Abstract

This paper presents the class of weighted discount functions, which contains the discount functions commonly used in economics and finance. Weighted discount functions also describe the discounting behavior of groups, and they admit a natural notion of group diversity in time preference. We show that more diverse groups discount less heavily, and make more patient decisions. Greater group diversity leads to more risk-taking and delayed investment in a real options setting. We further provide a general result on the timing behavior of groups, and link it to that of individuals who are time-inconsistent.

Key words: diversity; investment; time preference

JEL codes: G02; G11; D91
1 Introduction

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

First, and most importantly, 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 owns all the decision power, she may still attach some weight to the opinions of others when making her decision. CEOs consult with board members, experts, or friends. In modern societies, the members of a household make important decisions together. And also presidents collect the opinions of party members, advisers – or their spouse.

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

Third, our model takes into account the value of timing an investment optimally, as put forward in the real options approach proposed by Brennan and Schwartz (1985) and McDonald.

¹Frederick 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.
and Siegel (1986). For example, in some situations it can be beneficial to wait and see, rather than to decide for or against an investment right away. The real options literature has neither investigated the investment decisions of groups, nor has it given answers to the discount rate dilemma. This paper offers a solution to both issues. Our contribution to the literature on sustainable investment and cost-benefit analysis is, on the other hand, to incorporate insights from the real options literature regarding the investment timing. Moreover, we contribute to a recent and growing literature on the importance of team diversity for financial outcomes in general.\footnote{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, Kim and Starks 2016), studies on CEO power vis-à-vis the board (e.g., Adams et al. 2005, Fahlenbrach 2009, 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, Coles et al. forthcoming), and studies that combine several characteristics into an index (e.g., Giannetti and Zhao 2016, Bernile et al. 2016). Manconi et al. (2016) show that investing into 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.} To the best of our knowledge, our paper comprises the first theoretical contribution that establishes a connection between team diversity and investment decisions.

Fourth, we allow for the fact that some or all group members have less than fully rational time preferences. There is a close connection between the discounting behavior of groups and that of non-exponential, “behavioral” discounters, and we provide the necessary and sufficient condition for a behavioral discount function to be representable as a group discount function. Through this connection, we can generalize existing results on investment behavior to the case where discounting is less than fully rational. For example, we generalize the seminal real option result from McDonald and Siegel (1986) from exponential discounting to hyperbolic discounting.

The starting point of this paper is to provide an in-depth study of group discount functions (Section 2). Intuitively, the discount function of a group is given by a weighted average of the discount functions of its members, where more important or more valued members receive more weight. For sake of concreteness, first suppose that all group members are exponential discounters. 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 discount function (also weighted discount function) is given by

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

The idea of averaging discount functions to model social or group preferences goes back to seminal articles such as Marglin (1963) and Feldstein (1964). The current paper builds upon the literature on negotiation, voting, and preference aggregation, which is concerned with determining the weights of group members, and sets in once the weighting distribution $F$ is determined. We provide a comprehensive analysis of the class of weighted discount functions by proving a number of results about their importance, prevalence, and other properties. In particular, we explain how properties of the weighting distribution $F$ translate into properties of the discount function $h^F$, and vice versa.

The first main contribution of Section 2 is to introduce and characterize a notion of group diversity, which is naturally suggested by the definition of a weighted discount function. We define a group as being more diverse if its weighting distribution is more dispersed than that of another group, where dispersion is measured using a proper notion of stochastic dominance. Intuitively, greater group diversity means that there are stronger differences of opinions among its members about what discount rate to use. We then characterize group diversity (a property of the weighting distribution $F$) in terms of a property of the weighted discount function $h^F$. We prove that a group is more diverse than another group if and only if, at any future time, its discount factor is larger. Graphically, the more diverse a group is, the more elevated is its discount function.

We further clarify the relationship between group diversity and the important concepts of patience (Quah and Strulovici 2013) and decreasing impatience (Prelec 2004). Moreover, we show that all weighted discount functions exhibit decreasing impatience as defined by Prelec (2004), which complements Jackson and Yariv’s (2015) important result that a group discount function as above necessitates a present bias.

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 (i.e., group) disagreement, but also intra-personal disagreement. One might then wonder whether some of the commonly used discount functions can be written as a weighted discount function. That is, for a given, well-known discount function $h$, does there exist a distribution of opinions $F$ such that the resulting group discount function is given by $h = h^F$? While previous papers have given important examples for this to be the case, we provide the necessary and sufficient condition for a discount function to be of the weighted form. The mathematical result is not new and known as Bernstein’s Theorem in the literature on LaPlace transforms, but—to the best of our knowledge—it has not been applied to discount functions yet.

A first insight from the application of Bernstein’s Theorem is that the observation, that a behavioral discount function may be represented as a group discount function, is rather the rule than the exception. Consequently, the class of weighted discount functions is large. We show that the exponential (Samuelson 1937), hyperbolic (Harvey 1986), proportional (Mazur 1987; Harvey 1995), generalized hyperbolic (Loewenstein and Prelec 1992), and constant sensitivity (Ebert and Prelec 2007) discount functions are weighted discount functions. As will be 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

---

3The 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 proposed in that paper are also weighted discount functions with decreasing impatience.
weighted discount function. This “weighting iteration argument” 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 groups of non-exponential discounters – as long as they belong to our class of weighted discounters.

The previously mentioned result that weighted discount functions exhibit decreasing impatience constitutes the main challenge in studying a group’s investment decision, which we approach in Section 3. Except for the trivial case where the group consists of a single member, in which case the weighted discount function is exponential, the decision problem becomes time-inconsistent (Strotz 1955). For that reason, in our study of group investment behavior, we find ourselves facing similar difficulties as the behavioral literature on timing (or stopping) decisions under time-inconsistency. We study the group’s decision problem within the intra-personal game framework (e.g, Phelps and Pollak 1968; Laibson 1997). When making its decision, the group takes into account that the group’s future self may not agree with the planned 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 mathematical result of 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 results hold for arbitrary payoff or utility functions, and 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 general
weighted discount functions, we show that investment behavior is characterized by a threshold strategy. This means that it is desirable to invest once the project value is at or beyond a certain level, and to wait if it is beneath that value.

We then focus on some special weighted discount functions to obtain various analytical results. In particular, investment behavior under generalized hyperbolic discounting (Loewenstein and Prelec 1992) comprises an example of investment behavior under weighted discounting. This important special case has not been solved in the literature yet, and thus our results also break new ground in the behavioral literature. In particular, we are the first to obtain results on investment behavior under continuously changing levels of impatience, as is implied by generalized hyperbolic, constant absolute decreasing impatience (CADI), or constant relative decreasing impatience (CRDI) discounting (Bleichrodt et al. 2009).

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 admit a nice analytical representation, like hyperbolic discounting – as some economic results may involve this weighting distribution. For the discount function being the one proposed by Harris and Laibson (2013), we recover the results of Grenadier and Wang (2007) on the investment behavior of a sophisticated entrepreneur. In particular, as we show, their results indeed correspond to those obtained for a binomial weighting distribution \( F \). Finally, for the weighted discount function being the exponential one, our Bellman system for the real options case reduces to a single Bellman equation – the one studied extensively in Dixit and Pindyck (1994).

