Weighted discounting—On group diversity, time-inconsistency, and consequences for investment

Sebastian Ebert Wei Wei Xun Yu Zhou*

February 26, 2020

Abstract

This paper presents the class of *weighted discount functions,* which contains the discount functions commonly used in economics and finance. Weighted discount functions may describe the discounting behavior of groups, uncertainty about what discount rate to use, present-biased time preferences, and all of these simultaneously. As an application, we study investment behavior under weighted discounting and come up with the following general result: Greater group diversity, greater parameter uncertainty, and more present-biased time preferences lead to delayed investment or, equivalently, more risk-taking.

Keywords: diversity, hyperbolic discounting, investment, parameter uncertainty, time inconsistency, weighted discounting

JEL codes: D81; D90; G11

*Ebert: Frankfurt School of Finance and Management, Department of Economics, Adickesallee 32-34, 60322 Frankfurt, Germany. E-mail: s.ebert@fs.de. Wei: Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, Scotland, EH14, 4AS, UK. E-mail: wei.wei@hw.ac.uk. Zhou: Department of Industrial Engineering and Operations Research, Columbia University, New York, NY 10027, USA. E-mail: xz2574@columbia.edu. An earlier version of this paper circulated under the title “Discounting, Diversity, and Investment.” We thank an associate editor, two anonymous referees, Thomas Epper, Chiaki Hara, Oliver G. Spalt, Ivo Welch, Wei Xiong, and seminar audiences at Berkeley, Frankfurt School, Columbia University, Hamburg, USI Lugano, Tilburg, Singapore Management University, the National University of Singapore, Academia Sinica, and St. Gallen for helpful comments. We also thank conference participants of the 2017 Winter Workshop on Operations Research, Finance and Mathematics in Sapporo, Japan, and of the European Finance Association Annual Meeting in Mannheim, Germany. Ebert gratefully acknowledges financial support through a Veni grant by the Dutch Science Foundation (NWO). Wei gratefully acknowledges financial support through Oxford–Nie Financial Big Data Lab. Zhou gratefully acknowledges financial support through a start-up grant at Columbia University and through Oxford–Nie Financial Big Data Lab.

1 Introduction

How do individuals, firms, and public entities make investment decisions? The standard approach to measure the attractiveness of an investment is to choose some fixed discount rate and compute its net present value (NPV), that is, the sum of its discounted expected cash flows. This paper relaxes some of the simplifying assumptions of this approach.

First, we take into account that investment decisions may be taken by a group of people, rather than by a single individual. Even if a single decision-maker (DM) 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, and friends. In modern societies, the members of a household make important decisions together. Also, presidents collect the opinions of party members, advisers—or, sometimes, their spouses.

The second issue we 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 issues in sustainable investment and cost-benefit analysis.) Weitzman's survey evidence from more than 2,000 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 of opinion about the appropriate method of discounting.

Third, in addition to group disagreement, we allow for the possibility that individuals (including those in a group) may be uncertain about what discount rate to use and/or have present-biased time preferences. As we show, there is a close connection between the discounting behavior

[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.

of groups and that of present-biased discounters (such as hyperbolic discounters). We provide a necessary and sufficient condition that clarifies the exact relationship. Therefore, we allow for *both* inter-personal disagreement and intra-personal disagreement (i.e., parameter uncertainty) about what discount rate to use. From a modeling point of view, there is no essential difference between group discounting and present-biased discounting, and thus all results in this paper may be applied to both contexts and have proper implications and interpretations.

Fourth, our model takes into account the value of timing an investment optimally, as put forward in the real options approach originally 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 given answers to the discount rate dilemma and parameter uncertainty regarding what discount rate to use. This paper makes progress regarding all three issues. Moreover, we contribute to a recent and growing literature on the importance of team diversity for financial outcomes in general.[2] To the best of our knowledge, our paper offers the first theoretical contribution that establishes a connection between team diversity and investment decisions. At the same time, our results show that existing results on investment behavior and time-inconsistency (discussed below) have a corresponding interpretation in terms of team diversity. A contribution to behavioral finance is that our analysis extends the seminal real option result from McDonald and Siegel (1986) from exponential discounting to present-biased discounting, like general hyperbolic discounting.

The research reported in this paper was initially triggered by a simple observation, being that the hyperbolic discount function can be represented as the weighted sum of a continuum of exponential discounters, with the weighting distribution being the exponential one. This

[2]As Manconi et al. (2016) point out, examples include studies on the effects of women on boards (e.g., Adams and Ferreira 2009; Ahern and Dittmar 2012; Adams 2015; and Kim and Starks 2016), studies on CEO power vis-á-vis the board (e.g., Adams et al. 2005; Fahlenbrach 2009; and Bebchuk et al. 2011), studies on the nationality of board members (e.g., Masulis et al. 2007), studies on variation in expertise and prior work history (e.g., Güner et al. 2008), and studies that combine several characteristics into an index (e.g., Giannetti and Zhao 2019; and Bernile et al. 2016). Manconi et al. (2016) show that investing in firms with diverse executive teams is highly profitable. Our paper suggests that heterogeneity in time preferences may be an explanation for why diversity matters, as this heterogeneity has important consequences for a team's investment decisions.

motivated us to define and characterize a new class of discount functions called *weighted discount functions.* It turns out that, subject to some assumptions, this class of discount functions allows for a unified approach to the above four issues (group diversity in time preference, parameter uncertainty regarding the discount rate, present-biased time preferences, and inclusion of the option value of investment). We remark that the study of investment (and the study of investment in a real options model in particular), is merely one possible application of weighted discounting. We detail this application in this paper, because it illustrates well the power of the approach to generate economic insights. Weighted discounting, however, may yield new results for, and offer new perspectives on, many other economic problems that involve a temporal dimension. These results may be of normative relevance (in the paradigms of group decisions and intra-personal uncertainty, respectively) as well as of descriptive relevance (when studying present-biased time preferences and the resulting time-inconsistency and self-control issues).

For the sake of concreteness, we illustrate the main idea of weighted discounting within the group decision paradigm and (for now) suppose that all group members are exponential discounters who are certain about what discount rate to use. Member i ($i = 1, \ldots, n$) favors discount rate r_i, and thus her discount function is given by $e^{-r_i t}$. The group's *weighting distribution* $F(r)$ assigns a weight to each member or, equivalently, to each (positive) discount rate r. Letting f denote the probability mass function of F, the group's *weighted discount function* is given by

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

The idea of averaging discount functions to model social or group preferences goes back to seminal articles such as those of Marglin (1963) and Feldstein (1964). As will be explained, weighted discounting corresponds to utilitarian aggregation with weights given by f. In Section 2, we provide a comprehensive analysis of the *class of weighted discount functions* by presenting 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^F, and vice versa. Most results in Section 2 are variations or generalizations of existing results (sometimes restricted to specific members or a subclass of the weighted discount functions) or are applications of more abstract results from the statistics literature on stochastic orders. To the best of our knowledge, however, no unified treatment of weighted discounting has been given. By overcoming various technical difficulties, our results in Section 2 apply to the large class of weighted discount functions as defined in this paper, and their proofs are short and simple.

The first main contribution of Section 2 is to introduce a notion of *group diversity* and to show how greater group diversity reflects in a group's discount function. Our notion of group diversity 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. Intuitively, greater group diversity means that there are stronger differences of opinions among its members about what discount rate to use. For example, a group is more diverse if its weighting distribution is obtained as a mean-preserving spread (Rothschild and Stiglitz 1970) of the weighting distribution of the group in comparison. While we obtain results for much more general comparisons of weighting distributions—in particular, for weighting distributions that can be compared in terms of (an arbitrary degree of stochastic) *dominance,* see Fishburn (1980)—in our interpretations we often focus on the special case of our results that concern diversity. We show that greater group diversity (a property of the weighting distribution F) unambiguously translates into a more elevated discount function h^F. That is, more diverse groups discount less heavily. Likewise, we obtain a result that shows that *more patient* groups as defined recently by Quah and Strulovici 2013) have a more elevated discount function. Moreover, we show that all weighted discount functions exhibit decreasing impatience, as defined by Prelec (2004).

The second main contribution of Section 2 is to clarify the importance and prevalence of weighted discount functions. Souzou (1998) noted that we may also think of the weighted discount function in equation (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 (that is, group) disagreement, but also intra-personal disagreement. Then a more dispersed weighting distribution corresponds to *greater parameter uncertainty* (rather than group diversity).

Now, one might wonder whether some of the commonly used discount functions can be written as weighted discount functions. That is, for a given discount function h, does there exist a distribution of opinions/a distribution reflecting uncertainty about what discount rate to use, F, such that the resulting group discount function is given by $h = h^F$? The necessary and sufficient condition for a discount function to be of the weighted form is given by a mathematical result known as Bernstein's Theorem, which roughly says that the derivatives of the discount function are alternating in sign.

The first insight derived from the application of Bernstein's Theorem is that the observation, that a present-biased discount function may be represented as a group discount function, is rather the rule than the exception. Consequently, our class of weighted discount functions is large. We show that exponential (Samuelson 1937), pseudo-exponential (Ekeland and Lazrak 2006), hyperbolic (Harvey 1986), proportional (Mazur 1987; Harvey 1995), and generalized hyperbolic (Loewenstein and Prelec 1992) discount functions are all weighted discount functions.[3] As explained in Section 2, it may even be argued that "all of the commonly used discount functions" are weighted discount functions.

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 *weighting iteration argument,* as we call it, shows that Assumption 1 in Weitzman (2001, pp. 263–264), which restricts groups to "experts thinking in terms exponential discounting," and which Weitzman refers to as "enormously simplifying," is not required. His approach to resolve the discount rate dilemma applies more generally to

[3]The extension of quasi-hyperbolic discounting (Phelps and Pollak 1968; Laibson 1997) to continuous time (Harris and Laibson 2013) are likewise weighted discount functions. Bleichrodt et al. (2009) provide an overview of discount functions used in the literature. The two new discount functions with decreasing impatience proposed in that paper are also weighted discount functions.

groups of non-exponential discounters—as long as they belong to our class of weighted discounters. Furthermore, the mathematically simple weighting iteration argument shows that weighted discounting may capture group diversity, parameter uncertainty, and present-biased time preferences at the same time. In summary, weighted discounting arises for a number of economic reasons, and most existing intertemporal models implicitly assume weighted discounting already. Our results in Section 2 thus speak to a number of applications in economics and finance.

As mentioned before, investment in a real options model—the topic of Sections 3 and 4—is merely one possible application of weighted discounting that we chose to detail in this paper. This application illustrates well that thinking in terms of weighting discounting can bring significant advancements to a research field, both mathematically and economically.

The previously mentioned result that weighted discount functions exhibit decreasing impatience constitutes the main challenge in studying a group's investment decision. Except for the trivial case in which the group consists of a single member who discounts exponentially and without any parameter uncertainty—the case in which the weighted discount function is exponential—the decision problem becomes time-inconsistent (Strotz 1955). For that reason, in our study of group investment behavior and/or parameter uncertainty, we find ourselves facing similar difficulties known from the behavioral literature on timing (or stopping) decisions under time-inconsistency (e.g., O'Donoghue and Rabin 1999). We show how insights from behavioral finance can contribute to our understanding of *rational* investment decisions that are taken collectively or under parameter uncertainty.

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 its future selves may not agree with the planned investment 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 the technical Section 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 general results hold for arbitrary payoff or utility functions, and for stopping problems that are not specific to investment, in Section 4 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 arbitrary weighted discount functions, we derive the unique threshold strategy that describes investment behavior. 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.

The investment threshold can be derived in closed form. Interestingly, it involves the weighting distribution F, but *not* the weighted discount function h^F directly. This illustrates the mathematical virtue of writing discount functions in their weighted form—even if they already admit a nice analytical representation, like hyperbolic discounting—as some economic results may involve this weighting distribution. Having solved the investment problem for the large class of weighted discount functions, we can provide the most comprehensive analysis of the importance of time preferences for investment behavior to date. Turning to the comparative statics of investment behavior (i.e., the level of the investment threshold), we show that higher investment cost, expected growth of the project value, and project value volatility all lead to a higher threshold (i.e., later investment). This result generalizes extant results for exponential discounting to arbitrary weighted discount functions. Most interestingly, we can offer general results on the impact of group diversity, parameter uncertainty, and present-bias on investment behavior.

We proceed to show that greater group diversity and/or parameter uncertainty lead to *later investment.* Waiting longer to make a decision when disagreement and/or uncertainty is strong seems consistent with daily observations. Note that later investment—as characterized by a larger investment threshold—comes along with *more risk-taking.* Rather than taking a bird in the hand (the current value of the project), a more diverse group goes for two in the bush (a potentially higher project value in the future). The main driver behind the result that greater group diversity and/or greater parameter uncertainty lead to later investment and more risk-taking is the more elevated discount factor of a more diverse group or a more uncertain investor, as established in Section 2.

