proceed to study investment behavior in a real-options framework when the decision is taken by a group of people rather than by a single individual. Among other things, we show that more diverse groups hold the option to invest longer, leading to more patient investment behavior. Since most "behavioral" discount functions can be written as weighted discount functions, this paper likewise solves, in great generality, the long-standing optimal stopping problem under sophistication without commitment. Finally, we show that the so called "value matching" + "smooth pasting", which is a standard solving scheme for conventional time consistent optimal stopping problems, may not yields the desired solution for a time inconsistent stopping problem and the corresponding Bellman system may not admit any proper solution.

Chapter 7

Appendices

A Groups of non-exponential discounters

Consider an individual, i, with weighted discount function

$$h(t;i) = \int_0^\infty e^{-rt} dF(r;i)$$

where $F(r,i)$ denotes the weighting distribution used to build the discount function $h(t;i)$. A non-degenerate distribution $F(r;i)$ may reflect his uncertainty about what discount rate to use. Alternatively, a non-degenerate $F(r;i)$ may reflect individual i's bounded rationality. For example, if $F(r;i)$ is Gamma-distributed, we know that individual i is a generalized hyperbolic discounter.

We are interested in the discount function of a group of such individuals. As before, the group may consist of a finite number or a continuum of individuals. In either case, we let the indices i referring to the individuals be non-negative. Let us denote the discount function of this group by $\bar{h}(t)$. The weights of the individual group members are given by the weighting distribution $G(i)$. If the number of individuals is finite, then the probability mass function $g(i)$ of $G(i)$

gives the percentage weight of group member i, i.e., $\bar{h}(t) = \sum_i g(i)h(t;i)$. In the general case, we have

$$\bar{h}(t) = \int_0^\infty h(t;i)dG(i). \tag{A.1}$$

The following proposition shows that $\bar{h}(t)$ belongs also to the class of weighted discount functions.

Proposition A.1 *The group discount functions of weighted discounters are weighted discount functions. In particular, there exists a distribution $\bar{F}$ over exponential discount rates r such that*

$$\bar{h}(t) = \int_0^\infty e^{-rt} d\bar{F}(r).$$

Moreover, $\bar{F}$ can be computed as $\bar{F}(r) = \int_0^\infty F(r;i)dG(i)$.

Proof. The result follows from the following computation:

$$\begin{aligned}
\int_0^\infty e^{-rt} d\bar{F}(r) &= \int_0^\infty e^{-rt} d\left(\int_0^\infty F(r;i)dG(i)\right) \\
&= \int_0^\infty e^{-rt} \int_0^\infty dF(r;i)dG(i) \\
&= \int_0^\infty \int_0^\infty e^{-rt} dF(r;i)dG(i) \\
&= \bar{h}(t).
\end{aligned}$$

■

We close this section with a simple example that illustrates the above notation and analyzes the investment behavior of a group of non-exponential discounters within the real options framework of Section 4.1. Consider a group of just two members $i = 1, 2$ who receive weights $\delta \in (0, 1)$ and $1 - \delta$, respectively. Therefore, G in equation (A.1) is the weighting distribution of a pseudo-exponential discount function; see equation (2.2.1). Moreover, suppose that both members are hyperbolic discounters with the parametrization of Mazur (1987) and Harvey (1995) and parameters α_1 and α_2, respectively:

$$h(t;i) \equiv h(t;\alpha_i) = \frac{1}{1+\alpha_i t}.$$

From equation (2.2.4), these are weighted discount functions with exponentially distributed weighting distributions. That is, in the above notation, $F(r;i) \equiv F(r;\alpha_i) = 1 - e^{-\frac{r}{\alpha_i}}$ and the corresponding density functions are given by $f(r;i) \equiv f(r;\alpha_i) = \frac{1}{\alpha_i} e^{-\frac{1}{\alpha_i} r}$. Therefore, by Proposition A.1, the discount function of the group defined in equation (A.1),

$$\bar{h}(t) \equiv \bar{h}(t;\alpha_1,\alpha_2) = \delta \frac{1}{1+\alpha_1 t} + (1-\delta)\frac{1}{1+\alpha_2 t},$$

is also a weighted discount function. For that reason, our results on group investment behavior in Section 4.1 also apply to this group of behavioral investors. In order to compute the triggering boundary of this group, we need to know its weighting distribution $\bar{F}$. By Proposition A.1, its cdf is $\bar{F}(r) \equiv \bar{F}(r;\alpha_1,\alpha_2) = \delta\left(1 - e^{-\frac{r}{\alpha_1}}\right) + (1-\delta)\left(1 - e^{-\frac{r}{\alpha_1}}\right)$ and the corresponding density function is given by $\bar{f}(r) \equiv \bar{f}(r;\alpha_1,\alpha_2) = \delta \frac{1}{\alpha_1} e^{-\frac{1}{\alpha_1} r} + (1-\delta)\frac{1}{\alpha_2} e^{-\frac{1}{\alpha_2} r}$. Therefore, by equations (4.1.6) and (4.1.7), the triggering threshold is given by

$$x_* = \frac{\displaystyle\int_0^\infty \mu(r)(\delta \frac{1}{\alpha_1} e^{-\frac{1}{\alpha_1} r} + (1-\delta)\frac{1}{\alpha_2} e^{-\frac{1}{\alpha_2} r})dr}{\displaystyle\int_0^\infty \mu(r)(\delta \frac{1}{\alpha_1} e^{-\frac{1}{\alpha_1} r} + (1-\delta)\frac{1}{\alpha_2} e^{-\frac{1}{\alpha_2} r})dr - 1} I.$$