In Section 5, we study the comparative statics of investment behavior (i.e., the level of the threshold). Higher investment cost, expected growth of the project value, and project value volatility all lead to a higher threshold (i.e., later investment), which generalizes extant results for exponential discounting to arbitrary weighted discount functions. Moreover, for the patience ranking of discount functions proposed by Quah and Strulovici (2013), we show that more patient agents invest later.
Finally, we study the comparative statics of investment behavior with respect to group diversity. The main result is that greater group diversity leads to *later* investment. Waiting longer with making a decision when disagreement is strong seems consistent with daily observations. Earlier investment, in contrast, is characterized by a lower threshold, and thus comes along with less risk-taking – a bird in the hand (the current NPV of the project) is worth more than two in the bush (a potentially higher project NPV). Thus, the result that greater group diversity leads to more risk-taking in investment may be surprising. Within the intra-personal game framework, time-inconsistency that results from group disagreement restricts the set of feasible investment strategies, which should lead to *less* risk-taking. However, the direct effect of greater group diversity leading to less discounting (the discount function elevation effect) dominates the latter effect of greater time-inconsistency. Overall, the empirical prediction is that more diverse groups are more patient in their decision, willing to wait for a yet more profitable time to invest.

2 Weighted discount functions

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

2.1 General definition of weighted discount functions

The following definition extends equation (1) 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 $F(r)$ (or a density function $f(r)$) concentrated on $[0, \infty)$ such that

$$h(t) = \int_0^\infty e^{-rt} dF(r) = \int_0^\infty e^{-rt} f(r) dr.$$  \hspace{1cm} (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 $r$.

### 2.2 Prevalence and examples of weighted discount functions

As an introductory example, consider the so-called pseudo-exponential discount function. It is given by $h(t) = \delta e^{-rt} + (1 - \delta)e^{-(r+\lambda)t}$, $\lambda > 0$, $r > 0$, $0 < \delta < 1$ (Ekeland and Lazrak 2006; Karp 2007), i.e., it is given by the weighted average of just two opinions. The corresponding weighting distribution $F$ in equation (2) is a step function and can be written as

$$ F(x) = \begin{cases} 
0, & x < r \\
\delta, & r \leq x < r + \lambda \\
1, & x \geq r + \lambda.
\end{cases} \tag{3} $$

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

**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 function. Brockett and Golden (1987) refer to the class of
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 $F$.

Increasing utility functions with derivatives that alternate in sign as the class that contains “all commonly used utility functions.” In analogy, we may just refer to the class of weighted discount functions as the class that contains “all commonly used discount functions.”

While every weighting distribution $F$ defines a discount function $h^F$ – and while most discount functions are weighted discount functions – both the weighting distribution and the weighted discount function being available in closed form is rather the exception than the rule. In particular, it may not be possible to compute the integral in Definition 1 explicitly and, likewise, for a given
weighted discount function $h^F$, it may not be possible to recover the weighting distribution $F$ in closed form. The results in this paper, however, do not depend on whether a closed form expression of either the discount function or the weighting distribution is available. For example, the CADI function in Bleichrodt et al. (2009) may not be obtained as the group discount function of any nice distribution function $F$, but we can still study its implications for investment behavior and draw inference on the diversity of the group it represents. Likewise, the discount function obtained from a Chi-Square distribution may not be available in closed-form, but nevertheless we can study the investment decisions of a group with Chi-square-distributed opinions over discount rates.

There are also some weighting distributions that lead to nice weighted discount functions (and vice versa). We already gave the example of the pseudo-exponential discount function, which corresponds to a binary weighting distribution. This important example is easily generalized to finitely many group members, whose opinions are multinomially distributed. Even though not mentioned by Weitzman (2001), the weighted discount function obtained from the Gamma distribution (with proper re-parametrization as below) coincides with that of Loewenstein and Prelec (1992). In particular, the generalized hyperbolic discount function with parameters $\alpha > 0, \beta > 0$ can be written as (the first expression repeats the original definition of Loewenstein and Prelec (1992))

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

(4)

where

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

(5)

denotes the density function of the Gamma distribution with parameters $k$ and $\theta$ and where $\Gamma(k)$ denotes the Gamma function evaluated at $k$, i.e., $\Gamma(k) = \int_0^\infty x^{k-1}e^{-x}dx$. A simpler example is given by the discount function studied in Mazur (1987) and Harvey (1995), which corresponds
to the special case of Loewenstein and Prelec (1992) where $\alpha = \beta$. In that case, equation (4) becomes

$$h(t) = \frac{1}{1 + \alpha t} = \int_{0}^{\infty} e^{-rt} \left( \frac{1}{\alpha} e^{-\frac{r}{\alpha}} \right) dr,$$

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

As a final example, one can obtain the uniform-uncertainty discount function of Souzou (1998) from a uniform distribution of opinions over an interval $[r, \bar{r}]$:

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

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

### 2.3 Greater group diversity

In this subsection, we establish a notion of comparative group diversity in time preferences. The main result is that a group has a more elevated discount factor than another group (i.e., discounts less heavily at any future time) if and only if it is more diverse. Greater group diversity will be
captured by a more dispersed weighting distribution $F$ of the corresponding weighted discount function $h^F$, formally defined using stochastic dominance.

Recall that a second-order stochastically dominating distribution is preferred, for example, by all decision makers whose utility functions are increasing and concave. These decision makers dislike decreases in mean and increases in variance. Accordingly, second-order stochastically dominated distributions have lower mean and a higher variance than the distribution in comparison, and at least one of the two inequalities must be strict. We will work with the following, more general stochastic dominance notion, which is due to Fishburn (1976, 1980).5

**Definition 2 (Infinity-stochastic dominance.)** $F$ dominates $G$ via infinity-stochastic dominance (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, i.e., $\text{sgn} \, u^{(n)}(x) = (-1)^{n+1}$ for any $n \in \mathbb{N}^+, x \in (0, \infty)$.

Infinity stochastic dominance is the “weakest” stochastic dominance order in the sense that any finite stochastic dominance order (e.g., first-order stochastic dominance, second-order stochastic dominance,...) implies infinite stochastic dominance. Defining group diversity through infinity stochastic dominance will make our notion of group diversity less restrictive, as explained after its definition.6

**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 $G \preceq_{\infty SD} F$.

Economically, greater group diversity may refer to a distribution of opinions with higher variance (the second-order stochastic dominance case), but it may also refer to a distribution of opinions with lower skewness (the third-order stochastic dominance case). Even more generally, we can leave the type of dispersion (higher variance, lower skewness, higher kurtosis, or some or all of

5See Fishburn’s papers for the equivalent definition based on the cumulative distribution functions and the so-called “iterated integrals”

6At the first sight, the reader may feel that the “direction” of $\preceq_{\infty SD}$ should just be the other way around. The notation we chose is in line with the one that is standard in the stochastic dominance literature, where $G \preceq_{\infty SD} F$ means that $F$ is preferred over $G$ because $F$ has less risk (dispersion) in the stochastic dominance sense. We reinterpret this by saying that $G \preceq_{\infty SD} F$ means that group $G$ is more diverse than group $F$. 

12
The latter property of a more elevated discount function is central to many results of this paper. Figure 1 above illustrates Proposition 7 through the fact that the pseudo-exponential discount function $h^G$ lies above the exponential discount function $h^F$ at all times. Indeed, in that example, $G$ is more diverse than $F$ in the sense of second-order stochastic dominance, which implies infinite stochastic dominance, and is thus sufficient for a more elevated discount function. In particular, the binary distribution $G$ in Figure 1 constitutes a mean-preserving spread (Rothschild and Stiglitz 1970) of the degenerate distribution $F$. The intuition behind the proof for this special case is that exponential functions are convex so that, by Jensen’s inequality, averaging two discount factors whose discount rates average to 5%, results in a discount factor larger than a single exponential discount factor with rate 5%.