Another result on the comparative statics of investment shows that weighted discount functions with greater decreasing impatience—which induces a stronger present bias—also lead to later investment and more risk-taking. This statement relies on the reasonable ceteris paribus assumption that the discount functions in comparison share the same initial rate of time preference (Prelec 2004). This point is subtle and explained in more detail later. As a concrete illustration of this matter and our results on investment behavior more generally, we offer a detailed and somewhat exhaustive treatment of the investment behavior under the seminal model of generalized hyperbolic discounting (Loewenstein and Prelec 1992). We obtain clear-cut results for how the two parameters of the generalized hyperbolic discount function—which determine the level of decreasing impatience and the initial rate of time preference, respectively—affect investment. Taken together, therefore, the unified perspective on diversity, parameter uncertainty, and present bias offered through weighted discounting yields an unambiguous result: Greater group diversity, parameter uncertainty, and/or time-inconsistency all lead to later investment and more risk-taking.

2 Weighted discount functions

In this section, we formally define the class of weighted discount functions and then introduce our notions of dominance and diversity in time preferences. While—for sake of concreteness—in our interpretations we often refer to the group decision paradigm, the discussion likewise applies to the two other settings outlined in the introduction: intra-personal uncertainty about what discount rate to use and present-biased time preferences. We also illustrate the importance and prevalence of weighted discounting.

2.1 General definition of weighted discount functions

The following definition extends equation (1) to distributions of opinions that may be continuous.

Definition 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 concentrated on $[0, \infty)$ such that*

$$h(t) = \int_0^\infty e^{-rt} dF(r). \tag{2}$$

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. Mathematically, if the weighting distribution F has a density (or probability mass) function f, the corresponding weighted discount function is the Laplace transform of that density function. If F is a simple distribution, equation (2) becomes equation (1) in the introduction.

2.2 Prevalence and examples of weighted discount functions

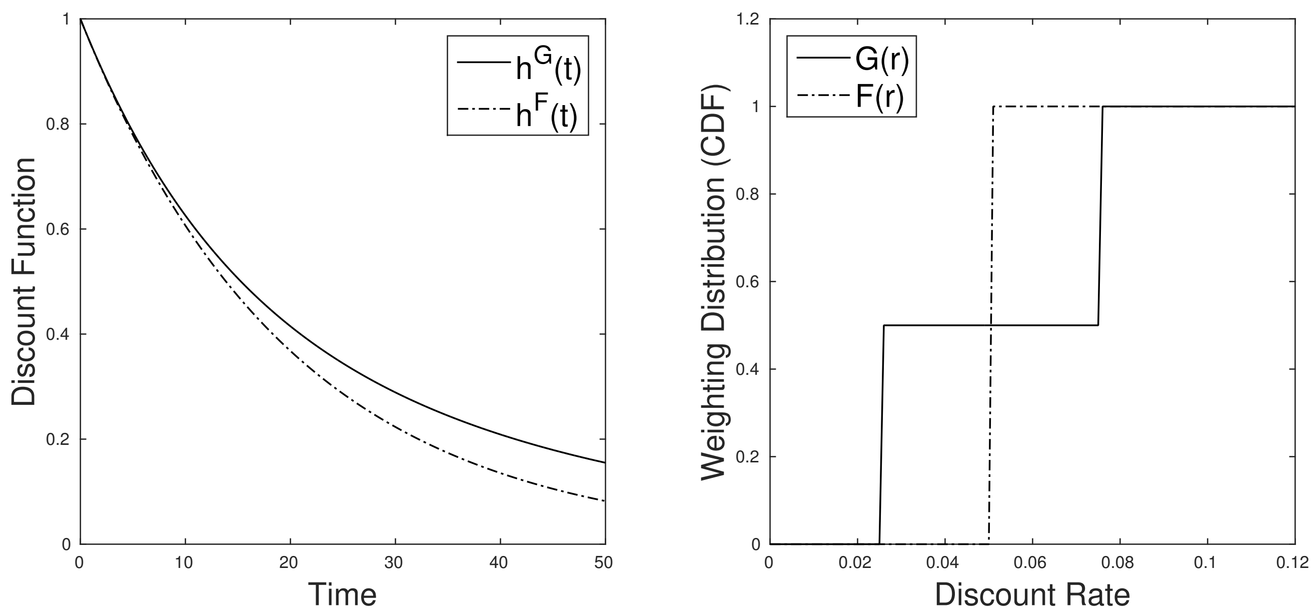
As an introductory example, consider the so-called pseudo-exponential discount function (Ekeland and Lazrak 2006; Karp 2007) given by

$$h(t) = \delta e^{-rt} + (1 - \delta) e^{-(r+\lambda)t} \text{ where } \lambda > 0, r > 0, 0 < \delta < 1, \tag{3}$$

which may describe a group with just two opinions, r and $r + \lambda$, having weights δ, and $1 - \delta$, respectively. The corresponding weighting distribution F in equation (2) is a step function and can be written as

$$F(x) = \begin{cases} 0, & x < r \\ \delta, & r \leq x < r + \lambda \\ 1, & x \geq r + \lambda. \end{cases} \tag{4}$$

Figure 1: Two discount functions and their respective weighting distribution



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 more elevated 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.

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. When $\delta = 1$, 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, that is, given by a step function with a single jump of size 1 at r. Figure 1 plots a pseudo-exponential and an exponential discount function (left panel) as well as their weighting distributions (right panel).

The following theorem is an application of Bernstein's (1928) Theorem to discount functions and provides the necessary and sufficient condition—sometimes referred to as *complete monotonicity*—for a given discount function to be a weighted discount function.[4]

[4] Hara (2008) uses the theorem to prove the existence of a representative consumer as well as to establish his result on decreasing discount rates of heterogeneous groups. Vanden (2015) applies the theorem in an asset-pricing context when deriving a stochastic discount factor.

Theorem 1 (Bernstein's Theorem Applied to Discount Functions) *A discount function h is a weighted 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 functions. 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." We may just refer to the class of weighted discount functions as the class that contains "all commonly used discount functions."[5]

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, for many weighting distributions, it may not be possible to compute the integral in Definition 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, even though the weighting distribution of the CADI function in Bleichrodt et al. (2009) may not be obtained in closed form, the results on investment behavior in Sections 3 and 4 apply and can be evaluated numerically. Likewise, the discount function obtained from a Weibull distribution may not be available in closed form, nevertheless, we can study the investment decisions of a group with Weibull-distributed opinions over discount rates.

There are also some weighting distributions that lead to explicitly expressible 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

[5]Ebert (forthcoming) characterizes the derivatives of the discount functions through risk-taking 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 (forthcoming) calls the discount functions that are decreasing with derivatives that alternate in sign—i.e., the weighted discount functions—*anti-mixed time risk averse.* Ebert (forthcoming) does not investigate the weighting representation in Definition 1, nor does he study group decisions, parameter uncertainty, or dynamic decisions such as stopping.

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).[6] In particular, the generalized hyperbolic discount function with parameters $\alpha > 0, \beta > 0$ can be written as (the first expression repeats the original definition of Loewenstein and Prelec 1992)

$$h(t) = \frac{1}{(1+\alpha t)^{\frac{\beta}{\alpha}}} = \int_0^\infty e^{-rt} f(r; \frac{\beta}{\alpha}, \alpha) dr \tag{5}$$

where

$$f(r; k, \theta) = \frac{r^{k-1} e^{-\frac{r}{\theta}}}{\theta^k \Gamma(k)} \tag{6}$$

denotes the density function of the Gamma distribution with parameters k and θ, and where $\Gamma(k)$ denotes the Gamma function evaluated at k; that is, $\Gamma(k) = \int_0^\infty x^{k-1} e^{-x} dx$. 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 (5) becomes

$$h(t) = \frac{1}{1+\alpha t} = \int_0^\infty e^{-rt} \left(\frac{1}{\alpha} e^{-\frac{1}{\alpha} r} \right) dr, \tag{7}$$

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, one can obtain the *uniform-uncertainty discount function* of Souzou (1998)

[6] Hara (2008) further elaborated on the connection between the Gamma distribution and hyperbolic discounting.

from a uniform distribution of opinions over an interval $[\underline{r}, \overline{r}]$:

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

We now present an economically important observation that results from the mathematically simple fact that any weighted average of weighted discount functions is still a weighted discount function. This implies that our results are not restricted to groups of exponential discounters. 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 what discount rate to use. By iterating the argument, it follows that the group discount function of individuals who are (i) weighted discounters and (ii) uncertain about which parameters to use is (still) a weighted discount function. We formalize this "weighting iteration argument" in Appendix B, where we also give an example.

Before concluding this section, we note that in Definition 1 and throughout this paper, the weighting distribution F is *exogenous.* This is a reasonable assumption when weighted discounting is supposed to capture intra-personal uncertainty about what discount rate to use. Similarly, when interpreting the results of this paper in the context of present-biased discounting, the weighting distribution is naturally exogenous in the sense that it is implicit from the discount function being assumed. In the group decision context, however, the assumption of an exogenous weighting distribution necessitates that the aggregation rule used in forming the weighted discount function is known and specified *a priori.* This assumption of our paper also underlies that of Weitzman (2001), who infers the weighting distribution from a histogram of surveyed opinions. An a priori determined weighting distribution seems appropriate for social cost-benefit analysis (e.g., investment in environmental projects). In the general case of group decisions in firms, however, it may be more realistic to endogenize the weighting distribution by considering more complex aggregation rules. Jouini and Napp (2007) and Hara (2008) consider complete competitive markets in which all DMs have exponential but heterogeneous discount functions. One of their results is that the consensus belief is, in fact, a weighted discount function if all the DM's utility functions are logarithmic. Moreover, in a simple setting, Jouini et al. (2010, Proposition 5.1) derive an aggregation rule explicitly. Recently, motivated by the works of Koopmans (1960) and Weitzman (2001), Chambers and Echenique (2018) develop an axiomatic approach to time preference aggregation and refer to Weitzman's (and thus our) approach as utilitarian.[7] Naturally, it is interesting to extend the analysis of the present paper to more complex aggregation rules that endogenize the weighting distribution.

2.3 Dominance and group diversity

In this subsection, we establish a notion of comparative group diversity in time preferences. Greater group diversity will be captured by a more dispersed weighting distribution G of the corresponding weighted discount function h^G. We show that a more diverse group has a more elevated discount factor (i.e., it discounts less heavily at any future time) than a less diverse group. In fact, we obtain a much more general result that says that less "dominant" weighted distributions—which include the more diverse ones—have a more elevated discount factor.

Recall that, in the expected utility model, a second-order stochastically dominating distribution is preferred, for example, by all DMs whose utility functions are increasing and concave. Roughly speaking, these DMs dislike decreases in mean and increases in variance. Accordingly, less second-order stochastically dominant distributions have lower mean and a higher variance than the distribution in comparison. The following, more general stochastic dominance notion is in the spirit of Fishburn (1980).[8]

Definition 2 ((Infinity-stochastic) dominance.) *Let F and G be weighting distributions of weighted discount functions h^F and h^G, respectively. We say that h^G (G) is* less (infinity-

[7]Consider $i = 1, 2, \ldots, n$ exponential discounters with discount rates r_i who experience utility $u(x)$ from an outcome x that is received at time t. Utilitarian aggregation with importance weights w_i yields aggregate utility $\sum_{i=1}^{n} w_i e^{-r_i t} u(x) = \left(\sum w_i e^{-r_i t}\right) u(x) := h(t)u(x)$, and h is a indeed a weighted discount function. In particular, the weighting distribution defining h is given by the probability mass function $f(r_i) := w_i$ (cf. equation (1)).

[8]Our definition is technically different from Fishburn's, in that his definition is based on the so-called iterated integrals. In the case of bounded supports of the distributions in comparison, the definitions are equivalent.

stochastically) dominant *than* h^F *(F), denoted as* $G \preceq_{\infty SD} F$*, if* $\int_0^\infty u(x)dF(x) \geq \int_0^\infty u(x)dG(x)$*, for all integrable functions* u *whose derivatives alternate in sign; that is,* $\operatorname{sgn} u^{(n)}(x) = (-1)^{n+1}$ *for any* $n \in \mathbb{N}^+, x \in (0, \infty)$.

A distribution that is less (infinity-stochastically) dominant has a lower mean, or equal mean and higher variance, or equal mean and variance and lower skewness ... as the distribution in comparison. Dominance is the "weakest" of the well-known stochastic dominance orders in the sense that any finite stochastic dominance order (i.e., first-order stochastic dominance, second-order stochastic dominance, and so on) implies it.