B On the well-posedness condition (4.1.2)

Proposition B.1 *If condition (4.1.2) holds, then* $\forall x \in [0, \infty)$,

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}[\int_0^\infty e^{-r\tau} dF(r)(X_\tau - I)] < \infty.$$

If $b > \inf\{r \in [0, \infty) : F(r) > 0\}$, *then* $\forall x \in \mathbb{R}^+$,

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}[\int_0^\infty e^{-r\tau} dF(r)(X_\tau - I)] = \infty.$$

Proof. For the sake of convenience, we denote $r_0 = \inf\{r \in [0, \infty) : F(r) > 0\}$. When $b < r_0$, by standard option pricing theory, we have

$$0 \le \sup_{\tau \in \mathcal{T}} \mathbb{E}[e^{-r\tau}(X_\tau - I)] \le \sup_{\tau \in \mathcal{T}} \mathbb{E}[e^{-r_0\tau}(X_\tau - I)] < \infty,$$

for $X_0 = x \in (0, \infty), r \ge r_0$. This yields that,

$$\begin{aligned}\sup_{\tau \in \mathcal{T}} \mathbb{E}[\int_0^\infty e^{-r\tau} dF(r)(X_\tau - I)] &\le \int_0^\infty \sup_{\tau \in \mathcal{T}} \mathbb{E}[e^{-r\tau}(X_\tau - I)] dF(r) \\ &\le \sup_{\tau \in \mathcal{T}} \mathbb{E}[e^{-r_0\tau}(X_\tau - I)] < \infty.\end{aligned}$$

Here, the second inequality is because of $\int_0^\infty dF(r) = 1$. When $b > r_0$, consider the case of $\tau_* = \infty$, then for any $r \ge b$, $\mathbb{E}[e^{-r\tau_*}(X_{\tau_*} - I)] \ge 0$; for any $r < b$, $\mathbb{E}[e^{-r\tau_*}(X_{\tau_*} - I)] = \infty$. Since when $b > r_0$, $F(b-) > 0$, then we have

$$\sup_{\tau \in \mathcal{T}} \mathbb{E}[\int_0^\infty e^{-r\tau} dF(r)(X_\tau - I)] \ge \int_0^b \mathbb{E}[e^{-r\tau_*}(X_{\tau_*} - I)] dF(r) = \infty.$$

This completes the proof. ■

Bibliography

Abdellaoui, M., A. Attema, and H. Bleichrodt (2010): "Intertemporal Tradeoffs For Gains and Losses: An Experimental Measurement of Discounted Utility," *Economic Journal*, 120, 845–866.

Attema, A., H. Bleichrodt, Y. Gao, Z. Huang, and P. Wakker (forthcoming): "Measuring Discounting without Measuring Utility," *American Economic Review*.

Back, K. and D. Paulsen (2009): "Open-loop equilibria and perfect competition in option exercise games," *Review of Financial Studies*, 22, 4531–4552.

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

Bateman, H. (1954): "Tables Of Intergral Transforms Volume 1," .

Bernstein, S. (1928): "Sur les fonctions absolument monotones," *Acta Mathematica*, 52, 1–66.

Bjork, T. and A. Murgoci (2010): "A general theory of Markovian time inconsistent stochastic control problems," *Available at SSRN 1694759*.

Björk, T., A. Murgoci, and X. Zhou (2014): "Mean-variance Portfolio Optimization with State Dependent Risk Aversion," *Mathematical Finance*, 24, 1–24.

Bleichrodt, H., K. Rohde, and P. Wakker (2009): "Non-hyperbolic Time Inconsistency," *Games and Economic Behavior*, 66, 27–38.

Brennan, M. and E. Schwartz (1985): "Evaluating Natural Resource investments," *Journal of Business*, 58, 135–157.

Brockett, P. L. and L. L. Golden (1987): "A Class of Utility Functions Containing All the Common Utility Functions," *Management Science*, 33, 955–964.

Caballé, J. and A. Pomansky (1996): "Mixed risk aversion," *Journal of Economic Theory*, 71, 485–513.

Caplin, A. and J. Leathy (2006): "The Recursive Approach to Time Inconsistency," *Journal of Economic Theory*, 131, 134–156.

Chiu, H. (2010): "Skewness Preference, Risk Taking and Expected Utility Maximization," *Geneva Risk and Insurance Review*, 35, 108–129.

Deck, C. and H. Schlesinger (2010): "Exploring higher order risk effects," *The Review of Economic Studies*, 77, 1403–1420.