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 (7) then says that group $G$ has a smaller discount rate if and only if it has a larger discount factor. In other words, in the exponential discounting case, comparing discount factors is equivalent to comparing discount rates. For general discount functions, however, comparing discount factors is not equivalent to comparing expected discount rates. For weighted discount functions, Theorem 2 establishes the necessary and sufficient condition for ordering discount factors. Ordering discount factors is equivalent to comparing the weighting distributions in the infinity-stochastic dominance sense or to comparing the diversity of the group’s opinions.
2.4 Decreasing impatience and other properties of weighed discount functions

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

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.

As regards the economic insight, the following result is similar to Proposition 1 in Jackson and Yariv (2014), who show that the discount functions of finite groups are present-biased.7

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

Proposition 1 illustrates the main difficulty that awaits us in Section 3 when studying the investment decisions of groups. Due to decreasing impatience, the group is in general time-inconsistent. The next result is on weighted discount functions that differ in their degree of decreasing impatience.

Proposition 2 (Greater Decreasing Impatience and Greater Group Diversity.) Consider weighted discount functions $h^F$ and $h^G$ with weighting distributions $F$ and $G$ and measures of decreasing impatience $P_F$ and $P_G$, respectively. Let the expectation of $F$ be no less than that of $G$. If $G$ is always more impatient than $F$, i.e., $P_G(t) \geq P_F(t)$ for all $t \geq 0$, then $G$ is more diverse.

7Our result generalizes and refines Weitzman's (2001) observation of time-varying discount rates for the special case where opinions $F$ are Gamma-distributed. Jackson and Yariv (2015) show that with any heterogeneity in time preferences, utilitarian aggregation necessitates a present bias. Gollier and Zeckhauser (2005) have noted that heterogeneity results in time-varying discount rates and time-inconsistency. Unlike these papers, we provide results on the concept of decreasing impatience as characterized by Prelec (2004), which is not cited by the other articles.
The assumption of $F$ having an equal or larger expectation means that the weighted average of group $F$'s discount \textit{rates} (rather than the group's discount factors) is equal or larger for $F$. Therefore, group $F$ on average discounts with a larger discount rate. The link to the discount function is as follows. It is easily verified that the negative of the expectation of $F$ can be computed as $(h^F)'(0)$. Therefore, an equal or larger expectation means that $h^F$ initially decreases more strongly. If, in addition, $h^F$ is less convex in the sense that its impatience is less decreasing, then $h^F$ must always lie below $h^G$. From Theorem 2, it follows that $h^G$ having more decreasing impatience and $G$ having less or equal expectation implies greater group diversity.

Finally, we show that greater patience, a novel concept defined and studied by Quah and Strulovici (2013), also implies greater group diversity.

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

$$\frac{h^G(t)}{h^F(t)}$$

is increasing in $t$.

**Proposition 3 (More patient groups are more diverse.)** More patient groups are more diverse.

While we believe that our results on weighted discount functions are interesting in their own right, we will illustrate their usefulness explicitly when studying the comparative statics of investment behavior in Section 5 with respect to group properties such as patience or diversity.

## 3 Time-consistent optimal stopping under weighted discounting

This section derives a general optimal stopping result under weighted discounting and sophistication without commitment, i.e., awareness of the time-inconsistency that results from decreasing impatience (Proposition 1)
3.1 Dynamics

Let \((\Omega, \mathcal{F}, \mathbb{P})\) denote a complete probability space which 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

\[
\frac{dX_t}{X_t} = b(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x_0
\]  

where the bounded function \(b\) describes the instantaneous conditional expected percentage change in \(X\) per unit of time and the bounded function \(\sigma\) is the instantaneous conditional standard deviation per unit time.\(^8\) \(X\) may describe the wealth of an agent gambling in casino, the quality of a current job offer, the project value in a real options setting, or the value of a stock or of the underlying of some derivative. 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 rule \(u\) defines a stopping time \(\tau^t_u\) after \(t\), via

\[
\tau^t_u = \inf\{s \geq t, u(s, X_s) = 1\}. 
\]  

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. Define the discount function

\(^8\)See Karatzas and Shreve (2006) for the usual technical conditions that need to be imposed on \(b\) and \(\sigma\). The results in this paper can be generalized to more complicated processes.
of self $t$ 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 $h_t(t) = 1$. If $X_t = x$, self $t$ seeks to maximize the weighted discounted payoff from its investment decision according to the stopping rule $u$:

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

(10)

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 the selves to prefer a different choice of $u$. In general, each self $t$ can only choose current (time-$t$) behavior. Plans when to stop (i.e., stopping times) made by self $t$ may be overturned 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:

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

(11)

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 constantly 0 or constantly 1. We call $u^{\epsilon,a}$ the $(\epsilon, a)$-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 (10) of the group as

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

$$= \mathbb{E}[h(\tau_u^0)G(X_{\tau_u^0})|X_0 = x]$$

$$= J(0, x; u).$$

(12)

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 $X_t = x$. 

17
of the process, and drop the time index from its objective functional

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

where $\tau_u = \inf\{s \geq 0, u(X_s) = 1\}$. With a slight abuse of notation, the stopping rule $u$ above is now independent of time $t$ and thus we write it as 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, i.e., 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.\(^9\)

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

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

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

The stopping rule $u^{\epsilon,a}$ in the above definition coincides with the equilibrium stopping rule $\hat{u}$ except for the (short) time interval $[t, t + \epsilon)$ for each $t \geq 0$. On that interval, $u^{\epsilon,a}$ is either constantly 0 or constantly 1. In particular, either $u^{\epsilon,1}$ or $u^{\epsilon,0}$ are different from the equilibrium strategy, i.e., the $u^{\epsilon,a}$ constitutes possible deviation strategies. If $\epsilon = 1$, the non-negativity of the ratio in (14) says that every future self (characterized by wealth $x$) prefers $\hat{u}$ over deviating according to $u^{\epsilon,1}$ or $u^{\epsilon,0}$. For an arbitrary short interval $\epsilon$ it follows that deviation is unattractive at the current value of $x$.\(^{10}\)

\(^9\)Here we follow the standard in the economics literature and focus on stationary Markov-perfect equilibria (e.g., Grenadier and Wang 2007, p.16; Harris and Laibson 2013, p.215). Stopping rules only depend on the current process value and all future selves pursue the same strategy.