The generality of the dominance ordering can make the interpretations of results on less dominant groups unspecific and difficult.[9] In particular, dominance may—but need not—rank distributions by their dispersion. As an example, the weighting distribution that yields 0.02 and 0.04 with equal probability is more (less) dominant than the less dispersed weighting distribution that yields 0.01 (0.05) for certain. Consequently, the dominance order is too general for the purpose of ranking groups according to their diversity. Definition 3 below, which is inspired by Ekern's (1980) concept of more nth-degree risk, imposes restrictions on the definition of dominance so that it can be interpreted in terms of diversity. For $n = 2$, Definition 3 says that G is a mean-preserving spread of F (thus having equal mean and larger variance than F; see Rothschild and Stiglitz 1970), and for $n = 4$, it says that G is an outer risk increase of F (thus having equal first three moments and larger kurtosis than F; see Menezes and Wang 2005).[10]

Definition 3 (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 there exists an even number* $n \in \mathbb{N}$ *such that* $\int_0^\infty u(x)dF(x) \geq \int_0^\infty u(x)dG(x)$*, for all integrable functions* u *whose* n*th derivative is negative for any* $x \in (0, \infty)$.

[9] This is different from the case of ranking monetary risks, which is commonly done using a stochastic dominance order. In that case, a more dominant distribution is often regarded as being preferable, and thus speaking of "dominance" is intuitive. At least in the context of this paper, we find it difficult to argue for a distribution over discount rates to be "preferable" over another one.

[10] One may argue that also other special cases of dominance than those described in Definition 3 could be interpreted in terms of diversity. Definition 3 could thus be considered a conservative notion of diversity. Our mathematical results throughout apply to the general notion of dominance (see Fact 1), and thus the precise choice of the diversity definition is merely of relevance as concerns their interpretation.

Comparing Definitions 2 and 3 immediately yields the following:

Fact 1 (More diverse groups are less dominant.) *If weighting distribution G is more diverse than weighting distribution F, then G is less dominant than F.*

Consequently, if we prove a result on the behavior of less dominant groups, as a special case, we obtain a result on the behavior of more diverse groups. The following result, a variation of a result in Fishburn (1980), shows that dominance is the weakest order that allows for a complete characterization on the ordering of the group discount factors.[11]

Theorem 2 (Less dominance is equivalent to a more elevated discount factor.) *Suppose F and G are weighting distributions and the corresponding discount functions are given by h^G and h^F, respectively. Then*

$$G \preceq_{\infty SD} F \text{ if and only if } h^G(t) \geq h^F(t), \forall t \geq 0. \tag{8}$$

The latter property of a *more elevated discount function* is central to many results of this paper. Figure 1 above illustrates Theorem 2 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 less dominant than F. In particular, G is more diverse than F as the binary distribution G in Figure 1 constitutes a mean-preserving spread 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 5% results in a discount factor larger than a single exponential discount factor with a rate of 5%.

An important, though trivial, special case of Theorem 2 is when both F and G are degenerate so that h^F and h^G are exponential discount functions. The equivalence (8) then says that group G has a smaller discount rate if and only if it has a larger discount factor. In other words, in

[11] Mathematically equivalent results for mixed risk averse *utility* functions (i.e., increasing utility functions with derivatives that alternate in sign) are found in Brockett and Golden (1987) and Thistle (1993), even though Thistle's result is based on Fishburn's definition of infinity stochastic dominance and only holds for distributions with finite support. Since the weighting distributions of many important discount functions (like the hyperbolic) have infinite support, the assumption of a finite support would be too restrictive for our purposes.

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 establishes the necessary and sufficient condition for ordering discount factors: Ordering discount factors is equivalent to comparing their weighting distributions in the dominance sense.

2.4 Decreasing impatience and other properties of weighted discount functions

In this section, we relate dominance (equivalently, the property of an elevated discount function) to two well-known and important concepts in the literature on time preferences. In particular, weighted discount functions allow for new insights regarding the notions of decreasing impatience and patience.

Definition 4 (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 shows that weighted discounting necessitates decreasing impatience in the sense of Prelec (2004).[12]

Proposition 1 (Weighted discount functions imply decreasing impatience.) *All weighted discount functions imply decreasing impatience.*

[12]This result generalizes and refines Weitzman's (2001) observation on time-varying discount rates, which is restricted to Gamma-distributed opinions F. Jackson and Yariv (2014, 2015) show that with any heterogeneity in time preferences, utilitarian aggregation necessitates a present bias. Gollier and Zeckhauser (2005) and Jouini and Napp (2007) have noted that heterogeneity results in time-varying discount rates and time-inconsistency, and Hara (2008) shows that the rate of pure time preference is decreasing under heterogeneity. Unlike these papers, we provide results on the classic concept of decreasing impatience as originally characterized by Prelec (2004), which is not cited in the other articles. Hara (2008, Example 2), however, proves a result in the spirit of Proposition 1 in the context of hyperbolic discounting.

It is well known that non-exponential discounting induces time-inconsistency. Proposition 1 clarifies that, more specifically, (non-exponential) weighted discounters are time-inconsistent due to decreasing impatience. The assumption of weighted discounting thus imposes a restriction on the type of time-inconsistency considered—the one consistent with present bias and the majority of empirical evidence. Yet, it should be kept in mind that it is not possible to study other causes of time-inconsistency within the realm of weighted discounting (in particular, increasing impatience and future bias).

Before presenting a relationship between decreasing impatience and dominance in Proposition 2 below, we give an important interpretation of the mean of a weighting distribution, that is, of the weighted mean of the discount rates employed by a group. It is easily shown that (see also the proof of Proposition 2) for a weighted discount function h^F we have

$$\int_0^\infty r dF(r) = -\frac{(h^F)'(0)}{h^F(0)} = -(h^F)'(0). \tag{9}$$

The middle expression in (9) is, by definition (cf. Prelec 2004, p. 524), the group's *initial rate of time preference;* that is, the rate of time preference at time zero. Therefore, the expectation of a weighting distribution F (the average discount rate used) equals the initial rate of time preference of the corresponding discount function h^F (indicating how quickly the discount factor declines at time zero). Note that present bias is caused by strictly decreasing impatience—regardless of the initial rate of time preference. In order to obtain a comparative statics (i.e., a ceteris paribus) result with respect to present *bias,* therefore, we need to keep the *present* (i.e., the initial rate of time preference) constant. We obtain the following result.

Proposition 2 (Greater present bias implies less dominance.) *Consider weighted discount functions h^F and h^G with equal-mean weighting distributions F and G and measures of decreasing impatience P_F and P_G, respectively. If $P_G(t) \geq P_F(t)$ for all $t \geq 0$, then G is less dominant than F.*

The proof in the appendix has the following graphical intuition. If G (i) has a lower average discount rate (i.e., has a discount function that decreases less at time zero) and (ii) has impatience decreasing more strongly at any time (i.e., is more convex in the sense of Prelec's measure), then h^G must always lie above h^F. By Theorem 2, then, G is dominated by F.

Finally, we clarify how our notion of greater diversity relates to that of greater patience, as recently put forward by (Quah and Strulovici 2013).

Definition 5 (Comparative patience (Quah and Strulovici 2013).) *The discount function h^G exhibits more patience than another discount function h^F if*

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

Quah and Strulovici (2013, Proposition 1) provide the following choice-based characterization of greater patience.

Proposition 3 (Choice-based characterization of greater patience (Quah and Strulovici 2013)). *Two discounted payoff maximizers with (not necessarily weighted) discount functions $h^F(t)$ and $h^G(t)$ are choosing between receiving fixed amounts π_1 at time t_1 and π_2 at time t_2, where $\pi_1 > 0$, $\pi_2 > 0$, and $t_2 > t_1 \geq 0$. Then h^G exhibits more patience than h^F if and only if whenever F chooses π_2, then G also chooses π_2.*

The following proposition shows that our results on the behavior of less dominant groups hold, in particular, for more patient groups.[13]

Proposition 4 (More patient groups are less dominant.) *More patient groups are less dominant.*

The converse is not true.[14] For the intuition, by Theorem 2, group G being less dominant than

[13] Proposition 4 coincides with a result from the statistics literature, saying that the Laplace ratio order induces the Laplace order; see, for example, Shaked and Shantikumar (2007, Theorem 5.B.10). We are not aware of any article applying this result, or assigning meaning to it, in the context of time preferences.

[14] A concrete counterexample is given by $h^G(t) = 0.5\exp(-at) + 0.5\exp(-bt)$ and $h^F(t) = 0.5\exp(-at) + 0.5\exp(-ct)$, where $0 < a < b < c$. It is straightforward to verify that $h^G(t) \geq h^F(t)$ so that, by Theorem 2, group G is less dominant than F. Moreover, $h^G(0)/h^GF0)$, $h^G(t) > h^F(t)$ for $t > 0$, and $\lim_{t\to\infty} h^F(t)/h^G(t) = 1$. Thus, $h^G(t)/h^F(t)$ cannot be increasing in t.

F means that

$$h^G(t) \geq h^F(t) \Longleftrightarrow \frac{h^G(t)}{h^F(t)} \geq 1 \Longleftrightarrow \frac{h^G(t)}{h^F(t)} \geq \frac{h^G(0)}{h^F(0)}. \tag{10}$$

Therefore, when G is less dominant, the ratio of discount factors in Definition 5 need not increase, but it must be smallest at time zero. Similarly, greater diversity would require the prediction for choice outlined in Proposition 3 to hold only for $t_1 = 0$ and $t_2 = t$ rather than for all $0 \leq t_1 < t_2$, illustrating that less dominance is a weaker requirement than (and thus a consequence of) greater patience.

3 Equilibrium stopping under weighted discounting: General results

We believe that writing discount functions in their weighted form may yield new perspectives on numerous economic (and associated mathematical) problems. The remainder of the paper is concerned with illustrating this point for one particular class of dynamic decision problems. In this section, we provide a general result on stopping decisions under weighted discounting. Section 4 will make use of this result to obtain new economic insights into the well-known real options investment problem.

We assume that the group whose stopping or timing decision we study is sophisticated without commitment. Sophistication refers to the assumption that the group—unless it is given by a single person—is aware of the time-inconsistency that results from decreasing impatience (Proposition 1). The assumption of no commitment means that the group, even though anticipating the desire to change its decision later, lacks the self-control to go through with the initially optimal decision once its preferences have changed due to time-inconsistency. As in the behavioral literature (e.g., O'Donoghue and Rabin 1999), the group's stopping decision is modeled as an intra-personal game. Following Strotz (1955) and many others since, equilibrium behavior is obtained as the solution to a system of Bellman equations. The main result of this section, Theorem 3, specifies such a Bellman system for an arbitrary weighted discount function.

Readers not interested in further technical details (in particular, in the precise equilibrium definition and the Bellman system characterizing equilibrium behavior in a continuous time stopping problem under weighted discounting), may skip the remainder of this section and proceed with Section 4.

3.1 Dynamics

Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a complete probability space that supports a standard Brownian motion $(W_t)_{t\geq 0}$ with its natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Let $X = \{X_t\}_{t\geq 0}$ denote the payoff value process, and suppose that its dynamics are given by

$$\frac{dX_t}{X_t} = b(X_t)dt + \sigma(X_t)dW_t, \; X_0 = x_0 \tag{11}$$

where the bounded function b describes the instantaneous conditional expected percentage change in X per unit of time and the bounded function σ is the instantaneous conditional standard deviation per unit time.[15] X may describe the accumulated wealth when gambling in a casino, the project value in a real options setting, or the value of a stock or of the underlying of some derivative. The payoff of investment at time t is given by some payoff function $G : [0, \infty) \to \mathbb{R}$. In the next section, we will focus on $G(X_t) = X_t - I$ where $I \in \mathbb{R}^+$ describes the investment cost. The results in this section, however, are derived for general payoff (or utility) functions G.

3.2 Stopping rules and equilibrium

We first define stopping rules and the corresponding (induced) stopping times.

Definition 6 *A* stopping rule *is a function of time and the process value,* $u : [0, \infty) \times [0, \infty) \to \{0, 1\}$*, where* 0 *indicates "continue" and* 1 *indicates "stop." For any time* $t \geq 0$*, each stopping*

[15]See Karatzas and Shreve (2006) for the usual technical conditions that need to be imposed on b and σ. The results in this paper can be generalized to more complicated processes.

rule u defines a stopping time τ_u^t *after t, via*

$$\tau_u^t = \inf\{s \geq t, u(s, X_s) = 1\}. \tag{12}$$

Let h denote a given weighted discount function with corresponding weighting distribution F. Consider some time $t \geq 0$ and refer to it as "self t" of the group. We assume that the discount function of self t is given by $h_t(s) := h(s-t)$, $s \geq t$. This means that self t treats calendar date t as the present, which is also reflected by the fact that $h_t(t) = 1$. Moreover, the specification that $h_t(s) = h(s-t)$ means that group preferences are time-invariant (e.g., Halevy 2015) and, in particular, that the weighting distribution F does not depend on calendar time.[16] If $X_t = x$, self t seeks to maximize the weighted discounted payoff from its investment decision according to the stopping rule u :

$$J(t, x; u) = \mathbb{E}[h(\tau_u^t - t) G(X_{\tau_u^t}) | X_t = x]. \tag{13}$$

Since h exhibits decreasing impatience (Proposition 1), the preferences of the selves change over time. In particular, self 0 is more patient at time t than self t is at time t. This may lead to the selves preferring a different choice of u. In general, each self t can only choose current (time-t) behavior, that is, $u(t, x)$. Plans *when* to stop (i.e., stopping times) made by self t may be overthrown by a group's future self $s > t$.

