Discounting, Values, and Decisions*

John K.-H. Quah
Oxford University

Bruno Strulovici
Northwestern University

March 24, 2013

Abstract

How do discount rates affect agents' decisions and valuations? This paper provides a common methodology and a systematic analysis of this question, considering stochastic and managed cash-flows, stochastic discount rates, time inconsistency, and including arbitrary learning and payoff or utility processes. We show that some of these features can lead to counter-intuitive answers (e.g., "a more patient agent stops *earlier*"), but we also establish, under simple conditions, theorems yielding robust and unifying answers on the impact of discount rates on control and stopping decisions and on valuations. We apply our theory to models of search, experimentation, bankruptcy, optimal growth, investment, and social learning.

*We are grateful for the suggestions and comments from the Editor and three anonymous referees, as well as comments from Chiaki Hara, Johannes Hörner, Paul Milgrom, Frank Riedel, Kareen Rozen, Larry Samuelson, Philipp Strack, Jean Tirole, and seminar participants at the World Congress of the Econometric Society (2010), Kyoto University, Kobe University, Konan University, Brown University, Yale University, and the SAET Conference (2011). Email addresses: john.quah@economics.ox.ac.uk and b-strulovici@northwestern.edu.

1 Introduction

How do discount rates affect agents' decisions and valuations? As economists, we are routinely confronted with this question: whether they concern savings and refinancing decisions, the exercise of options, or experimentation with reforms or new technologies, discount rates are at the heart of economic analysis. Given its importance, one might expect there to be a unified and general treatment of this question, but that is not the case. While there exist papers that deal with the general relationship between net present value, discount rates, and optimal stopping (in particular, Samuelson (1937), Arrow and Levhari (1969), and Sen (1975)), those works only consider the problem in a non-stochastic context.[1] When the environment is stochastic, answers have been given in different models using different methods, but there have been few attempts at an overarching and systematic analysis; the aim of this paper is to do precisely that.

With the multitude of phenomena that could arise in a general setting, it is not clear that robust comparative statics for discount rates are even possible. Indeed, it is known that, even in a deterministic context, a lower discount rate can *reduce* the present value of a project and precipitate its termination if it has a negative termination value (Sen, 1975) or if there are restrictions on the dates at which the project can be interrupted (Samuelson, 1937). Arrow and Levhari (1969) have shown that for deterministic pure stopping problems with no termination values and no restrictions on stopping times, both the stopping time and the net present value are increasing with patience. However, most economic applications do not fit in that framework. We consider more realistic environments where (i) cash/utility flows are stochastic and there are stochastic (possibly correlated) termination values; (ii) the agent can manage or influence in some way his cash flows (so his decision problem involves both stopping and control); (iii) discount rates are also stochastic and correlated with payoff or utility flows; and (iv) the agent is time inconsistent.[2] For reasons explained in Section 2.1, standard comparative statics techniques are ill suited to analyze the impact of discounting on decisions. In this paper, we develop a unifying methodology that achieves precisely that.

General results on discounting, stopping, and control

[1]These papers were concerned with finding conditions guaranteeing the monotonicity of the value of capital with respect to interest rates and the uniqueness of an *internal rate of return*. In addition to the applications provided in the present paper, our results also provide a general answer to these questions.

[2]Our analysis focuses on a single agent, but the methodology promises fruitful applications to the analysis of dynamic games (see the brief discussion in the Conclusion (Section 6)). For example, in a game where players choose effort and efforts are strategic complements, a change in the discount rate that leads to greater effort from one player will lead to higher equilibrium efforts for all players.

For pure stopping problems with deterministic discounting, we prove that even with stochastic cash/utility flows and stochastic (but non-negative) termination values, the agent with the lower discount rate stops later almost surely; similarly the net present value of the project also increases with a lower discount rate. Many economic decisions (such as search and experimentation, of which more below) can be modeled as pure stopping problems, so our general result will apply in all of those contexts. There are also problems where the agent has the ability to manipulate cash flows or utility flows; in formal terms these are problems with both stochastic control and stopping. For these problems, one could show that when the utility flow is a function of some Markov process, the *continuation domain* (i.e., the set of states at which stopping is suboptimal) is always shrinking with the discount rate. However, this does not necessarily mean that the more patient agent stops later and we give an example of a counter-intuitive situation where the opposite happens. To restore a positive result, more structure has to be imposed on the agent's decision environment. In a large class of models, utility depends on some one-dimensional state variable that follows a Markov process, such as the wealth of the agent, or the stock of capital available. We show that, as long as the utility of the agent is increasing in the state, a lower discount rate leads to (a) higher "effort" to increase the state, (b) a higher state trajectory (pointwise, almost surely), and (c) later project interruption.