\(^{10}\)Note that at each single point in time, there is only one deviation strategy. If the equilibrium stopping rule is “stop” at current wealth, then the deviation strategy is “continue”, and vice versa. Therefore, the above definition indeed considers all possible deviation strategies at a given process value $x$. Our equilibrium definition is consistent with those from time-inconsistent control problems (see Ekeland and Pirvu 2008; Björk et al. 2014) when interpreting a stopping rule as a binary control.
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.\textsuperscript{11}

**Theorem 3 (Equilibrium Characterization)** Consider the performance functional (13) with weighted discount function
\[ h(t) = \int_0^\infty e^{-rt} dF(r), \]
and
\[ V(x) = \int_0^\infty w(x; r) dF(r). \]
Let
\[ A = \frac{1}{2} \sigma(x)^2 \frac{\partial^2}{\partial x^2} + b(x) x \frac{\partial}{\partial x} \]
and suppose that \((V, w, \hat{u})\) solves
\[ \max \{ AV(x) - \int_0^\infty rw(x; r) dF(r), G(x) - V(x) \} = 0, \] \hspace{1cm} (15)
\[ \hat{u}(x) = \begin{cases} 1 & \text{if } V(x) = G(x), \\ 0 & \text{otherwise}, \end{cases} \] \hspace{1cm} (16)
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 \(V(x)\), i.e.,
\[ V(x) = \mathbb{E}[h(\tau_\hat{u}) G(X_{\tau_\hat{u}})]. \]

To interpret Theorem 3, let us call
\[ \mathcal{S} = \{ x \in [0, \infty) : V(x) = G(x) \} \]
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
\[ \hat{u}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{S}, \\ 0 & \text{if } x \in \mathcal{C} \end{cases} \] \hspace{1cm} (17)
and the corresponding stopping time given by
\[ \tau_\hat{u} = \inf\{s \geq 0 : X_s \in \mathcal{S} \} \]

\textsuperscript{11}The value function \(V\) in the theorem must satisfy some regularity conditions that are not restrictive from an economic point of view.
with \( X_0 = x \).

Equations (15)–(16) constitute the so-called Bellman system, a system of coupled equations. Note that (16) is an equation (rather than a definition of \( \hat{u} \)) since its right hand side involves \( V \) that 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 \( x \in S \) where the agent will stop and the values \( x \in C \) where she 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 \( V(x) = w(x; r) \), i.e., the value function \( V \) is given by that member's expected discounted payoff and equation (15) becomes the well-known Bellman equation (e.g., Dixit and Pindyck 1994)

\[
\max\{AV(x) - rV, G(x) - V(x)\} = 0.
\]

Note that the above equation is independent of \( \hat{u} \) and can be solved once the model primitives are given. Then, \( \hat{u} \) is obtained immediately from equation (16).\(^{12}\)

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

\[
V(x) = \mathbb{E}[h(\tau_{\hat{u}})G(X_{\tau_{\hat{u}}})] = \mathbb{E}\left[\int_{0}^{\infty} e^{-r\tau_{\hat{u}}} dF(r)G(X_{\tau_{\hat{u}}})\right]
\]

\[
= \int_{0}^{\infty} \mathbb{E}[e^{-r\tau_{\hat{u}}}G(X_{\tau_{\hat{u}}})]dF(r)
\]

\[
= \int_{0}^{\infty} w(x; r)dF(r).
\]

This calculation shows that the weighted form of the discount function carries over to the value

\(^{12}\)In other words, in that case equation (16) 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.
function. Moreover, it clarifies the motivation behind defining the function $w(x; r)$. To derive the Bellman equation (15), 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

$$Aw(x; r) - rw(x; r) = 0, x \in C$$

with the boundary conditions $w(0; r) = 0$ and $w(x; r)|_S = G(x)|_S$. 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, but only the value function itself is. Since $V$ is the weighted average of the $w$, by integrating the differential equation of $w$ against $F$, $V$ must satisfy

$$\int_0^\infty (Aw(x; r) - rw(x; r)) dF(r) = 0, x \in C$$

$$\iff \left( A \int_0^\infty w(x; r) dF(r) \right) - \int_0^\infty rw(x; r) dF(r) = 0, x \in C$$

$$\iff AV(x) - \int_0^\infty rw(x; r) dF(r) = 0, x \in C.$$

Equation (15) is then obtained by comparing the value of continuation and stopping.

We close this section by considering the example of the pseudo-exponential discount function. In that case, the group consists of two group members with discount rates $r$ and $r + \lambda > r$ whose weights in the decision process are $\delta$ and $1 - \delta$, respectively; see equation (3). In that case, we
can recover the equilibrium stopping rule from the following four equations:

\[
\max \{ AV(x) - (\delta w(x; r) + (1 - \delta)(r + \lambda)w(x; r + \lambda)), G(x) - V(x) \} = 0,
\]

\[
w(x; r) = \mathbb{E}[e^{-r\tau_u}G(X_{\tau_u})],
\]

\[
w(x; r + \lambda) = \mathbb{E}[e^{-(r+\lambda)\tau_u}G(X_{\tau_u})];
\]

\[
\hat{u}(x) = \begin{cases} 
1 & \text{if } \delta w(x; r) + (1 - \delta)w(x; r + \lambda) = G(x), \\
0 & \text{otherwise}, 
\end{cases}
\]

along with the boundary condition \(V(0) = \max\{G(0), 0\}\). We note that, in this case, the last equation is indeed coupled with the other ones.

### 4 Dynamic investment under weighted discounting

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

#### 4.1 Economic setup

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

\[
\frac{dX_t}{X_t} = bdt + \sigma dW_t. \tag{19}
\]

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

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

To ensure the well-posedness of the problem, let

$$b < \inf\{ r \in [0, \infty) : F(r) > 0 \}. \quad (20)$$

When $F$ is a step function jumping at $r = r_0$ (the exponential discounting case), condition (20) is reduced to the standard condition $b < r_0$ (e.g., Dixit and Pindyck 1994, p.141).\(^\text{13}\)

### 4.2 Investment behavior under weighted discounting

In this subsection, we solve the Bellman system described in Theorem 3 explicitly when $X$ follows a geometric Brownian Motion (so that $A = \frac{1}{2} \sigma^2 x^2 \frac{d^2}{dx^2} + bx \frac{d}{dx}$) and when $G(x) = x - I$, for arbitrary weighted discount functions.

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

$$\frac{1}{2} \sigma^2 x^2 w_{xx}(x; r) + bxw_x(x; r) - rw(x; r) = 0, \quad (21)$$

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

\(^\text{13}\)Economically, the more general condition (20) 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 condition (20).
differential equation (21) is given by

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

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

\[ \frac{1}{2} \sigma^2 \mu^2 + (b - \frac{1}{2} \sigma^2) \mu - r = 0, \quad \text{i.e.,} \quad \mu(r) = \frac{-(b - \frac{1}{2} \sigma^2) + \sqrt{(b - \frac{1}{2} \sigma^2)^2 + 2\sigma^2 r}}{\sigma^2}. \quad (22) \]

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

\[ V(x) = \int_0^\infty \left( \frac{x}{x^*} \right)^{\mu(r)} dF(r)(x^* - I). \quad (23) \]

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

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

It follows that

\[ x^* = \frac{A}{A - 1}, \quad (24) \]

where

\[ A = \int_0^\infty \mu(r)dF(r). \quad (25) \]

In the appendix we prove:

**Proposition 4 (Investment behavior under weighted discounting.)** With the notation above, the triple

\[ (V^e(x), w^e(x; r), \hat{u}(x)) = \begin{cases} (V(x), w(x; r), 0) & \text{if } x < x^*, \\ (x - I, x - I, 1) & \text{otherwise} \end{cases} \]

solves the Bellman system (15)–(16) with its boundary condition. Hence, \( \hat{u} \) is an equilibrium
stopping rule and $V^e$ is the corresponding value function.

Note that the expression for the triggering threshold, equation (24), involves the weighting distribution $F$, but not the corresponding discount function $h^F$ directly. Therefore, if one studied investment under, say, hyperbolic discounting without involving the weighted representation, it seems difficult to arrive at the solution, which involves the weighting distribution. This illustrates the mathematical virtue of writing even discount functions that admit a nice analytical representation (like hyperbolic discounting) in the weighted form.