Given a stopping rule u and a self $t \geq 0$, we can define the following stopping rule to be used from t on:

$$u^{\epsilon,a}(s, x) = \begin{cases} u(s, x), & \text{if } s \in [t+\epsilon, \infty) \\ a & \text{if } s \in [t, t+\epsilon), \end{cases} \tag{14}$$

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

[16] The point of this assumption is to illustrate that time-inconsistency of the group arises even if the (non-degenerate) distribution of opinions about what discount rate to use, F, does not change with time. It would be interesting to study decisions for non-constant, possibly endogenous dynamics of F. In this first treatment of stopping under weighted discounting, we focus on the simplest case in which F is independent of calender time.

constantly 0 or constantly 1. We call $u^{\epsilon,a}$ the (ϵ, a)-*deviation* from u.

Given the infinite horizon and the stationarity of the process X, we need to consider only *stationary* stopping rules u, which are functions of the state variable x only. To see this, let us rewrite the objective functional (13) of the group as

$$\begin{aligned} J(t, x; u) &= \mathbb{E}[h(\tau_u^t - t)G(X_{\tau_u^t})|X_t = x] \\ &= \mathbb{E}[h(\tau_u^0)G(X_{\tau_u^0})|X_0 = x] \\ &= J(0, x; u). \end{aligned} \tag{15}$$

Hence, each self t faces the same decision problem, which only depends on the current state $X_t = x$ but not on time t directly. We can thus identify group self t by the current state $X_t = x$ of the process and drop the time index from its objective functional

$$J(x; u) = \mathbb{E}[h(\tau_u)G(X_{\tau_u})] \tag{16}$$

where $\tau_u = \inf\{s \geq 0, u(X_s) = 1\}$. The stopping rule u in equation (16) is now stationary, namely it is independent of time t. The argument above thus shows that any given non-stationary stopping rule has a unique stationary version such that the equality (15) holds. With a slight abuse of notation, we still denote this stationary stopping rule as u, which is now a function of the process value x alone. The sophisticated group anticipates the disagreement between its current and future selves. Therefore, it searches for a stopping rule $\hat{u}$ that all possible future selves x are willing to go through with; that is, no future self x wishes to deviate from $\hat{u}$. In other words, the group plays a game with its future selves, and behavior is described by the equilibrium of that game.[17]

[17] Definition 7 is consistent with the equilibrium definition studied in time-inconsistent control problems (see Ekeland and Pirvu 2008; Björk et al. 2014) when interpreting a stopping rule as a binary control. Moreover, Definition 7 is in line with Strotz's (1955) "strategy of consistent planning," nowadays better known as an "intra-personal equilibrium." More specifically, the Strotzian solution stipulates that for any given self t : (i) all the future selves will commit to the strategy and (ii) any deviation from the strategy will make the deviating self t worse off. In a continuous-time setting, however, any fixed t (or, equivalently, a state x in the setting of this paper) *alone* has no influence whatsoever on the final payoff function because it has a measure of zero. Therefore, in order to capture the condition (ii) above in continuous time, one considers a small "alliance" of self t : the interval $[t, t+\epsilon)$. The (ϵ, a)-deviation, $u^{\epsilon,a}$, is, then, a strategy for which the alliance $[t, t+\epsilon)$ deviates to a while all

Definition 7 (Equilibrium stopping rule) *The stopping rule $\hat{u}$ is an equilibrium stopping rule if*

$$\liminf_{\epsilon\to 0} \frac{J(x;\hat{u}) - J(x;u^{\epsilon,a})}{\epsilon} \geq 0 \tag{17}$$

where $u^{\epsilon,a}$ is the (ϵ,a)-deviation from $\hat{u}$ with $\epsilon > 0$ and $a \in \{0,1\}$.

Note that $u^{\epsilon,a}$ is well defined only when it, together with $\hat{u}$, is non-stationary; see (14). However, we have explained earlier that each non-stationary stopping rule has a unique stationary version, due to the infinite horizon and the stationarity of the process X. Hence, in the above, both $\hat{u}$ and $u^{\epsilon,a}$ are understood to be the stationary versions. Intuitively, condition (17) says that every future self (characterized by wealth x) prefers $\hat{u}$ over deviating according to $u^{\epsilon,1}$ or $u^{\epsilon,0}$ for the short time ϵ during which it has decision power.

3.3 Equilibrium characterization

We now present the main result of this section, a general method to find equilibrium stopping rules and the corresponding stopping times under weighted discounting.[18]

Theorem 3 (Equilibrium Characterization) *Consider the performance functional (16) with weighted discount function $h(t) = \int_0^\infty e^{-rt}dF(r)$, a stopping rule $\hat{u}$, and functions $w(x;r) = \mathbb{E}[e^{-r\tau_{\hat{u}}}G(X_{\tau_{\hat{u}}})]$ and $V(x) = \int_0^\infty w(x;r)dF(r)$. Let $\mathcal{A} = \frac{1}{2}\sigma(x)^2x^2\frac{\partial^2}{\partial x^2} + b(x)x\frac{\partial}{\partial x}$ and suppose that $(V, w, \hat{u})$ solves*

$$\max\{\mathcal{A}V(x) - \int_0^\infty rw(x;r)dF(r), G(x) - V(x)\} = 0, \tag{18}$$

$$\hat{u}(x) = \begin{cases} 1 & \textit{if } V(x) = G(x) \\ 0 & \textit{otherwise,} \end{cases} \tag{19}$$

subject to $V(0) = \max\{G(0), 0\}$. Then $\hat{u}$ is an equilibrium stopping rule, and the value function of the problem is given by $V(x)$; that is, $V(x) = \mathbb{E}[h(\tau_{\hat{u}})G(X_{\tau_{\hat{u}}})]$.

selves beyond $t + \epsilon$ stay committed to $\hat{u}$. Definition 7 posits that such a "deviation-in-alliance" from equilibrium is worse in a first-order sense.

[18] The value function V in the theorem must satisfy some regularity conditions that are not restrictive from an economic point of view.

To interpret Theorem 3, let us call

$$\mathcal{S} = \{x \in [0, \infty) : V(x) = G(x)\} \text{ and } \mathcal{C} = \{x \in [0, \infty) : V(x) > G(x)\}$$

the *stopping region* and *continuation region* of the stopping problem, respectively. Then, the equilibrium stopping rule $\hat{u}$ can be written as

$$\hat{u}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{S} \\ 0 & \text{if } x \in \mathcal{C}, \end{cases} \tag{20}$$

and the corresponding stopping time given by

$$\tau_{\hat{u}} = \inf\{s \geq 0 : X_s \in \mathcal{S}\},$$

with $X_0 = x$.

Equations (18)–(19) constitute the so-called Bellman system, a system of coupled equations. Note that (19) is an *equation* (rather than a definition of $\hat{u}$) since its right-hand side involves V, which in turn depends on $\hat{u}$ through $w(x; r)$. Therefore, the equilibrium stopping rule $\hat{u}$ is *part* of the solution of the Bellman system.

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 tells us how we can obtain an equilibrium stopping rule $\hat{u}$ or, equivalently, the values of $x \in \mathcal{S}$, where the group will stop, and the values $x \in \mathcal{C}$, where it will continue. The function $w(x; r) = \mathbb{E}[e^{-r\tau_{\hat{u}}} G(X_{\tau_{\hat{u}}})]$ 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 $V(x) = w(x; r)$; that is, the value function V is given by that member's expected discounted payoff and equation (18) becomes the well-known Bellman equation (e.g., Dixit and Pindyck 1994): $\max\{\mathcal{A}V(x) - rV, G(x) - V(x)\} = 0$.[19]

[19] Note that the equation above is *independent* of $\hat{u}$ and can be solved once the model primitives are given. Then,

Now consider the case of several group members. Along the equilibrium stopping rule $\hat{u}$, the value function $V(x)$ can be written as

$$\begin{aligned} V(x) &= \mathbb{E}[h(\tau_{\hat{u}})G(X_{\tau_{\hat{u}}})] = \mathbb{E}[\textstyle\int_0^\infty e^{-r\tau_{\hat{u}}}dF(r)G(X_{\tau_{\hat{u}}})] \\ &= \textstyle\int_0^\infty \mathbb{E}[e^{-r\tau_{\hat{u}}}G(X_{\tau_{\hat{u}}})]dF(r) \\ &= \textstyle\int_0^\infty w(x;r)dF(r). \end{aligned} \tag{21}$$

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 (18), 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

$$\mathcal{A}w(x;r) - rw(x;r) = 0, x \in \mathcal{C}$$

with the boundary conditions $w(0;r) = 0$ and $w(x;r)|_{\mathcal{S}} = G(x)|_{\mathcal{S}}$. 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; 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

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

Equation (18) is then obtained by comparing the value of continuation and stopping.

$\hat{u}$ is obtained immediately from equation (19). In other words, in that case, equation (19) 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.

4 Investment behavior under weighted discounting

This section applies the general stopping result from Section 3 to solve the the standard irreversible investment problem of Brennan and Schwartz (1985), McDonald and Siegel (1986), and Dixit and Pindyck (1994) for general time preferences as given by an arbitrary weighted discount function. After generalizing well-known comparative statics results from exponential to weighted discounting, we present new results on the impact of group diversity, parameter uncertainty, and present bias on investment behavior.

4.1 Economic setup

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

$$\frac{dX_t}{X_t} = bdt + \sigma dW_t. \tag{22}$$

Investment in X can be made at any time t at cost I so that $G(x) = x - I$. The performance functional of the group, whose time preferences are described by a weighted discount function h—being obtained from some weighting distribution F—is then given by

$$J(x; u) = \mathbb{E}\left[\int_0^\infty e^{-r\tau_u} dF(r)(X_{\tau_u} - I)\right]. \tag{23}$$

To ensure the well-posedness of the problem, let

$$b < \inf\{r \in [0, \infty) : F(r) > 0\}. \tag{24}$$

When F is a step function jumping at $r = r_0$ (the exponential discounting case), condition (24) is reduced to the standard condition $b < r_0$ (e.g., Dixit and Pindyck 1994, p.141).[20]

[20]Economically, the more general condition (24) 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 not explode. In Appendix C, we provide a formal result on the sufficiency and necessity of the well-posedness

4.2 The decision to invest: The investment threshold under weighted discounting

By solving the Bellman system described in Theorem 3 explicitly for the real option setting specified in Section 4.1, we obtain the following result.

Proposition 5 (Investment behavior under weighted discounting.) *Consider a weighted discounter with weighting distribution F. Investing once the process X hits the* investment threshold x_* *given by*

$$x_* = \frac{\int_0^\infty \mu(r) dF(r)}{\int_0^\infty \mu(r) dF(r) - 1} I, \tag{25}$$

where

$$\mu(r) = \frac{-(b - \frac{1}{2}\sigma^2) + \sqrt{(b - \frac{1}{2}\sigma^2)^2 + 2\sigma^2 r}}{\sigma^2}, \tag{26}$$

constitutes the unique one-sided threshold equilibrium strategy.

Note that the expression for the investment threshold, equation (25), involves the weighting distribution F, but *not* the corresponding discount function h^F directly. Therefore, if one studied investment under, say, generalized hyperbolic discounting, without involving the weighted representation, it would seem very difficult to arrive at the solution. Recalling the notation from equation (5)), equation (25) becomes

$$x_* = \frac{\displaystyle\int_0^\infty \mu(r) \frac{r^{\frac{\beta}{\alpha}-1} e^{-\frac{r}{\alpha}}}{\alpha^{\frac{\beta}{\alpha}} \Gamma(\frac{\beta}{\alpha})} dr}{\displaystyle\int_0^\infty \mu(r) \frac{r^{\frac{\beta}{\alpha}-1} e^{-\frac{r}{\alpha}}}{\alpha^{\frac{\beta}{\alpha}} \Gamma(\frac{\beta}{\alpha})} dr - 1} I. \tag{27}$$

More generally, the fact that the explicit solution to a problem may involve the weighting distribution rather than the weighted discount function illustrates a possible virtue of writing discount functions in their weighted form, even if the resulting expression looks complicated at first sight.

Let us also compare our general result for weighted discounting to the standard result for exponential discounting. The investment threshold for exponential discounting is easily recovered

condition (24).

from equation (25) by letting $h(t) = e^{-r_0 t}$, so that we obtain the well-known

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

4.3 The impact of the economic environment on investment behavior

Before we turn to the impact of discounting behavior on investment, we note that the investment threshold x_* in equation (25) depends on the project return rate and volatility (through the function $\mu(r)$ in equation (26)), as well as on the entry cost I. The following proposition generalizes well-known results from the real options literature under standard, exponential discounting, to weighted discount functions.