In his presidential address to the American Finance Association, Cochrane (2011) argues that changes in discount rates matter much more than changes in cash flows in explaining asset return variation. Cochrane's address focuses on stochastic discount rates, which are pervasive in financial and macroeconomic analysis. In this context, we may consider an agent to be more patient than another if he has a lower discount rate almost surely. However, even with this strong definition of greater patience, one could build examples where the correlation between cash flows and discount rates leads to the more patient agent stopping sooner. To restore a positive result, additional conditions are needed; we show it suffices that (loosely speaking) lower discount rates are correlated with higher cash flows.

An oft-studied theme in behavioral economics is time inconsistent behavior, which can have a major impact on the timing of decisions (see, e.g., Akerlof (1991), O'Donoghue and Rabin (1999)). However, the issue of how time inconsistent agents may be ranked by their patience, and whether such a ranking is related to decisions in a way similar to that between the discount rate and decisions for time consistent agents, is a largely unexplored issue. We define an agent to be more patient than another if every one of his future selves is more patient than the corresponding future self of the other agent. We prove that for all stochastic pure stopping problems, the more patient agent stops earlier, provided the agent (and everyone of his future selves) uses a discount function obeying a certain condition satisfied by hyperbolic and also quasi-hyperbolic discount functions. This result is easily shown for naive agents; for sophisticated agents whose actions result from an equilibrium amongst the agent's multiple selves, the proof is more challenging. Using the same methods, we also show that for a given time inconsistent agent, increasing his sophistication (in the sense defined by O'Donoghue and Rabin (2001)) leads to earlier stopping almost surely.

Applications to specific models

Our results can be applied to yield general results on the impact of discounting in a number of important economic models.

(1) The standard economic model of search is an example of a pure stopping problem. For certain highly stylized search models, it is known that a more patient agent will search for a longer period. However, in many applications (e.g., to labor and housing markets), it is reasonable to suppose that the agent learns about conditions affecting search outcomes along the search process and he can also become better at searching, which reduces his search cost. In formal terms, this means that search costs may be stochastic and non-stationary over time, and the distribution of the termination value can also be state and time dependent. In some situations, objects encountered in the past can be recalled while in other situations, it may be reasonable to exclude this possibility; does this influence the impact of discount rates on optimal stopping? Our general result tells us that, with or without these complexities, the more patient agent stops searching later.

(2) It is standard in economics to model experimentation as a bandit problem (see Rothschild (1974)). In its simplest form, the agent has to decide when to switch from playing the 'risky' arm whose payoff distribution is unknown to playing the 'safe' arm, which has a known payoff distribution. This is a stochastic pure stopping problem and it follows immediately from our result that the more patient agent will switch to the safe arm, i.e., stop experimenting, later. In a multi-armed bandit problem with independent arms, this result can be extended to show that the duration of experimentation on each risky arm (and hence the total time spent on experimentation) also increases with patience.

(3) The recent literature on corporate default has emphasized the role of shareholders' optimizing behavior on the timing of bankruptcies. These models typically assume that shareholders's only decision is to decide on when to default; in cases where the setup is sufficiently simple, an explicit solution to this problem can be obtained (see, for example, Leland (1994)). More realistically, debt payments may depend in sophisticated ways on past performance and shareholders also make decisions on (say) project technologies that have

an impact on cash flow. Thus the shareholders' problem involves both optimal control and stopping. Whether or not a model with these features can be solved explicitly, our results guarantee that a lower discount rate leads to later bankruptcy.

(4) A basic question in growth theory is whether a more patient representative agent leads to 'capital deepening', i.e., a higher capital stock at all times (see, for example, Amir et al. (1991)). Models giving this result are typically deterministic, involve strong assumptions on the form of the utility and production functions, and restrict the agent's decision to that of deciding on the level of consumption/saving at each period. We show, in a continuous time economy with a stochastic production technology, that a more patient representative agent leads to capital deepening almost surely; this result does not require shape restrictions on the utility and production functions (thus permitting increasing returns to scale) and also allows the agent's control variable to be multi-dimensional (thus allowing for other decision variables, besides consumption, to affect capital accumulation).

(5) In consumption-based asset pricing models, the cash flow of a project is discounted by a stochastic discount function that is proportional to the marginal utility of the economy's representative agent at each time and state. We establish a systematic relation between an agent's risk attitude and the "patience" implied by his discount function: the less risk averse agent, the more patient is the discount function, which in turn means that projects will have higher net present value and terminate later.