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

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

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

We close this section by presenting analytical solutions for real-option investment under hyperbolic and pseudo-exponential discounting. For the generalized hyperbolic discount function from equation (4), the triggering threshold $x^*$ is given by

$$x^* = \frac{\int_0^\infty \mu(r) \frac{r^{\frac{\beta}{\alpha}} - 1}{\alpha^{\frac{\beta}{\alpha}} \Gamma\left(\frac{\beta}{\alpha}\right)} dr}{\int_0^\infty \mu(r) \frac{r^{\frac{\beta}{\alpha}} - 1}{\alpha^{\frac{\beta}{\alpha}} \Gamma\left(\frac{\beta}{\alpha}\right)} dr - 1} I. \quad (26)$$

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

For the pseudo-exponential discount function \( h(t) = \delta e^{rt} + (1 - \delta) e^{(r + \lambda)t}, \lambda > 0, r > 0, 0 < \delta < 1, \) the triggering boundary \( x^* \) is given by

\[
x^* = \frac{\delta \mu(r) + (1 - \delta) \mu(\lambda + r)}{\delta \mu(r) + (1 - \delta) \mu(\lambda + r) - 1}.
\]

The representation of the triggering boundary \( x^* \) for pseudo-exponential discounting is the same as the one obtained by Grenadier and Wang (2007) who considered an investment timing problem in the context of stochastic quasi-hyperbolic discounting.\(^{14}\) Harris and Laibon (2013) noted that the pseudo-exponential discount function is the expectation of the stochastic quasi-hyperbolic discount function. They argue that the two types of discount functions lead to the same equilibrium in a consumption-savings model. The above example on pseudo-exponential discounting shows that this intuition carries through to an investment timing problem.

5 Comparative statics of investment behavior

In this section, we study the investment behavior under weighted discounting in more detail. We generalize some well-known comparative static results from exponential discount functions to arbitrary weighted discount functions. In particular, we show how the triggering boundary \( x^* \) in equation (24), which describes investment behavior in our model, changes as the economic environment changes. Then, we obtain new results by turning to comparative statics with respect to decreasing impatience or group diversity, for example.

\(^{14}\)Grenadier and Wang (2007) employ a continuous-time version of quasi-hyperbolic discounting. In this model, the quasi-hyperbolic discount function takes a stochastic form, i.e., at time \( t \) the agent’s self \( n \) applies the discount factor \( D_n(t, s) \) given by

\[
D_n(t, s) = \begin{cases} 
 e^{-r(s-t)} & \text{if } s \in [t_n, t_{n+1}), \\
 \delta e^{-r(s-t)} & \text{if } s \in [t_{n+1}, \infty), 
\end{cases}
\]

where \( \{t_n\}_{n \geq 1} \) is a sequence of arrival times which follow a Poisson process with intensity \( \lambda \) and which are independent of the payoff process \( X \).
5.1 The impact of the economic environment on investment behavior

Before we turn to the impact of discounting behavior on investment, we note that the triggering boundary $x_*$ in equation (24) depends on the project return rate and volatility, determined by the function $\mu(r)$ in equation (22), 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 5 (Comparative statics for entry cost and project dynamics.)** Suppose that the factor $A$ defined by equation (25) is finite.\(^{15}\) Then,

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

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

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

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

5.2 The impact of group diversity 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 greater group diversity, as defined in Section 2 via stochastic dominance deteriorations, leads to later investment.

**Proposition 6 (More diverse groups invest later.)** Suppose $F$ and $G$ are weighting distributions. Then,

$$G \preceq_{\infty SD} F \Rightarrow x^G_* \geq x^F_*,$$

where $x^G_*(x^F_*)$ is the triggering boundary defined by (24) with the weighting distribution $G$ ($F$).

\(^{15}\)If $A = \infty$, then $x_* = I$, which means that the agent invests in the project whenever the payoff process is greater than the entry cost.
In general, under greater group diversity it is more difficult to find a consensus investment strategy. In other words, greater group diversity comes along with more restrictions on the equilibrium stopping rule. Therefore, as explained in the introduction, one might have expected that greater group diversity leads to less risk-taking and thus earlier stopping. However, this effect is dominated by the fact that greater group diversity is equivalent to a more elevated discount factor (Theorem 2). From Theorem 2 and Proposition 6 it follows that

**Proposition 7 (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 triggering boundary defined by (24) with weighting distribution $G(F)$.

Note that Proposition 6 makes an assumption on the weighting distribution $F$ while Proposition 7 makes an assumption on the discount function $h^F$, and both proposition’s implication is on the triggering boundary $x_*$ in equation (24). 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 6 is relatively simple while a direct proof of Proposition 7 is quite complicated. We give such a direct proof in the appendix, because a comparison with the simple proof of Proposition 6 illustrates well 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 1 in the appendix offers an alternative and less intuitive expression for $x_*$ in equation (24) that, however, involves the original discount function $h^F$ rather than the corresponding weighting distribution $F$. 

28
Table 1: Parameters used for the scenarios in Figure 2

<table>
<thead>
<tr>
<th></th>
<th>$r$</th>
<th>$\lambda$</th>
<th>$\delta$</th>
<th>$M$</th>
<th>$SD$</th>
<th>$S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td>0.0825</td>
<td>0</td>
<td>1</td>
<td>0.0825</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FSD</td>
<td>0.0525</td>
<td>0</td>
<td>1</td>
<td>0.0525</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SSD</td>
<td>0.03</td>
<td>0.045</td>
<td>0.5</td>
<td>0.0525</td>
<td>0.0225</td>
<td>0</td>
</tr>
<tr>
<td>TSD</td>
<td>0.0075</td>
<td>0.0562</td>
<td>0.2</td>
<td>0.0525</td>
<td>0.0225</td>
<td>-1.5000</td>
</tr>
</tbody>
</table>

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 (3). The parameters of the weighting distribution are shown in the first three columns, while the remaining three columns show its mean $M$, standard deviation $SD$, and standardized skewness $S$. The scenarios are chosen such that the weighting distribution in scenario Benchmark is less first-order diverse than scenario FSD; which in turn is less second-order diverse than scenario SSD; which in turn is less third-order diverse than scenario TSD.

5.3 The impact of group diversity on investment behavior: An example

In this subsection, we illustrate the quantitative impact of group diversity on the triggering boundaries through an example. To this end, let us consider pseudo-exponential discounting once more. The decision-making group consists of two members with discount rates $r$ and $r + \lambda > r$ and whose weights in the decision process are $\delta$ and $1 - \delta$, respectively; see equation (3). The fact that the weighting distribution is binary allows for a simple and intuitive characterization of stochastic dominance relationships through statistical moments. In particular, the weighting distribution with parameters $\delta$, $r$, and $r + \lambda$ can be re-parametrized in terms of its mean $M$, standard deviation $SD$, and standardized skewness $S$. Moreover, it can be shown that a binary distribution first- (second-, third-) order dominates another binary distribution with lower mean (greater variance, lower skewness), given that the other two moments are the same. In the following, we can thus conduct comparative statics with respect to the moments of the weighting distribution, rather than having to deal with 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 which 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, 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, but note that this result can be proven (and interpreted) in terms of our result that greater group diversity leads to later investment: In the terminology of stochastic dominance, the FSD weighting distribution is first-order “riskier”, i.e., the group in the FSD scenario is first-order more diverse.