Proposition 6 (Comparative statics for entry cost and project dynamics.) *Suppose that $\int_0^\infty \mu(r) dF(r)$ is finite.*[21] *Then*

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

(ii) x_* increases with the entry cost I, and

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

In the remainder of this section, we conduct an in-depth analysis of how differences in discounting impact investment behavior.

4.4 The impact of dominance on investment behavior

The following result clarifies the impact of the group decomposition (i.e., the impact of the weighting distribution F) on the investment decision. In particular, it shows that less dominance (and thus greater diversity in particular) leads to later investment.

[21] If not, then $x_* = I$, which means that the group invests in the project whenever the payoff process is greater than the entry cost.

Proposition 7 (Less dominant groups invest later.) *Suppose F and G are weighting distributions. Then*

$$G \preceq_{\infty SD} F \Longrightarrow x_*^G \geq x_*^F,$$

where $x_*^G (x_*^F)$ is the investment threshold defined by equation (25), with the weighting distribution G (F).

Less dominance is equivalent to a more elevated discount factor (Theorem 2). Theorem 2 and Proposition 7 thus imply the following:

Corollary 1 (More elevated discount factors imply later investment.) *Consider weighted discount functions h^F and h^G with weighting distributions F and G. Then*

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

where $x_*^G (x_*^F)$ is the investment threshold defined by equation (25), with weighting distribution $G(F)$.

Note that Proposition 7 makes an assumption on the weighting distribution F, while Corollary 1 makes an assumption on the discount function h^F, and both propositions' implication is on the investment threshold x_* in equation (25). Because x_* is expressed in terms of the weighting distribution F rather than in terms of the corresponding weighted discount function h^F, the proof of Proposition 7 is relatively simple, while a *direct* proof of Corollary 1 is quite complicated. We present the direct proof in the appendix, because a comparison with the simple proof of Proposition 7 illustrates the mathematical virtue of writing discount functions in the weighted form. Moreover, the more complicated proof contains an additional result that we state separately. Lemma 2 in the appendix offers an alternative and less intuitive, semi-analytical expression for x_* in equation (25), which involves the original discount function h^F rather than the corresponding weighting distribution F.

We now illustrate the quantitative impact of dominance and group diversity on the investment thresholds through an example. To this end, let us consider the pseudo-exponential discount

Table 1: Parameters used for the scenarios in Figure 2

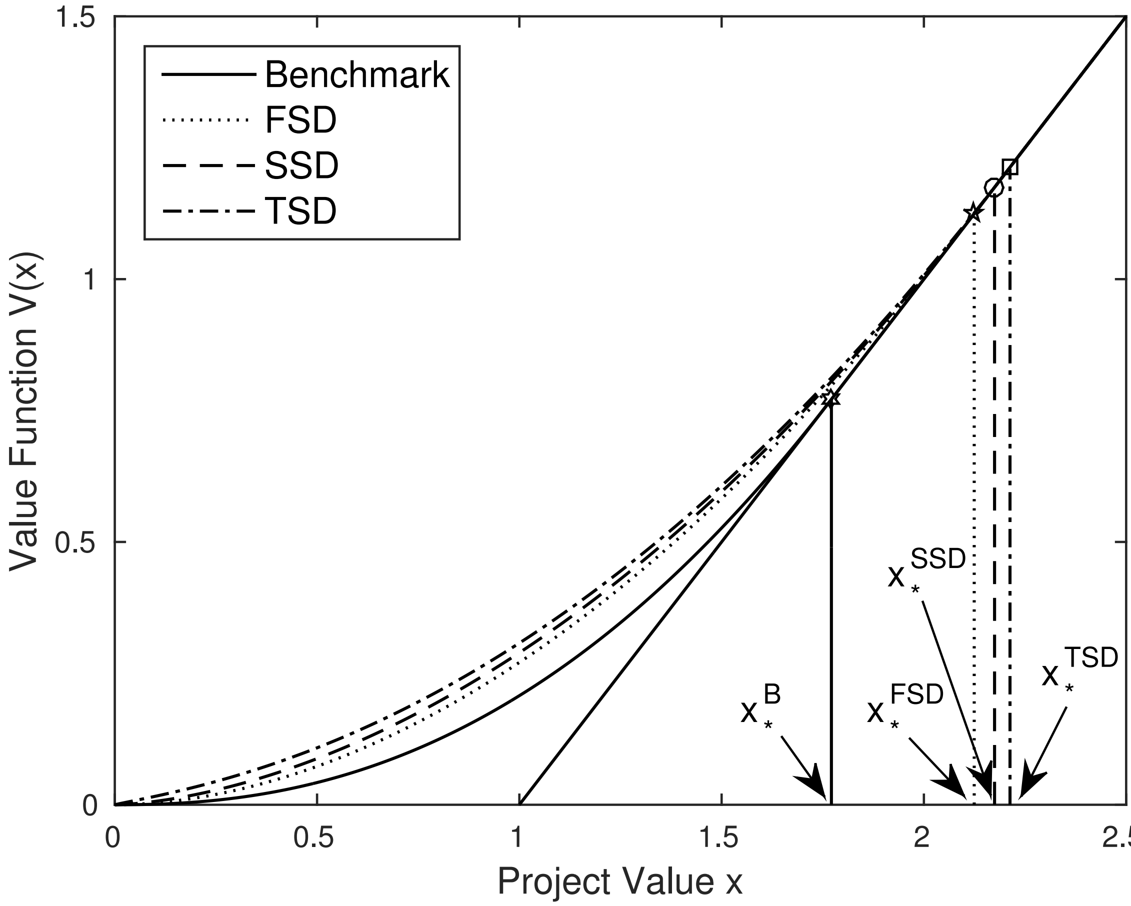
	r	λ	δ	M	SD	S
Benchmark	0.0825	0	1	0.0825	0	0
FSD	0.0525	0	1	0.0525	0	0
SSD	0.03	0.045	0.5	0.0525	0.0225	0
TSD	0.0075	0.0562	0.2	0.0525	0.0225	-1.5000

Notes. The table defines the four scenarios denoted by Benchmark, FSD, SSD, and TSD studied in Figure 2. Each scenario is characterized by a different parametrization of the weighting distribution of a pseudo-exponential discount function as defined in equation (4). The parameters of the weighting distribution are shown in columns two to four, 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 stochastically) dominant than scenario FSD, which in turn is less (second-order stochastically) dominant than scenario SSD, which in turn is less (third-order stochastically) dominant than scenario TSD.

function from equation (3) once more. In particular, 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. The fact that the corresponding weighting distribution given by (4) is binary allows for a simple and intuitive characterization of stochastic dominance relationships through statistical moments. 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 (see Ebert 2015 for the closed-form expressions). 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 more complicated stochastic dominance shifts.

Figure 2 plots the value function of the real option investment problem under pseudo-exponential discounting with four different weighting distributions that differ in their first three moments. Table 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,

Figure 2: The impact of dominance and diversity on investment behavior



Notes. The figure illustrates that less dominance leads to a riskier 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 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 (i.e., less first-order stochastic dominance). $x_*^{FSD} < x_*^{SSD}$ means that a group of two individuals who disagree about the discount rate stops later than a (more second-order stochastically dominant and, in particular, less diverse) single individual whose discount rate corresponds to the group average. $x_*^{SSD} < x_*^{TSD}$ illustrates that a group that is less (third-order stochastically) dominant—which means that there is one member with low weight in the decision but with an extremely low opinion about the appropriate discount rate—stops later than the dominating group whose weighting distribution has equal mean and variance but 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 $I = 1$.

because it has a larger expected discount rate, while standard deviation and skewness of the discount rate distribution are the same. Figure 2 shows that the FSD value function 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.

Next, the weighting distribution of the SSD scenario is less (second-order stochastically) dominant than that of the the FSD scenario. In particular, the SSD weighting distribution is more diverse than that of the FSD scenario. Indeed, the SSD weighting distribution has greater variance than the weighting distribution of the FSD scenario but identical mean and skewness. Figure 2 confirms that greater diversity leads to later investment: $x_*^{SSD} > x_*^{FSD}$. Finally, the figure illustrates that a weighting distribution with lower skewness—that is, with less (third-order stochastic) dominance—results in later investment: $x_*^{TSD} > x_*^{SSD}$. This means that if one 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, given that the group mean and standard deviation of the discount rate are the same.

Turning to the interpretation of our results in terms of present bias, Proposition 2 and Proposition 7 yield the following result.

Corollary 2 (More present-biased groups invest later.) *Consider weighted discount functions h^F and h^G, with equal-mean weighting distributions F and G and measures of decreasing impatience P_F and P_G, respectively. Then*

$$P_G(t) \geq P_F(t) \text{ for all } t \geq 0 \implies x_*^G \geq x_*^F,$$

where $x_*^G(x_*^F)$ is the investment threshold defined by equation (25), with weighting distribution $G(F)$.

As an illustration of Corollary 2, we show that we get unambiguous predictions for how investment depends on the preference parameters α and β of the famous generalized hyperbolic discounting model (recall equations (5) and (6)). From Prelec (2004, p.524), we know that α captures decreasing impatience, because $P(t) = \frac{\alpha}{1+\alpha t}$. The parameter β captures the investor's initial rate of time preference or, equivalently, the group's weighted mean opinion about the appropriate discount rate (cf. (9)): $\rho(0) = \int_0^\infty r dF(r) = \beta$.

Corollary 3 (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 investment threshold* x_* *defined by equation (25)*

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

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

Figure 3 illustrates the quantitative impact of the parameters α and β on the investment threshold x_*. Consistent with Corollary 3, we indeed see that greater decreasing impatience (captured by α) leads to later investment (left panel of Figure 3), while a greater mean of the weighting distribution (captured by β) leads to earlier investment.

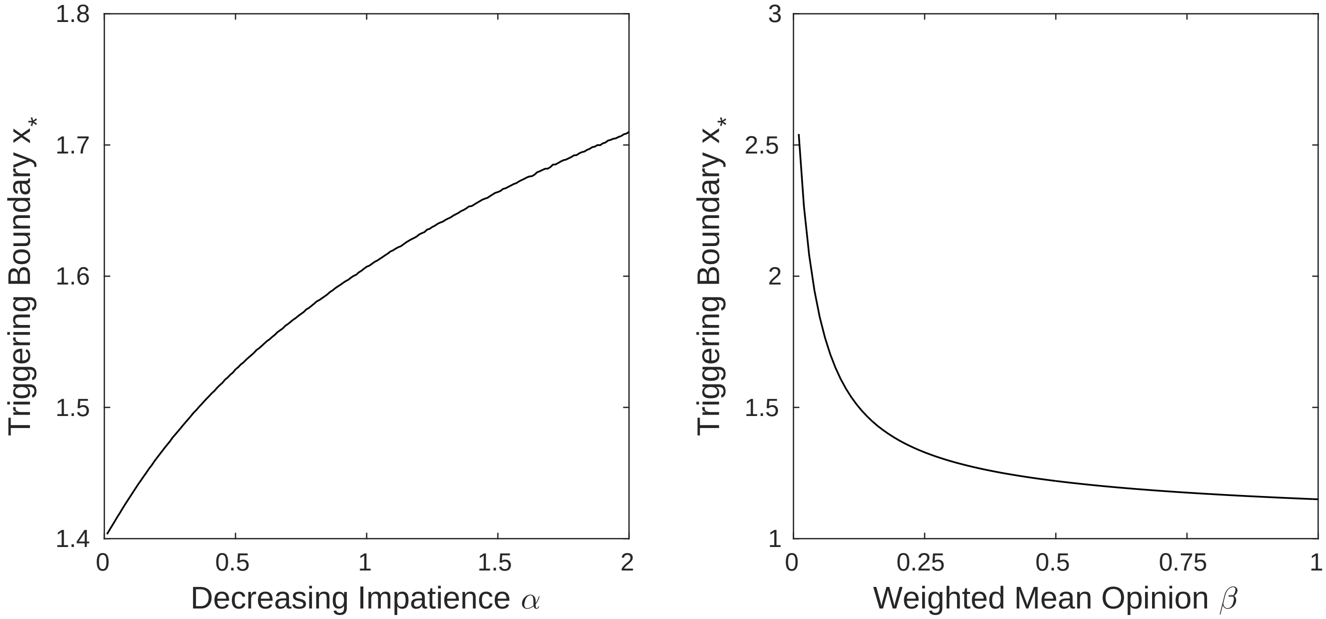
4.5 Comparing the investment behavior of weighted and exponential discounters

Does time-inconsistency—as it arises with weighted discounting due to strictly decreasing impatience—result in sooner or later investment compared to investment under standard, time-consistent exponential discounting? We obtain the following result, which compares the investment threshold of three time-consistent and an arbitrary weighted discounter.