Denuit, M. and J. Dhaene (2003): "Simple characterizations of comonotonicity and countermonotonicity by extremal correlations," *Belgian Actuarial Bulletin*, 3, 22–27.

Dixit, A. and R. Pindyck (1994): *Investment under Uncertainty*, vol. 15, Princeton University Press.

Ebert, J. and D. Prelec (2007): "The Fragility of Time: Time-Insensitivity and Valuation of the Near and Far Future," *Management Science*, 53, 1423–1438.

Ebert, S. (2015a): "Decision Making When Things Are Only a Matter of Time," Working Paper.

——— (2015b): "On Skewed Risks in Economic Models and Experiments," *Journal of Economic Behavior and Organization*, 4, 85–97.

Eeckhoudt, L. and H. Schlesinger (2006): "Putting risk in its proper place," *The American Economic Review*, 96, 280–289.

Ekeland, I. and A. Lazrak (2006): "Being Serious About Non-Commitment: Subgame Perfect Equilibrium in Continuous Time," Working Paper.

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

Falk, A., A. Becker, T. Dohmen, D. Huffman, and U. Sunde (2016): "The Nature and Predictive Power of Preferences: Global Evidence," Working Paper.

Feldstein, M. (1964): "The Social Time Preference Discount Rate in Cost-Benefit Analysis," *Economic Journal*, 74, 360–379.

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40, 351–401.

Friedman, A. (2010): *Variational principles and free-boundary problems*, Courier Corporation.

GOLLIER, C. (2014): *Pricing the Future: The Economics of Discounting and Sustainable Developement*, Princeton University Press.

GOLLIER, C. AND R. ZECKHAUSER (2005): "Aggregation of Heterogeneous Time Preferences," *Journal of Political Economy*, 113, 878–896.

GRENADIER, S. AND N. WANG (2007): "Investment under Uncertainty and Time Inconsistent Preferences," *Journal of Financial Economics*, 84, 2–39.

HARDY, G. H., J. E. LITTLEWOOD, AND G. PÓLYA (1952): *Inequalities*, Cambridge university press.

HARRIS, C. AND D. LAIBSON (2013): "Instantaneous Gratification," *Quarterly Journal of Economics*, 128, 205–248.

HARVEY, C. (1986): "Value Functions for Infinite-Period Planning," *Management Science*, 32, 1123–1139.

——— (1995): "Proportional Discounting of Future Costs and Benefits," *Mathematics of Operations Research*, 20, 381–399.

INGERSOLL, J. (1987): *Theory of Financial Decision Making*, Rowman & Littlefield, Lanham, MD.

JACKSON, M. AND L. YARIV (2015): "Collective Dynamic Choice: The Necessity of Time-Inconsistency," *American Economic Journal: Microeconomics*, 4, 150–178.

KARATZAS, I. AND S. SHREVE (2006): *Springer, 2nd printing edition*, Brownian Motion and Stochastic Calculus.

KARP, L. (2007): "Non-constant discounting in continuous time," *Journal of Economic Theory*, 132, 557–568.

KRYLOV, N. V. (2008): *Controlled diffusion processes*, vol. 14, Springer Science & Business Media.

LAIBSON, D. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112, 443–378.

LOEWENSTEIN, G. AND D. PRELEC (1992): "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics*, 37, 573–597.

MARGLIN, S. (1963): "The Social Rate of Discount and The Optimal Rate of Investment," *Quarterly Jounral of Economics*, 77, 95–111.

MAZUR, J. (1987): "Social Influences on Risk Attitudes: Applications in Economics," in *Quantitative Analyses of Behavior, Vol.5*, ed. by M. Commons, J. Mazur, J. Nevin, and H. Rachlin, Lawrence Erlbaum, Hillsdale, NJ, 55–73.

MCDONALD, R. AND D. SIEGEL (1986): "The Value of Waiting to Invest," *Quarterly Journal of Economics*, 101, 707–727.

PHELPS, E. AND R. POLLAK (1968): "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies*, 35, 185–199.

PRELEC, D. (2004): "Decreasing Impatience: A Criterion for Non-Stationary Time Preference and "Hyperbolic" Discounting," *Scandinavian Journal of Economics*, 106, 511–532.

QUAH, J. AND B. STRULOVICI (2013): "Discounting, Values, and Decisions," *Journal of Politic Economy*, 121, 896–939.

RAMSEY, F. (1928): "A Mathematical Theory of Saving," *Economic Journal*, 38, 543–559.

Rothschild, M. D. and J. E. Stiglitz (1970): “Increasing Risk: I. A Definition,” *Journal of Economic Theory*, 2, 225–243.

Samuelson, P. (1937): “A Note on Measurement of Utility,” *Review of Economic Studies*, 4, 155–161.

Strotz, R. (1955): “Myopia and Inconsistency in Dynamic Utility Maximization,” *Review of Economic Studies*, 23, 165–180.

Weitzman, M. (2001): “Gamma Discounting,” *American Economic Review*, 91, 260–271.