Next, the weighting distribution of the FSD scenario second-order stochastically dominates that of the SSD scenario. Equivalently, the SSD weighting distribution is second-order more diverse than that of the FSD scenario. This is because the SSD weighting distribution has a greater standard deviation than the weighting distribution of the FSD scenario, but identical mean and skewness. Greater second-order diversity also leads to later investment: \( x_{*}^{SSD} > x_{*}^{FSD} \). Finally, the figure illustrates that a weighting distribution with lower skewness (increased third-order diversity) results in later investment: \( x_{*}^{SSD} > x_{*}^{FSD} \). This means that if one of the group members has small weight \( \delta \), but a strong opinion in favor of a very small discount rate, this results in later investment compared to when the weights on group members are more symmetric – even if the group mean and standard deviation of the discount rate are the same.

### 5.4 The impact of patience and decreasing impatience on investment behavior

The result that greater group diversity leads to later investment is also useful to prove a number of other results on investment behavior. This is because greater group diversity is the weakest property that we defined in Section 2. Therefore, first, Proposition 2 and Proposition prop:SDLT yield the following result.

**Proposition 8 (Decreasing impatience and later investment.)** Consider weighted discount
Figure 2: The impact of group diversity on investment behavior

Notes. The figure illustrates that greater group diversity leads to a more risky investment decision (i.e., to stopping later, at a higher threshold level). $x^*_i$ denotes the threshold for scenario $i$ ($i = B, FSD, SSD, TSD$) which is defined in Table 1). The figure also plots the value function for each scenario, whose intersection with the payoff schedule determines the threshold. $x^*_B < x^*_FSD$ means that we have later stopping under exponential discounting for smaller discount rates (which comprise a first-order stochastic dominance deterioration). $x^*_FSD < x^*_SSD$ means that a group of two individuals who disagree about the discount rate stops later than a single individual whose discount rate corresponds to the group average. $x^*_SSD < x^*_TSD$ illustrates that a group who is more diverse in the sense that there is one member with low weight in the decision, but an extremely low opinion about the appropriate discount rate, stops later than a group whose disagreement is more symmetric. The project drift and volatility assumed are $b = 0.01$ and $\sigma = 0.2$, respectively. The project entry cost is set to $I = 1$. 
function \(h^F\) and \(h^G\) with weighting distributions \(F\) and \(G\) and measures of decreasing impatience \(P_F\) and \(P_G\), respectively. Let the expectation of \(F\) be no less than that of \(G\) and \(P_G(t) \geq P_F(t)\) \(\forall t \geq 0\). Then \(x^G_* \geq x^F_*\), where \(x^G_*(x^F_*)\) is the triggering boundary defined by (24) with discount function \(h^G(h^F)\).

Second, Proposition 3 implies that more patience implies greater group diversity. Therefore, from Proposition 7 we further obtain the following result.

**Proposition 9 (More patient groups invest later.)** For weighted discount functions \(h^F\) and \(h^G\) with weighting distributions \(F\) and \(G\) we have

\[
h^G \succeq_p h^G \implies x^G_* \geq x^F_*,
\]

where \(x^G_*(x^F_*)\) is the triggering boundary defined by (24) with discount function \(h^G(h^F)\).

### 5.5 The impact of discount function parameters on investment behavior

Finally, we explain how one can study the impact of other properties of the discount function, that are captured by its parameters, on investment behavior. For sake of concreteness, we do this with reference to generalized hyperbolic discounting; see equation (4). From Prelec (2004, p.524), we know that the parameter \(\alpha\) captures decreasing impatience. To the best of our knowledge, a clear economic interpretation of the parameter \(\beta\) has not yet been proposed\(^{16}\), but we give one here. \(\beta\) is equal to the expectation of the generalized hyperbolic weighting distribution (the Gamma distribution with parameters as in equation (4)). That is, \(\beta\) identifies the group’s weighted mean opinion about the appropriate discount rate. This once more illustrates the added value that may arise from interpreting discounting behavior through both the weighted discount function and its weighting distribution, as proposed in this paper.

\(^{16}\)Loewenstein and Prelec (1992) merely use the parameter for functional flexibility. Prelec (2004) notes that \(\beta\) controls time preference, but this is only true once \(\alpha\) has been fixed. In other words, many different values of \(\beta\) can imply same time preference, and vice versa.
Figure 3: The impact of decreasing impatience, $\alpha$, and the weighted mean of the group discount rate, $\beta$, on the triggering boundary under generalized hyperbolic discounting (Loewenstein and Prelec 1992)

\[ h(t; \alpha, \beta) = \frac{1}{(1+\alpha t)^{\frac{\beta}{\alpha}}} \] with $\alpha, \beta > 0$. Then the triggering boundary $x_*$ defined by equation (24)

(i) increases with the degree of decreasing impatience, $\alpha$.

(ii) decreases with the group weighted mean opinion about the discount rate, $\beta$.

Figure 3 illustrates the quantitative impact of the parameters $\alpha$ and $\beta$ on the triggering boundary $x_*$. We see that greater decreasing impatience leads to later investment (left panel of Figure 3). Since the weighting distribution mean is fixed to $\beta = 0.14$, Proposition 2 and Theorem 2 apply, saying that the discount factor, at any date $t$, is increasing with greater decreasing impatience $\alpha$. Proposition 7 then says that the larger discount factor implies later stopping (larger $x_*$). As
regards the parameter $\beta$, we see that a greater mean of the weighting distribution leads to earlier stopping ($\text{lower } x^*_s$). This latter observation is intuitively consistent with our result that less group diversity (and recall that a first-order stochastic dominance “risk” decrease is an instance of less group diversity) leads to earlier stopping.

## 6 Conclusion

We have introduced and analyzed a new class of discount functions, the class of *weighted discount functions*, which includes all of the most commonly used discount functions. Weighted discount functions may describe: the discounting behavior of groups; uncertainty about what discount rate to use; “behavioral” time preferences such as hyperbolic discounting; and, through an iteration argument, all of these things simultaneously.

We have proved 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 group 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 more diverse groups wait longer with investing, thereby taking greater risk than less diverse groups. This result is a consequence of an optimal stopping result for arbitrary weighted discount functions. The general result likewise clarifies investment behavior in a single-agent setting where the agent is, for example, a hyperbolic discounter who is aware of his time-inconsistency, but lacks commitment.

On a more general level, our analysis shows that exploiting the link between behavioral and collective time preferences can offer valuable insights. The study of the class of weighted discount functions necessarily makes implications for either of the two fields, and shows how the literature streams on behavioral and group decision making can benefit from one another.
A Proofs

A.1 Proof of Theorem 2

Suppose first that \( G \preceq_{\infty \text{SD}} F \). Note that \( u(x) := -e^{-xt} \) defines a mixed risk averse utility function (i.e., is increasing with derivatives that alternate in sign) on \([0, \infty)\) for each fixed \( t > 0 \). Therefore, Definition 2 yields that \( h_F(t) \leq h_G(t), \forall t \geq 0 \).

Now suppose that \( h_F(t) \leq h_G(t) \) for all \( t \geq 0 \). For any mixed risk averse utility function \( u \), 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

\[
  u(x) = \int_a^x u'(t)dt + u(a) \\
  = \int_a^x \int_0^\infty e^{-ts}dF_u(s)dt + u(a) \\
  = \int_0^\infty \int_a^x e^{-ts}dtdF_u(s) + u(a) \\
  = \int_0^\infty \frac{1}{s}(e^{-as} - e^{-xs})dF_u(s) + u(a).
\]

Hence,

\[
  \int_0^\infty u(x)dG(x) = \int_0^\infty \int_0^\infty \frac{1}{s}(e^{-as} - e^{-xs})dF_u(s)dG(x) + u(a) \\
  = \int_0^\infty \frac{1}{s}(e^{-as} - \int_0^\infty e^{-xs}dG(x))dF_u(s) + u(a) \\
  = \int_0^\infty \frac{1}{s}(e^{-as} - h_G(s))dF_u(s) + u(a).
\]

Comparing with the analogous expression for \( F \), because \( h_F(t) \leq h_G(t) \) for all \( t \geq 0 \) it follows that \( G \preceq_{\infty \text{SD}} F \). This completes the proof.
A.2 Proof of Proposition 1

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

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

so that

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

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

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

which concludes the proof.