Proposition 8 (Comparing the investment behavior of weighted and exponential discounters.) *Consider a weighted discount function* h^F, *with weighting distribution* F. *Let* $r_{\min} = \inf\{r \in [0,\infty) : F(r) > 0\}$ *and* $r_{\max} = \sup\{r \in [0,\infty) : F(r) > 0\}$, *and* $r_{avg} = \int_0^\infty r dF(r)$ *refer to the average discount rate implied by* F. *Then*

$$x_*^{r_{\max}} \leq x_*^{r_{avg}} \leq x_*^F \leq x_*^{r_{\min}},$$

Figure 3: The impact of decreasing impatience, α, and the weighted mean of the group discount rate, β, on the investment threshold under generalized hyperbolic discounting (Loewenstein and Prelec 1992)



Notes. The figure plots the variation of the investment threshold with respect to the degree of decreasing impatience 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. 2016).

where x_*^F *denotes the investment threshold defined by equation (25), with discount function* h^F. *Likewise,* $x_*^{r_{\max}}$ *(*x_*^{avg}*,* $x_*^{r_{\min}}$*) denotes the investment threshold defined by equation (25), with exponential discount function having discount rate* $r_{\max}$ *(*r_{avg}*,* $r_{\min}$*).*

The result that a group invests earlier (later) than a person whose discount rate is larger (lower) or equal than that of every group member seems expected. The outcome of the comparison with an exponential discounter who uses the same discount rate on average is an application of our main comparative statics result (Proposition 7) that less dominance leads to later investment. In particular, F is a mean-preserving spread of the constant equal to its mean, $\int_0^\infty r dF(r)$, and thus is more diverse and less dominant.

Grenadier and Wang (2007) argue that time-inconsistency (due to present-biasedness) leads to *earlier investment.* It can be shown that this result stems from a comparison analogous to that underlying the last inequality of Proposition 8.[22] In the last paragraph of this section, we will argue that a comparison based on the middle inequality in Proposition 8 is more appropriate so that, in fact, present-biasedness and time-inconsistency lead to *later investment.*

While the last inequality of Proposition 8 generalizes Proposition 4 of Grenadier and Wang (2007), the other comparisons in Proposition 8 are also conceptually new.[23] The question now is: In order to study the impact of time-inconsistency on investment behavior, which of the three exponential discounters—all of which are time-consistent—is the weighted, time-inconsistent discounter best compared to? While it seems intuitive that r_{avg} constitutes a more "balanced" choice than either of the two extreme discount rates $r_{\min}$ or $r_{\max}$, the choice of r_{avg} can also be given firm decision-theoretic support. In order not to confound the effect of time-inconsistency on investment behavior with that of the initial rate of time preference (recall the discussion preceding Proposition 2), the initial rate of time preference of the time-inconsistent DM and the time-consistent DM in comparison should be the same. This is *only* the case for an exponential discounter with discount rate r_{avg}, because the initial rate of time preference of a weighted discounter is given by (recall equation (9))

$$\rho_F(0) = \int_0^\infty r dF(r) = r_{\text{avg}}.$$

In summary, the time-inconsistency introduced through present-biased weighted discounting leads to *later* investment.

[22]More specifically, Grenadier and Wang (2007) consider a stochastic version of the quasi-hyperbolic discounting model, which was later shown to be equivalent to pseudo-exponential discounting by Harris and Laibson (2013). For the case of pseudo-exponential discounting (equation (3)), we have $r_{\min} = r$, $r_{\text{avg}} = \delta r + (1-\delta)(r+\lambda)$, and $r_{\max} = r + \lambda$. Therefore, the last inequality of Proposition 8 indeed generalizes Proposition 4 in Grenadier and Wang (2007), from pseudo-exponential to arbitrary weighted discounting.

[23]Note that, at the time, Grenadier and Wang (2007) could not exploit the equivalence of their quasi-hyperbolic discounting model with the pseudo-exponential discounting model pointed out later by Harris and Laibson (2013). That equivalence suggests a comparison with $r_{\min}$ (being a combination of the two quasi-hyperbolic discounting parameters β and δ.). Likewise, the weighted representation of discounting is analyzed only in this paper. Given the state of research in 2007 and before, therefore, the idea to compare with r_{avg} was not obvious at all.

5 Conclusion

We have introduced the class of *weighted discount functions,* which includes all of the most commonly used discount functions. Weighted discount functions may describe the discounting behavior of groups, uncertainty about what discount rate to use, present-biased time preferences such as hyperbolic discounting, and, through an iteration argument, all of these things simultaneously.

We have collected a number of results on the class of weighted discount functions. Their definition suggests a natural notion of group diversity in time preferences. We show that greater group diversity results in a more elevated discount function so that more diverse groups discount outcomes at any future time by less. We study the implications of group diversity for investment in a real options framework and find that greater group diversity leads to delayed investment. This delay in investment comes with greater risk. The same result applies to investment under intra-personal uncertainty about what discount rate to use: Greater parameter uncertainty leads to delayed investment and more risk-taking. These results are the consequence of an equilibrium stopping result for arbitrary weighted discount functions. Our most general result also clarifies investment behavior in a single-person setting where the DM is, for example, a hyperbolic discounter. This paper thus further contributes to the theoretical literature in behavioral economics and finance by solving the long-standing investment problem under hyperbolic discounting (and many other parametrizations that imply present-biased preferences). Contrary to existing results, we illustrate that time-inconsistency arising from present-biased time preferences leads to delayed rather than premature investment.

On a more general level, our analysis shows that exploiting the link between parameter uncertainty about what discount rate to use, present-biased time preferences, and collective time preferences can offer valuable insights. Investment and equilibrium stopping in a real option setting, as studied in the later sections of this paper, are merely one possible application. Any result on the class of weighted discount functions necessarily makes implications for either of the three fields, thereby illustrating how these seemingly diverse literature streams can benefit from one another.

A Proofs

A.1 Proof of Theorem 2

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 $t > 0$. For this u, $h^F(t) \leq h^G(t)$, $\forall t \geq 0$ follows from the integral inequality in Definition 2

Now suppose that $h^F(t) \leq h^G(t)$ for all $t \geq 0$. For any mixed risk averse utility function u, the first derivative of u satisfies the assumptions of Bernstein's Theorem (Theorem 1). Thus, there exists a distribution function F_u such that $u'(t) = \int_0^\infty e^{-ts} dF_u(s)$. For $0 < a \leq x$, 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,

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

Comparing with the analogous expression for F, 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.

A.2 Proof of Proposition 1

We need to show that $P(t) = -\frac{(\ln h(t))''}{(\ln h(t))'} \geq 0$. The claim immediately follows from the following result, which derives a novel representation of Prelec's measure of decreasing impatience for weighted discount functions.

Lemma 1 *Consider a weighted discount function h^F, with weighting distribution F and measure of decreasing impatience P_F. Then*

$$P_F(t) = \frac{\mathbb{V}[\xi_t]}{\mathbb{E}[\xi_t]}$$

where ξ_t is a random variable with distribution function $F_{\xi_t}(x) = \frac{\int_0^x re^{-rt}dF(r)}{\int_0^\infty e^{-rt}dF(r)}$.

Proof of Lemma 1. The result follows from two simple calculations. First,

$$-(\ln h(t))' = -\frac{h'(t)}{h(t)} = \frac{\int_0^\infty re^{-rt}dF(r)}{\int_0^\infty e^{-rt}dF(r)} = \mathbb{E}[\xi_t]$$

and, second,

$$\begin{aligned}(\ln h(t))'' = \left(\frac{h'(t)}{h(t)}\right)' &= \left(\frac{\int_0^\infty re^{-rt}dF(r)}{\int_0^\infty e^{-rt}dF(r)}\right)' \\ &= \frac{\int_0^\infty r^2e^{-rt}dF(r)}{\int_0^\infty e^{-rt}dF(r)} - \left(\frac{\int_0^\infty re^{-rt}dF(r)}{\int_0^\infty e^{-rt}dF(r)}\right)^2 \\ &= \mathbb{E}[\xi_t^2] - \mathbb{E}[\xi_t]^2 = \mathbb{V}[\xi_t].\end{aligned}$$

A.3 Proof of Proposition 2

Let $\rho_i = -(\ln h^i)', i = F, G$, denote the rate of time preference of weighting distributions F and G. We prove the more general result that only requires that the mean of F is equal or larger than that of F. This assumption yields $(h^F)'(0) = -\int_0^\infty rdF(r) \leq -\int_0^\infty rdG(r) = (h^G)'(0)$, so that

$$\rho_F(0) = -\frac{(h^F)'(0)}{h^F(0)} \geq -\frac{(h^G)'(0)}{h^G(0)} = \rho_G(0). \tag{28}$$

Since $P_i = -\frac{\rho_i'}{\rho_i} = -(\ln \rho_i)', i = F, G$, the assumption that F exhibits less DI than G; that is, $P_F(t) \le P_G(t) \forall t \ge 0$ can be restated as

$$(\ln \rho_F(t))' \ge (\ln \rho_G(t))'. \tag{29}$$

Therefore, (28) and (29) together imply that $\ln \rho_F(t) \ge \ln \rho_G(t), \forall t \ge 0$, which is equivalent to

$$\rho_F(t) \ge \rho_G(t), \ \forall t \ge 0. \tag{30}$$

Since $\rho_i = -(\ln h^i)'$ and $h^i(0) = 1, i = F, G$, we have $\ln h^F(t) \le \ln h^G(t)$, and the result follows from Theorem 2.

A.4 Proof of Proposition 4

By assumption, $\frac{h^G}{h^F}$ is increasing. The claim follows from equation (10) and Theorem 2.

A.5 Proof of Theorem 3

We have proven in equation (21) that

$$\mathbb{E}[h(\tau_{\hat{u}})G(X_{\hat{u}})] = \int_0^\infty w(x;r)dF(r) \equiv V(x).$$

It now suffices to show that $\hat{u}$ is an equilibrium stopping rule, namely, that it satisfies equation (17), subject to condition (14) in Definition 7.

In the case of $X_0 = 0$, due to the boundedness of b and σ, the process X will always stay at 0. Hence the stopping rule $\hat{u}$, determined by the boundary condition $V(0) = \max\{G(0), 0\}$, is trivially an equilibrium stopping rule.

In the case of $X_0 = x > 0$, if $a = 1$, then $J(x; u^{\epsilon,a}) = G(x)$. Due to the Bellman equation (18), we have $V(x) \ge G(x)$. This shows that equation (17) is satisfied.

We now turn to the case of $a = 0$. For notational convenience, we define $\tau_u^t = \inf\{s \ge t :$

$u(X_s) = 1\}$, $\forall t > 0$, where $X_0 = x$. Noting that, by construction, $\tau_{u^{\epsilon,0}} \equiv \tau^{\epsilon}_{\hat{u}}$, we have

$$\begin{aligned}
J(x; u^{\epsilon,0}) =& \mathbb{E}[h(\tau_{u^{\epsilon,0}})G(X_{\tau_{u^{\epsilon,0}}})] \\
=& \mathbb{E}[h(\tau^{\epsilon}_{\hat{u}} - \epsilon)G(X_{\tau^{\epsilon}_{\hat{u}}})] + \mathbb{E}[(h(\tau^{\epsilon}_{\hat{u}}) - h(\tau^{\epsilon}_{\hat{u}} - \epsilon))G(X_{\tau^{\epsilon}_{\hat{u}}})] \\
=& \mathbb{E}[\mathbb{E}[h(\tau^{\epsilon}_{\hat{u}} - \epsilon)G(X_{\tau^{\epsilon}_{\hat{u}}})|X_\epsilon]] + \mathbb{E}[(h(\tau^{\epsilon}_{\hat{u}}) - h(\tau^{\epsilon}_{\hat{u}} - \epsilon))G(X_{\tau^{\epsilon}_{\hat{u}}})] \\
=& \mathbb{E}[V(X_\epsilon)] + \mathbb{E}[(h(\tau^{\epsilon}_{\hat{u}}) - h(\tau^{\epsilon}_{\hat{u}} - \epsilon))G(X_{\tau^{\epsilon}_{\hat{u}}})] \\
=& \mathbb{E}[V(X_\epsilon)] + \mathbb{E}\left[\left(\int_0^\infty e^{-r\tau^{\epsilon}_{\hat{u}}} dF(r) - \int_0^\infty e^{-r(\tau^{\epsilon}_{\hat{u}} - \epsilon)} dF(r)\right) G(X_{\tau^{\epsilon}_{\hat{u}}})\right] \\
=& \mathbb{E}[V(X_\epsilon)] + \int_0^\infty (e^{-r\epsilon} - 1)\mathbb{E}[e^{-r(\tau^{\epsilon}_{\hat{u}} - \epsilon)} G(X_{\tau^{\epsilon}_{\hat{u}}})] dF(r) \\
=& \mathbb{E}[V(X_\epsilon)] + \int_0^\infty (e^{-r\epsilon} - 1)\mathbb{E}[\mathbb{E}[e^{-r(\tau^{\epsilon}_{\hat{u}} - \epsilon)} G(X_{\tau^{\epsilon}_{\hat{u}}})|X_\epsilon]] dF(r) \\
=& \mathbb{E}[V(X_\epsilon)] + \int_0^\infty (e^{-r\epsilon} - 1)\mathbb{E}[w(X_\epsilon; r)] dF(r).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\liminf_{\epsilon \to 0} \frac{J(x; \hat{u}) - J(x; u^{\epsilon,0})}{\epsilon} =& \liminf_{\epsilon \to 0} \frac{V(x) - \mathbb{E}[V(X_\epsilon)]}{\epsilon} \\
&+ \liminf_{\epsilon \to 0} \int_0^\infty \frac{(1 - e^{-r\epsilon})}{\epsilon} \mathbb{E}[w(X_\epsilon; r)] dF(r) \\
=& - (\mathcal{A}V)(x) + \int_0^\infty r w(x; r) dF(r) \geq 0,
\end{aligned}$$

where the inequality follows from the Bellman equation (18). This completes the proof.