Organization of the paper

Section 2 is devoted to the pure stopping problem, beginning with a discussion of the deterministic case to provide motivation. Problems involving both stopping and control are dealt with in Section 3, while Section 4 considers pure stopping problems with stochastic discounting (in addition to stochastic payoffs). Most of our results do not depend on discrete time versus continuous time specifications. For simplicity, counterexamples are mostly given in discrete time, while general results are mostly stated in continuous time. The exception is Section 3.2 and applications of its main result; the continuous time formulation seems crucial in that case. Time inconsistency is discussed in Section 5; to keep things technically manageable, we use a discrete time formulation in that section. All the substantive sections begin with a formulation of the general theory followed by applications to widespread economic models. Some of the simpler proofs are contained in the main part of the paper; all others are found in the Appendix. Section 6 concludes.

2 The Stochastic Pure Stopping Problem

In this section we consider how the value and optimal stopping time of a pure stopping problem varies with the discount rate. This issue has been fully solved in the case where payoffs are deterministic but not when it is stochastic. Nonetheless, we shall begin with a brief discussion of the deterministic case because this simpler setting allows us both to highlight some of the difficulties involved in establishing a general result and also to develop some intuition for why a general result may be possible.

2.1 Background: the deterministic case

Suppose there is a project that gives a cash flow – which can be either positive or negative – of $\pi(s)$ at time s over some interval $[0, \bar{t}]$. The project can be stopped at no cost at some time before $\bar{t}$. An agent with a discount rate of r chooses the stopping time t to maximize the discounted cash flow $U(t,r) = \int_0^t e^{-rs}\pi(s)ds$. The value of the project, as a function of r is $V(r) = U(\tau(r), r)$, where $\tau(r)$ is the optimal stopping time at the discount rate r. The question we wish to ask is the following: how do τ and $V(r)$ vary with r?

The first answer to this problem was provided by Arrow and Levhari (1969), who showed that *τ and V are both decreasing in r*. The proof given by these authors is rather complicated and involves ancillary assumptions, even though the reader may at first blush wonder why this problem is not completely straightforward. It is not straightforward because common methods for dealing with comparative statics problems are not applicable in this case. To see this note that $U(t,r)$ is not typically concave in t; indeed, since $\frac{\partial U}{\partial t}(t,r) = \pi(t)$, U has a local maximum at t^* whenever $\pi(t^*) = 0$ and π is locally decreasing. What is worse, the local maxima do not vary with r, so there is no straightforward way of determining how $\tau(r)$, which is the time at which U is *globally* maximized, varies with r.

A more sophisticated approach may be to check for single crossing differences (see Milgrom and Shannon (1994)). Using this approach, a sufficient condition for $\tau(r)$ to be decreasing in r is the following: for all $t_2 > t_1$ and $r' > r$,

$$U(t_1, r) \geq (>)\, U(t_2, r) \Longrightarrow U(t_1, r') \geq (>)\, U(t_2, r') \tag{1}$$

We now provide an example which illustrates why single crossing differences can fail and also hints at a way out of the problem. Consider a project, active over the time interval $[0, 3]$, which has the following payoff: $\pi(t) = A > 0$ for $t \in [0, 1]$, $\pi(t) = -B < 0$ for $t \in (1, 2]$ and $\pi(t) = C > 0$ for $t \in (2, 3]$. The local maxima of $U(t, r)$ are at $t = 1$ and 3, and depending on r, either could be the global maximum. Now it is straightforward to check that

$$U(0,r) - U(2,r) = r^{-1}(1-e^{-r})\left(Be^{-r} - A\right).$$

which, when B is large, goes from positive to negative values as r increases, so there is a failure of single crossing differences. Put another way, *if* the agent can only stop at $t = 0$ or 2, then less patience can lead to later stopping, which is the opposite of what we want. But in fact, the agent can stop at any time between 0 and 3, so it does not matter that $t = 2$ becomes preferred to $t = 0$ as r increases because $t = 0$ is never optimal. On the other hand,

$$U(1,r) - U(3,r) = r^{-1}e^{-r}(1-e^{-r})\left(-Ce^{-r} + B\right),$$

which *does* go from negative to positive as r increases (and thus obeys (1)). This means that if stopping at $t = 1$ is preferred to stopping at $t = 3$ for some value of r, it will remain preferred as r increases: the impatient stops earlier.

This example suggests that what is needed to establish the monotonicity of $\tau(r)$ is something weaker than single crossing differences — a concept that restricts payoff comparisons to those pairs of choice variables that are relevant for global optimality. In Quah and Strulovici (2009), we develop a way of ranking payoff functions (called the *interval dominance order*) which has this feature and we use it to establish the monotonicity of $\tau(r)$. The techniques in that paper can be developed (in a non-trivial way) to obtain a general result relating patience and stochastic optimal stopping, even though the stochastic nature of the problem precludes a straightforward application of the interval dominance order.