A.3 Proof of Proposition 2

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

$$g_F(0) = -\frac{(h^F)'(0)}{h^F(0)} \geq -\frac{(h^G)'(0)}{h^G(0)} = g_G(0). \quad (27)$$
Since \( P_i = -\frac{g_i'}{g_i} = -(\ln g_i)' \), \( i = F, G \), the assumption that \( F \) exhibits less DI than \( G \), i.e., \( P_F(t) \leq P_G(t) \forall t \geq 0 \), can be restated as

\[
(\ln g_F(t))' \geq (\ln g_G(t))'.
\]

Therefore, (27) and (28) together imply that \( \ln g_F(t) \geq \ln g_G(t) \), \( \forall t \geq 0 \), which is equivalent to

\[
g_F(t) \geq g_G(t), \quad \forall t \geq 0,
\]

Since \( g_i = -(\ln h_i)' \) and \( h_i(0) = 1 \), \( i = F, G \), we have \( \ln h_F(t) \leq \ln h_G(t) \), and the result follows from Theorem 2.

**A.4 Proof of Proposition 3**

The assumption that \( h_F \preceq_P h_G \) says that \( \frac{h_G}{h_F} \) is increasing. Therefore, the fact that \( \frac{h_G(0)}{h_F(0)} = 1 \) yields that \( h_F(t) \leq h_G(t) \), \( \forall t \geq 0 \), and the result follows from Theorem 2.

**A.5 Proof of Theorem 3**

We have proven in equation (18) that

\[
\mathbb{E}[h(\tau_u)G(X_u)] = \int_0^\infty w(x; r) dF(r) = V(x).
\]

It now suffices to show that \( \hat{u} \) is an equilibrium stopping rule, namely, that it satisfies equation (14) subject to condition (11) in Definition 7.

In the case of \( X_0 = 0 \), due to the boundedness of \( b \) and \( \sigma \), 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 \( J(x; w^{(a)}) = G(x) \). Due to the Bellman equation (15), we have \( V(x) \geq G(x) \). This shows that equation (14) is satisfied.
We prove that the triple \((A)\) Proof of Proposition 4 where the inequality follows from the Bellman equation (15). This completes the proof.

Therefore, \(\liminf_{\epsilon \to 0} \frac{J(x; \hat{u}) - J(x; u^{c,0})}{\epsilon} = \liminf_{\epsilon \to 0} \frac{V(x) - \mathbb{E}[V(x)]}{\epsilon} + \liminf_{\epsilon \to 0} \int_{0}^{\infty} \frac{(1 - e^{-r\epsilon})}{\epsilon} \mathbb{E}[w(x; r)]dF(r) = -(AV)(x) + \int_{0}^{\infty} rw(x; r)dF(r) \geq 0.\)

where the inequality follows from the Bellman equation (15). This completes the proof.

**A.6 Proof of Proposition 4**

We prove that the triple \((V^{c}(x), w^{c}(x; r), \hat{u})\) defined in Proposition 4 indeed solves the Bellman equation with its boundary condition.

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

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 (21), we have that

$$\int_0^\infty \left[ \frac{1}{2} \sigma^2 x^2 w_{xx}(x; r) + bxw_x(x; r) \right] dF(r) = \int_0^\infty rw(x; r)dF(r)$$

$$\iff \frac{1}{2} \sigma^2 x^2 V_{xx}(x) + bxV_x(x) - \int_0^\infty rw(x; r)dF(r) = 0.$$
As $\mu$ is a concave function and because $\mu(r) > 1$ for all $r > b$, we have
\[
\int_0^\infty (r - b) \frac{\mu(r) dF(r)}{\int_0^\infty \mu(r) dF(r) - 1} dF(r) \geq (r_m - b) \frac{\mu(r_m)}{\mu(r_m) - 1},
\]
where $r_m = \int_0^\infty r dF(r)$. Finally, it remains to prove that
\[
(r_m - b) \frac{\mu(r_m)}{\mu(r_m) - 1} \geq r_m,
\]
which is equivalent 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 (33). This completes the proof.

A.7 Proof of Proposition 5

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 $\mu$, we only need to check the monotonicity of $\mu$ with respect to $b$ and $\sigma^2$.

For any fixed $r > b$, we redefine the function $\mu(r)$ defined by equation (22) as the function of $b$ and $\sigma^2$ by $\nu(b, \sigma^2)$. By simple calculation, we have
\[
\frac{\partial \nu}{\partial b} = \frac{1}{\sigma^2} \left( -1 + \frac{b - \frac{1}{2} \sigma^2}{\sqrt{(b - \frac{1}{2} \sigma^2)^2 + 2 \sigma^2 r}} \right) \leq 0 \text{ and }
\frac{\partial \nu}{\partial \sigma^2} = \frac{1}{\sigma^4 \sqrt{(b - \frac{1}{2} \sigma^2)^2 + 2 \sigma^2 r}} \left( b \sqrt{(b - \frac{1}{2} \sigma^2)^2 + 2 \sigma^2 r} + \frac{1}{2} \sigma^2 b - b^2 - \sigma^2 r \right).
\]
If $b \leq 0$, it is easy to see that $\frac{\partial \nu}{\partial \sigma^2} \leq 0$. If $b > 0$, since $b < r$, after some algebra we have

$$b\sqrt{(b - \frac{1}{2} \sigma^2)^2 + 2 \sigma^2 r + \frac{1}{2} \sigma^2 b - b^2 - \sigma^2 r} \leq b(b + \frac{1}{2} \sigma^2) + \frac{1}{2} \sigma^2 b - b^2 - \sigma^2 r$$

$$= \sigma^2(b - r) < 0.$$ 

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

**A.8 Proof of Proposition 6**

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

$$x^X_* = \frac{\mathbb{E}[\mu(X)]}{\mathbb{E}[\mu(X)] - 1} - 1$$

and

$$x^Y_* = \frac{\mathbb{E}[\mu(Y)]}{\mathbb{E}[\mu(Y)] - 1} - 1,$$

where $\mu$ is defined by equation (22). The main idea of the proof is to note that $\mu$ satisfies $\text{sgn } \mu^{(n)} = (-1)^{n+1}$ for all $n \in \mathbb{N}^+$, i.e. $\mu(r)$ can be interpreted as a mixed risk averse “utility function over discount rates.” Then, by definition of infinity stochastic dominance, $F^Y \preceq_{\infty SD} F^X$ implies $\mathbb{E}[\mu(X)] \geq \mathbb{E}[\mu(Y)]$. This yields the conclusion and completes the proof.

**A.9 Proof of Proposition 7**

The result follows from Proposition 6 and Theorem 2, but here we give the direct proof motivated in the main text. We first prove the following lemma, which provides a semi-analytic representation of the triggering boundary $x_*$ in terms of the weighted discount function $h^F$ (rather than in terms of its weighting distribution $F$, as does equation (24)).