A.6 Proof of Proposition 5

The proof is structured into three parts. In Part 1, we derive the explicit expression of the investment threshold, equation (25). In Part 2, we verify that this threshold is indeed part of a solution to the Bellman system in Theorem 3. In Part 3, we prove that this threshold is unique.

Proof of Part 1 (derivation of the investment threshold). 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 $r \geq 0$. With the infinitesimal generator of the geometric Brownian motion being $\mathcal{A} =$

$\frac{1}{2}\sigma^2 x^2 \frac{d^2}{dx^2} + bx\frac{d}{dx}$, for $x < x_*$, we obtain as usual (e.g., Dixit and Pindyck 1994, p.141ff) that

$$\frac{1}{2}\sigma^2 x^2 w_{xx}(x;r) + bxw_x(x;r) - rw(x;r) = 0, \tag{31}$$

with boundary conditions $w(0;r) = 0$ and $w(x_*;r) = x_* - I$. Letting $\mu(r)$ denote the positive square root of the fundamental quadratic

$$\frac{1}{2}\sigma^2\mu^2 + (b - \frac{1}{2}\sigma^2)\mu - r = 0,$$

the solution to the ordinary differential equation (31) is given by

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

and the explicit expression for $\mu(r)$, equation (26), follows immediately. Second, the equilibrium stopping rule maximizes the weighted average of the $w(x;r)$, which is given by

$$V(x) = \int_0^\infty \left(\frac{x}{x_*}\right)^{\mu(r)} dF(r)(x_* - I). \tag{32}$$

The smooth-pasting condition is thus not imposed on w but on the value function V :

$$\int_0^\infty w_x(x_*;r)dF(r) = 1.$$

This condition is used to identify the triggering threshold x_*; that is,

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

Solving for x_* yields equation (25).

Proof of Part 2 (x_* is part of a solution to the Bellman system). We prove that 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 (18)–(19) with its boundary condition, as specified in Theorem 3. Then, $\hat{u}$ must be an equilibrium stopping rule, V^e the corresponding value function, and (an) equilibrium behavior is given by stopping once the process X hits x_*.

From the boundary condition $w(0;r) = 0$, it is easy to see that $V(0) = 0$, which satisfies the boundary condition $V(0) = \max\{0, x - I\}|_{x=0}$. 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 (31), we have

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

To obtain the Bellman equation (31), it remains to check that

$$V(x) - (x - I) > 0 \tag{33}$$

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

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

or, equivalently,

$$\int_0^\infty (r - b)x dF(r) \geq \int_0^\infty rI dF(r) \tag{34}$$

on the stopping region $[x_*, \infty)$. As regards inequality (33), due to condition (24), we have that $\mu(r) > 1$ and thus $V(x)$ is a convex function. Then the linearity of the payoff function, together with the value matching and smooth pasting conditions, yields inequality (33).

As regards inequality (34), 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 (34) 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 (34) 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

$$\int_0^\infty (r-b)\frac{\int_0^\infty \mu(r)dF(r)}{\int_0^\infty \mu(r)dF(r)-1}dF(r) \geq (r_m - b)\frac{\mu(r_m)}{\mu(r_m)-1}, \tag{35}$$

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

$$(r_m - b)\frac{\mu(r_m)}{\mu(r_m)-1} \geq r_m, \tag{36}$$

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

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

Note that $f(b) = 0$. Therefore, $f(r) \geq 0$, which yields inequality (36). This completes Part 2 of the proof.

Proof of Part 3 (Uniqueness of the investment threshold). Suppose $\bar{x}$ is an equilibrium investment threshold. We have to show that $\bar{x} = x_*$. A stopping rule that induces $\bar{x}$ is given by $\bar{u}$ defined by $\bar{u}(x) = 0$ if $x < \bar{x}$ and $\bar{u}(x) = 1$ otherwise. Let $\bar{V}(x) := J(x;\bar{u})$.

It follows from the proof of Part 1 that the smooth pasting principle (i.e., $D_x^- \bar{V}(\bar{x}) = D_x^+ \bar{V}(\bar{x}) = 1$, where $D_x^- \bar{V}(\bar{x}) = \lim_{x\to\bar{x}-} \bar{V}_x(x)$ and $D_x^+ \bar{V}(\bar{x}) = \lim_{x\to\bar{x}+} \bar{V}_x(x)$) yields a unique solution. Hence, if we can prove that the value function $\bar{V}$ is obtained by the smooth pasting principle, then the conclusion that $\bar{x} = x_*$ follows.

To complete the proof, it suffices to exclude the following possibilities: (a) $\bar{x} = \infty$, that is, the stopping region is empty; (b) $\bar{x} = 0$, that is, the continuation region is empty; and (c) $\bar{x} \in (0, \infty)$ and $D_x^- \bar{V}(\bar{x}) \neq D_x^+ \bar{V}(\bar{x})$. This is because only when one of these possibilities occurs does $\bar{V}$ not satisfy the smooth pasting principle.

Proof that (a) is impossible: $\bar{x} \neq \infty$. If not, we have $\tau_{\bar{u}} = \infty$. Then the well-posedness condition (24) ensures that $e^{-r\infty}(X_\infty - I) := \limsup_{t \to \infty} e^{-rt}(X_t - I) = 0$, which implies that $\bar{V}(x) = 0$. Therefore, for $x_0 > I$, we have that

$$\liminf_{\epsilon \to 0} \frac{J(x_0; \bar{u}) - J(x_0; \bar{u}^{\epsilon,1})}{\epsilon} = \liminf_{\epsilon \to 0} \frac{0 - (x_0 - I)}{\epsilon} = -\infty,$$

where $\bar{u}^{\epsilon,1}$ is the deviation stopping rule defined in equation (17). This contradicts the definition of an equilibrium stopping rule and thus $\bar{x} \neq \infty$.

Proof that (b) is impossible: $\bar{x} \neq 0$. If not, it is easy to see that

$$J(x; \bar{u}^{\epsilon,0}) = \int_0^\infty \mathbb{E}[e^{-r\epsilon}(X_\epsilon - I) | X_0 = x] dF(r) = \int_0^\infty x e^{(b-r)\epsilon} dF(r) - \int_0^\infty e^{-r\epsilon} I dF(r)$$

and $J(x; \bar{u}) = x - I$. Then some simple algebra yields that

$$\liminf_{\epsilon \to 0} \frac{J(x; \bar{u}) - J(x; \bar{u}^{\epsilon,0})}{\epsilon} = \int_0^\infty (r - b) x dF(r) - I \int_0^\infty r dF(r).$$

Note that condition (24) yields that $\int_0^\infty (r - b) dF(r) > 0$. Therefore $\liminf_{\epsilon \to 0} \frac{J(x;\bar{u}) - J(x;\bar{u}^{\epsilon,0})}{\epsilon} < 0$, if $x < \frac{I \int_0^\infty r dF(r)}{\int_0^\infty (r-b) dF(r)}$, which shows that $\bar{x} \neq 0$.

Proof that (c) is impossible. Suppose that $\bar{x} \in (0, \infty)$ and $D_x^- \bar{V}(\bar{x}) \neq D_x^+ \bar{V}(\bar{x})$. Since $\bar{V}(x) \geq x - I$ and $\bar{V}(\bar{x}) = \bar{x} - I$, we have

$$D_x^- \bar{V}(\bar{x}) < 1. \tag{37}$$

Similar to the proof of Theorem 3, we have that for any $x > 0$,

$$\liminf_{\epsilon \to 0} \frac{J(x;\bar{u}) - J(x;\bar{u}^{\epsilon,0})}{\epsilon} = \liminf_{\epsilon \to 0} \frac{\bar{V}(\bar{x}) - \mathbb{E}[\bar{V}(X_\epsilon)|X_0 = \bar{x}]}{\epsilon} + \int_0^\infty r\bar{w}(x;r)dF(r),$$

where $\bar{w}(x;r) = \mathbb{E}[e^{-r\tau_{\bar{u}}}(X_{\tau_{\bar{u}}} - I)|X_0 = x]$.

Then it follows from (generalized) Ito's formula (see, e.g., Chapter 3 of Karatzas and Shreve 2006) that

$$\begin{aligned}
\liminf_{\epsilon \to 0} \frac{J(\bar{x};\bar{u}) - J(\bar{x};\bar{u}^{\epsilon,0})}{\epsilon} &= \liminf_{\epsilon \to 0} \frac{1}{\epsilon}\mathbb{E}(\int_0^\epsilon (-bX_s D_x^- \bar{V}(X_s) - \frac{1}{2}\sigma^2 X_s^2 D_{xx}^- \bar{V}(X_s))ds \\
&\quad - \frac{1}{2}(D_x^+\bar{V}(\bar{x}) - D_x^-\bar{V}(\bar{x}))L_\epsilon^{\bar{x}}|X_0 = \bar{x}) + \int_0^\infty rdF(r)(\bar{x} - I) \\
&= \liminf_{\epsilon \to 0} \frac{1}{\epsilon}\mathbb{E}(\int_0^\epsilon (-bX_s D_x^- \bar{V}(X_s) - \frac{1}{2}\sigma^2 X_s^2 D_{xx}^- \bar{V}(X_s))ds|X_0 = \bar{x}) \\
&\quad - \limsup_{\epsilon \to 0} \frac{1}{\epsilon}\mathbb{E}(\frac{1}{2}(1 - D_x^-\bar{V}(\bar{x}))L_\epsilon^{\bar{x}}|X_0 = \bar{x}) + \int_0^\infty rdF(r)(\bar{x} - I), \qquad (38)
\end{aligned}$$

where the process L_t^x is the local time of $\{X_t\}_{t\geq 0}$, $D_{xx}^-\bar{V}(x) = \lim_{y\to x-} \bar{V}_{xx}(y)$, and $D_{xx}^+\bar{V}(x) = \lim_{y\to x+} \bar{V}_{xx}(y)$. Moreover, by Tanaka's formula (see, e.g., Chapter 3 of Karatzas and Shreve 2006), L_t^x can be written as

$$L_t^x = 2(X_t - x)^+ - 2(X_0 - x)^+ - 2\int_0^t I_{X_s \geq x} dX_s,$$

and thus

$$\mathbb{E}[L_\epsilon^{\bar{x}}|X_0 = \bar{x}] = 2\mathbb{E}[(X_\epsilon - \bar{x})^+ - \int_0^\epsilon bX_s I_{X_s \geq \bar{x}} ds|X_0 = \bar{x}].$$

Using standard calculations it is easy to see that

$$\mathbb{E}[(X_\epsilon - \bar{x})^+|X_0 = \bar{x}] = \bar{x}e^{-b\epsilon}N(d_1) - \bar{x}N(d_2)$$

where $N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x e^{-\frac{1}{2}y^2}dy, d_1 = \frac{b+\frac{\sigma^2}{2}}{\sigma}\sqrt{\epsilon}, d_2 = \frac{b-\frac{\sigma^2}{2}}{\sigma}\sqrt{\epsilon}$.

Therefore,

$$\lim_{\epsilon\to 0+}\frac{1}{\epsilon}\mathbb{E}[(X_\epsilon-\bar{x})^+|X_0=\bar{x}]=\lim_{\epsilon\to 0+}\frac{\bar{x}\sigma}{2\sqrt{2\pi\epsilon}}=\infty. \tag{39}$$

Then it follows from (37), (38), and (39) that

$$\liminf_{\epsilon\to 0}\frac{J(\bar{x};\bar{u})-J(\bar{x};\bar{u}^{\epsilon,0})}{\epsilon}=-\infty,$$

which shows that $\bar{u}$ cannot be an equilibrium and thus completes the proof.