**Lemma 1 (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 triggering boundary \( x_* \) in equation (24). Then

\[
x_* = \max\left\{ \frac{-2(b - \frac{1}{2}\sigma^2)}{\sigma^2}, 0 \right\} + B \max\left\{ \frac{-2(b - \frac{1}{2}\sigma^2)}{\sigma^2}, 0 \right\} + B - 1
\]

(34)

where

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

(35)

Proof of the Lemma. After some algebra, the factor \( A \) defined by equation (25) can be written as

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

(36)

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

\[
\sqrt{r + C} = \frac{1}{2\sqrt{C}} \int_0^r \frac{1}{\sqrt{1 + \frac{1}{C} s}} ds + \sqrt{C}.
\]

(37)

Moreover,

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

(38)

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

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

Next, we plug equation (38) into equation (37) to obtain

\[
\sqrt{r + C} = \frac{1}{2\sqrt{\pi}} \int_0^r \int_0^\infty t^{-\frac{1}{2}}e^{-Ct}e^{-st} ds dt + \sqrt{C}
\]

\[
= \frac{1}{2\sqrt{\pi}} \int_0^\infty t^{-\frac{1}{2}}e^{-Ct} \int_0^r e^{-st} ds dt + \sqrt{C}
\]

\[
= \frac{1}{2\sqrt{\pi}} \int_0^\infty t^{-\frac{1}{2}}e^{-Ct}(1 - e^{-rt}) dt + \sqrt{C},
\]

42
and hence

\[
\int_0^\infty \sqrt{C + rdF(r)} = \int_0^\infty \frac{1}{2\sqrt{\pi}} \int_0^\infty t^{-\frac{3}{2}} e^{-Ct(1 - e^{-rt})} dt dF(r) + \sqrt{C}
\]

\[
= \frac{1}{2\sqrt{\pi}} \int_0^\infty t^{-\frac{3}{2}} e^{-Ct(1 - \int_0^\infty e^{-rt} dF(r))} dt + \sqrt{C}
\]

\[
= \frac{1}{2\sqrt{\pi}} \int_0^\infty t^{-\frac{3}{2}} e^{-Ct(1 - h(t))} dt + \sqrt{C}.
\]

Finally, we plug the last equality into equation (36) to obtain

\[
A = \frac{-(b - \frac{1}{2}\sigma^2)}{\sigma^2} + \frac{|b - \frac{1}{2}\sigma^2|}{\sigma^2} + \frac{1}{\sqrt{2\pi}\sigma} \int_0^\infty t^{-\frac{3}{2}} e^{-\frac{(b - \frac{1}{2}\sigma^2)^2}{2\sigma^2}} t(1 - h(t)) dt
\]

\[
= \max\left\{ \frac{-2(b - \frac{1}{2}\sigma^2)}{\sigma^2}, 0 \right\} + \frac{1}{\sqrt{2\pi}\sigma} \int_0^\infty t^{-\frac{3}{2}} e^{-\frac{(b - \frac{1}{2}\sigma^2)^2}{2\sigma^2}} t(1 - h(t)) dt,
\]

which completes the proof of the lemma.

We now turn to the proof of Proposition 7. From equation (35) in Lemma 1 it follows, because \( h^F \geq h^G \), that \( B_F \leq B_G \) with \( B_i, i = F, G \) defined in equation (35). Then, from equation (34) in Lemma 1 it follows that the triggering boundary is decreasing in the factor \( B \), which completes the proof.

### A.10 Proof of Proposition 8

The result is an immediate consequence of Proposition 2 and Proposition 6.

### A.11 Proof of Proposition 9

Proposition 3 states that more patience implies greater group diversity. The claim thus follows from Proposition 6.

### A.12 Proof of Proposition 10

Noting that

\[
h(t) = e^{-\int_0^t \frac{a}{1+as} ds},
\]

43
it is easy to see that for $\beta \leq \beta_1$ and $\alpha \leq \alpha_1$ the functions

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

and

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

are increasing in $t$. This yields

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

Then, the properties of the triggering boundaries (statements (i) and (ii)) follow from Proposition 9.

### 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 bounded rationality. For example, if $F(r; i)$ is Gamma-distributed, we know that individual $i$ is a generalized hyperbolic discounter.

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

$$\bar{h}(t) = \int_0^\infty h(t; i) dG(i). \quad (39)$$

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

**Proposition 11** 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:

$$\int_0^\infty e^{-rt} d\bar{F}(r) = \int_0^\infty e^{-rt} \left( \int_0^\infty F(r; i) dG(i) \right)$$

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

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

$$= \bar{h}(t).$$

We close this section with a simple example that illustrates the above notation and analyzes the investment behavior of a group of non-exponential discounters within the real options framework of Section 4. 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 (39) is the weighting distribution of a pseudo-exponential discount function; see equation (3). Moreover, suppose that both members are hyperbolic discounters with the parametrization of Mazur (1987) and Harvey (1995) and
parameters $\alpha_1$ and $\alpha_2$, respectively:

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

From equation (6), these are weighted discount functions with exponentially distributed weighting distributions. That is, in the above notation, $F(r; i) \equiv F(r; \alpha_i) = 1 - e^{-\frac{r}{\alpha_i}}$ and the corresponding density functions are given by $f(r; i) \equiv f(r; \alpha_i) = \frac{1}{\alpha_i} e^{-\frac{r}{\alpha_i}}$. Therefore, by Proposition 11, the discount function of the group defined in equation (39),

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

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 triggering boundary of this group, we need to know its weighting distribution $\bar{F}$. By Proposition 11, its cdf is $\bar{F}(r) \equiv \bar{F}(r; \alpha_1, \alpha_2) = \delta \left(1 - e^{-\frac{r}{\alpha_1}}\right) + (1 - \delta) \left(1 - e^{-\frac{r}{\alpha_2}}\right)$ and the corresponding density function is given by $\bar{f}(r) \equiv \bar{f}(r; \alpha_1, \alpha_2) = \delta \frac{1}{\alpha_1} e^{-\frac{r}{\alpha_1}} + (1 - \delta) \frac{1}{\alpha_2} e^{-\frac{r}{\alpha_2}}$. Therefore, by equations (24) and (25), the triggering threshold can be computed as

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

C On the well-posedness condition (20)

The following proposition shows that equation (20) indeed is the necessary and sufficient condition for our generalized real options investment problem to be well-defined.

**Proposition 12** If condition (20) holds, then $\forall x \in [0, \infty)$,

$$\sup_{\tau \in \mathcal{T}} E\left[\int_0^\infty e^{-\tau r} dF(r)(X_\tau - I)\right] < \infty.$$
If $b > \inf\{r \in [0, \infty) : F(r) > 0\}$, then $\forall x \in \mathbb{R}^+$,

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

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

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

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

$$\sup_{\tau \in T} \mathbb{E}\left[ \int_0^\infty e^{-r\tau} dF(r)(X_\tau - I) \right] \leq \int_0^\infty \sup_{\tau \in T} \mathbb{E}[e^{-r\tau}(X_\tau - I)] dF(r) \leq \sup_{\tau \in T} \mathbb{E}[e^{-r_0\tau}(X_\tau - I)] < \infty.$$  

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 $r \geq b$, $\mathbb{E}[e^{-r\tau}(X_\tau - I)] \geq 0$; for any $r < b$, $\mathbb{E}[e^{-r\tau}(X_\tau - I)] = \infty$. Since when $b > r_0$, $F(b- > 0$, then we have

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

This completes the proof.  

**References**