A.7 Proof of Proposition 6

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 the $\mu(r)$ defined in equation (26), we only need to check the monotonicity of $\mu(r)$ with respect to b and σ^2.

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

$$\begin{aligned}\frac{\partial\nu}{\partial b}&=\frac{1}{\sigma^2}\left(-1+\frac{b-\frac{1}{2}\sigma^2}{\sqrt{(b-\frac{1}{2}\sigma^2)^2+2\sigma^2 r}}\right)\leq 0 \text{ and}\\ \frac{\partial\nu}{\partial\sigma^2}&=\frac{1}{\sigma^4\sqrt{(b-\frac{1}{2}\sigma^2)^2+2\sigma^2 r}}\left(b\sqrt{(b-\frac{1}{2}\sigma^2)^2+2\sigma^2 r}+\frac{1}{2}\sigma^2 b-b^2-\sigma^2 r\right).\end{aligned}$$

If $b\leq 0$, it is easy to see that $\frac{\partial\nu}{\partial\sigma^2}\leq 0$. If $b>0$, since $b<r$, define $f(r)=b\sqrt{(b-\frac{1}{2}\sigma^2)^2+2\sigma^2 r}+\frac{1}{2}\sigma^2 b-b^2-\sigma^2 r$. Then, after some algebra, we have

$$f'(r)=\frac{b\sigma^2}{\sqrt{(b-\frac{1}{2}\sigma^2)^2+2\sigma^2 r}}-\sigma^2<\frac{b\sigma^2}{\sqrt{(b-\frac{1}{2}\sigma^2)^2+2\sigma^2 b}}-\sigma^2=\frac{b\sigma^2}{b+\frac{1}{2}\sigma^2}-\sigma^2<0.$$

Then $\frac{\partial\nu}{\partial\sigma^2}<0$ follows from $f(b)=0$. This completes the proof.

A.8 Proof of Proposition 7

Let X and Y be two random variables with distributions F^X and F^Y, respectively. Then the investment thresholds can be written as

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

where μ is defined by equation (26). The main idea of the proof is to note that μ satisfies $\operatorname{sgn} \mu^{(n)} = (-1)^{n+1}$ for all $n \in \mathbb{N}^+$; that is, $\mu(r)$ can be interpreted as a mixed risk averse "utility function over discount rates." Then by definition of infinity stochastic dominance, $F^Y \preceq_{\infty SD} F^X$ implies that $\mathbb{E}[\mu(X)] \geq \mathbb{E}[\mu(Y)]$. This yields the conclusion and completes the proof.

A.9 Proof of Corollary 1

The result follows from Proposition 7 and Theorem 2, but here we give the direct proof motivated in the main text. We first prove the following lemma, which is of its own interest. It provides a semi-analytic representation of the investment threshold x_* in terms of the weighted discount function h^F (rather than in terms of its weighting distribution F, as does equation (25)).

Lemma 2 (Threshold in terms of the discount function rather than the weighting distribution.) *Consider a weighted discount function h^F with weighting distribution F, which yields the investment threshold x_* in equation (25). Then*

$$x_* = \frac{\max\{\frac{-2(b-\frac{1}{2}\sigma^2)}{\sigma^2}, 0\} + B}{\max\{\frac{-2(b-\frac{1}{2}\sigma^2)}{\sigma^2}, 0\} + B - 1} I, \tag{40}$$

where

$$B = \frac{1}{\sqrt{2\pi}\sigma} \int_0^\infty t^{-\frac{3}{2}} e^{-\frac{(b-\frac{1}{2}\sigma^2)^2}{2\sigma^2} t} (1 - h^F(t)) dt. \tag{41}$$

Proof of Lemma 2. Using some algebra, it can be shown that

$$\int_0^\infty \mu(r) dF(r) = \frac{-(b - \frac{1}{2}\sigma^2)}{\sigma^2} + \frac{\sqrt{2}}{\sigma} \int_0^\infty \sqrt{r + C} dF(r), \tag{42}$$

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

$$\sqrt{r+C} = \frac{1}{2\sqrt{C}} \int_0^r \frac{1}{\sqrt{1+\frac{1}{C}s}} ds + \sqrt{C}. \tag{43}$$

Moreover,

$$\frac{1}{\sqrt{1+\frac{1}{C}s}} = \int_0^\infty e^{-st} f(t; \frac{1}{2}, \frac{1}{C}) dt, \tag{44}$$

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 $\frac{1}{C}$; that is,

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

Next, we plug equation (44) into equation (43) to obtain

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

Hence,

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

Finally, we plug the last equality into equation (42) 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 Corollary 1. From equation (41) in Lemma 2, it follows, because $h^F \geq h^G$, that $B_F \leq B_G$, with B_i $(i = F, G)$ defined in equation (41). Then, from equation (40) in Lemma 2, it follows that the investment threshold is decreasing in the factor B, which completes the proof.

A.10 Proof of Corollary 2

The result is an immediate consequence of Proposition 2 and Proposition 7.

A.11 Proof of Corollary 3

Consider two hyperbolic discount functions h^i $(i = F, G)$ with parameters α^i and β_i and measures of decreasing impatience $P_i(t)$. To prove (i), suppose that $\beta_F = \beta_G$ and $\alpha_F \leq \alpha_G$. Note that $\beta_F = \beta_G$ implies that the expectation of F is equal to that of G and

$$\alpha_G \geq \alpha_F \iff \frac{\alpha_G}{1+\alpha_G t} \geq \frac{\alpha_F}{1+\alpha_F t} \iff P_G(t) \geq P_G(t). \tag{45}$$

Therefore, part (i) follows from Corollary 2. To prove (ii), suppose that $\beta_F \leq \beta_G$ and $\alpha_F = \alpha_G$. Note that $\beta_F \leq \beta_G$ implies that the expectation of F is no less than that of G, and, analogously to the computations in (45), we find that $\alpha_F = \alpha_G$ yields $P_G(t) = P_G(t)$. Thus, part (ii) also follows from Corollary 2.

A.12 Proof of Proposition 8

Note that we have the following relationship (the second inequality follows from Theorem 2, as F is more diverse than its mean r_{avg}) :

$$e^{-r_{\max}t} \leq e^{-r_{\text{avg}}t} \leq h^F(t) = \int_0^\infty e^{-rt} dF(r) \leq \int_0^\infty e^{-r_{\min}t} dF(r) = e^{-r_{\min}t}.$$

Therefore, the result follows from Corollary 1.

B Groups of non-exponential discounters

In this section, we formalize the "weighting iteration argument" and illustrate it through a concrete example. 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 present-bias. 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; that is, $\bar{h}(t) = \sum_i g(i)h(t;i)$. In the general case, we have

$$\bar{h}(t) = \int_0^\infty h(t;i) dG(i). \tag{46}$$

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

Proposition 9 *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 notation above and analyzes the investment behavior of a group of non-exponential discounters within the real options framework of Section 4. 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 (46) is the weighting distribution of a pseudo-exponential discount function (see equation (4)). 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 (7), these are weighted discount functions with exponentially distributed weighting distributions. That is, in the notation above, $F(r;i) \equiv F(r;\alpha_i) = 1 - e^{-\frac{r}{\alpha_i}}$, 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 9, the discount function of the group defined in equation (46),

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

is also a weighted discount function. For that reason, our results on group investment behavior in Section 4 also apply to this group of "behavioral" investors. In order to compute the investment threshold of this group, we need to know its weighting distribution $\bar{F}$. By Proposition 9, its cumulative density function 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 equation (25), the triggering threshold can be computed as

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

C On the well-posedness condition (24)

The following proposition shows that equation (24) is indeed the necessary and sufficient condition for our generalized real options investment problem to be well-defined.

Proposition 10 *If condition (24) 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 arguments, 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

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

Here, the second inequality follows 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 $r < b$, $\mathbb{E}[e^{-r\tau_*}(X_{\tau_*} - I)] = \infty$. Moreover, when $b > r_0$, $F(b-) > 0$, so that

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

This completes the proof.

References

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.

Adams, R. (2015): "Women on Boards: The Superheroes of Tomorrow?" Working paper.

Adams, R., H. Almeida, and D. Ferreira (2005): "Powerful CEOs and Their Impact on Corporate Performance," *Review of Financial Studies*, 18, 1403–1432.

Adams, R. and D. Ferreira (2009): "Women in the Boardroom and Their Impact on Governance and Performance," *Journal of Financial Economics*, 94, 291–309.

Ahern, K. and A. Dittmar (2012): "The Changing of the Boards: The Impact on Firm Valuation of Mandated Female Board Representation," *Quarterly Journal of Economics*, 127, 137–197.

Attema, A., H. Bleichrodt, Y. Gao, Z. Huang, and P. Wakker (2016): "Measuring Discounting without Measuring Utility," *American Economic Review*, 106, 1476–1494.

Bebchuk, L., K. Cremers, and U. Peyer (2011): "The CEO Pay Slice," *Journal of Financial Economics*, 102, 199–221.

Bernile, G., V. Bhagwat, and S. Yonker (2016): "Board Diversity, Firm Risk, and Corporate Policies," Working paper.

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

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.

Chambers, C. and F. Echenique (2018): "On Multiple Discount Rates," *Econometrica*, 86, 1325–1346.

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

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

——— (forthcoming): "Decision Making When Things Are Only a Matter of Time," *Operations Research*.

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.

Ekern, S. (1980): "Increasing Nth Degree Risk," *Economics Letters*, 6, 329–333.

Fahlenbrach, R. (2009): "Founder-CEOs, Investment Decisions, and Stock Market Performance," *Journal of Financial and Quantitative Analysis*, 44, 439–466.

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.

Fishburn, P. (1980): "Continua of Stochastic Dominance Relations for Unbounded Probability Dstributions," *Journal of Mathematical Economics*, 3, 94–100.

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

Giannetti, M. and M. Zhao (2019): "Board Diversity and Firm Performance Volatility," *Journal of Financial and Quantitative Analysis*, 54, 1117–1155.

Gollier, C. (2014): *Pricing the Future: The Economics of Discounting and Sustainable Development*, 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.

Güner, A., U. Malmendier, and G. Tate (2008): "Financial Expertise of Directors," *Journal of Financial Economics*, 88, 323–354.

Halevy, Y. (2015): "Time Consistency: Stationarity and Time Invariance," *Econometrica*, 83, 335–352.

Hara, C. (2008): "Complete Monotonicity of the Representative Discount Factor," *Journal of Mathematical Economics*, 44, 1321–1331.

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.

Jackson, M. and L. Yariv (2014): "Present Bias and Collective Dynamic Choice in the Lab," *American Economic Review*, 104, 4184–4204.

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

Jouini, E., J. Marin, and C. Napp (2010): "Discounting and Divergence of Opinion," *Journal of Economic Theory*, 145, 830–859.

Jouini, E. and C. Napp (2007): "Consensus Consumer and Intertemporal Asset Pricing with Heterogeneous Beliefs," *Review of Economic Studies*, 74, 1149–1174.

Karatzas, I. and S. Shreve (2006): *Brownian Motion and Stochastic Calculus*, 2nd edition, Springer, New York, NY.

Karp, L. (2007): "Non-constant Discounting in Continuous Time," *Journal of Economic Theory*, 132, 557–568.

Kim, D. and L. Starks (2016): "Gender Diversity on Corporate Boards: Do Women Contribute Unique Skills?" *American Economic Review*, 106, 267–271.

Koopmans, T. (1960): "Stationary Ordinal Utility and Impatience," *Econometrica*, 28, 287–309.

Laibson, D. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112, 443–477.

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

Manconi, A., A. Rizzo, and O. Spalt (2016): "Diversity Investing," Working paper.

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

Masulis, R., C. Wang, and F. Xie (2007): "Corporate Governance and Acquirer Returns," *Journal of Finance*, 62, 1851–1889.

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, New Jersey, 55–73.

McDonald, R. and D. Siegel (1986): "The Value of Waiting to Invest," *Quarterly Journal of Economics*, 101, 707–727.

Menezes, C. F. and X. H. Wang (2005): "Increasing Outer Risk," *Journal of Mathematical Economics*, 41, 875–886.

O'Donoghue, T. and M. Rabin (1999): "Doing It Now or Later," *American Economic Review*, 89, 103–124.

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 Political 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.

SHAKED, M. AND J. SHANTIKUMAR (2007): *Stochastic Orders*, Springer, New York.

SOUZOU, P. (1998): "On Hyperbolic Discounting and Uncertain Hazard Rates," *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265, 2015–2020.

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

THISTLE, P. (1993): "Negative Moments, Risk Aversion, and Stochastic Dominance," *Journal of Financial and Quantitative Analysis*, 28, 301–311.

VANDEN, J. (2015): "General Properties of Isoelastic Utility Economies," *Mathematical Finance*, 25, 187–219.

WEITZMAN, M. (2001): "Gamma Discounting," *American Economic Review*, 91, 260–271.