2.2 Stochastic optimal stopping

For $T = [0,\bar{t}]$ or $T = [0,\infty)$, let $\{u_t\}_{t\in T}$ and $\{G_t\}_{t\in T}$ denote real-valued stochastic processes adapted to the filtration $\mathcal{F}$ of some filtered probability space $(\Omega, P, \mathcal{F} = \{\mathcal{F}_t\}_{t\in T})$. The value function of a pure stopping problem is given by

$$V(\alpha) = \sup_{\tau\in\mathcal{T}} U(\tau;\alpha) = E\left[\int_0^\tau \alpha(s)u_s ds + \alpha(\tau)G_\tau\right] \tag{2}$$

where $\mathcal{T}$ denotes the set of stopping times adapted to $\mathcal{F}$ and taking values in T. We assume that the discount function $\alpha : T \to (0,\infty)$ is deterministic (for now) and strictly positive, with the normalization $\alpha(0) = 1$. We also assume that $E[\int_0^{\bar{t}} \alpha(s)|u_s|ds]$ is finite, $\alpha(t)$ and $|G_t(\omega)|$ are uniformly bounded, α and G have bounded variation, that all processes are right-continuous and have left limits, and that the filtration $\mathcal{F}$ is right continuous and satisfies the usual conditions (see, e.g., Karatzas and Shreve (1991)). These assumptions guarantee that all expectations (including conditional expectations) are (almost everywhere) well-defined, that Fubini's Theorem can be used, and that Lebesgue-Stieltjes integrals with respect to α and G are well defined.

The discount function α gives, at each time t, the discount factor between time 0 and time t. *By our definition*, the stopping time τ that solves (2) is optimal from the perspective of an agent at $t = 0$. The agent is time-consistent, i.e., will not revise at time $s > 0$ the optimal stopping time chosen initially, if his discount function at time s is given by $\bar{\alpha}(s,t) = \alpha(t)/\alpha(s)$ where t (with $t > s$) refers to the time that has elapsed from time 0. It would be reasonable to restrict discount functions to be decreasing in t, though our results do not in fact require that assumption.

For expositional simplicity, we assume that the agent stops if and only if he is indifferent between stopping and continuing, and denote by $\tau(\alpha)$ the resulting stopping time. All of our results can be stated more generally by endowing the set of stopping times with a lattice structure and making statements in terms of the strong set order. Formally, $\tau(\alpha)$ is the essential infimum of all stopping times that are optimal when the discount function is α. The right continuity of G and of the filtration $\{\mathcal{F}_t\}_{t\in T}$ guarantee that $\tau(\alpha)$ is an optimal stopping time.[3]

The stopping problem as stated assumes that the flow payoff after stopping is zero. When the agent is making a purely financial decision (such as a firm deciding on when to invest in a project, or a landlord deciding on when to sell an apartment, etc.), then the zero payoff simply means that the cash flow ceases when the project stops. The agent's objective in (2) is to maximize the discounted cash flow and $V(\alpha)$ can then be interpreted as the present value of the project. In other settings, the post-stopping zero flow payoff is a normalization and what happens after stopping depends on the setting. For example, imagine an agent who is engaging in some activity X that gives him a stochastic consumption of x_t at time t. He has the option of switching, without cost but only once, away from this activity to another activity Y that gives a stochastic consumption of y_t at time t.[4] The agent has a discount function α and is risk averse, with the Bernoulli utility function w. Thus the agent's problem is to choose a stopping (or, in this context, a switching) time τ that is adapted to

[3]The essential infimum of an arbitrary family of random variables is uniquely defined up to a set of measure zero (see Durrett (2004)). The right continuity of G and of the filtration $\{\mathcal{F}_t\}_{t\in T}$ guarantee that the infimum is itself optimal (see Lamberton and Lapeyre (2007)).

[4]Note that y_t does not depend on the time at which the agent switches from X to Y.

$\mathcal{F}$ and maximizes his lifetime utility

$$E\left[\int_0^\tau \alpha(s)w(x_s)ds + \int_\tau^{\bar{t}} \alpha(s)w(y_s)ds\right]. \tag{3}$$

Clearly, we can reformulate this problem by treating the agent as choosing τ to maximize

$$E\left[\int_0^\tau \alpha(s)\left(w(x_s) - w(y_s)\right)ds\right]; \tag{4}$$

this problem then takes precisely the form (2), with $u_s = w(x_s) - w(y_s)$, $G_t \equiv 0$, and a payoff of zero after stopping. Note also that the optimal stopping time is invariant to affine transformations of w and, consequently, to linear – but not affine – transformations of u_s.

2.3 Ordering discount functions

We endow discount functions with the following relation, which (as one could easily check) is a partial order.

Definition 1 (Patience Order) *The discount function β exhibits* greater patience *than another discount function α (denoted by $\alpha \prec \beta$) if*

$$\frac{\beta(s)}{\alpha(s)} \text{ is weakly increasing in } s.$$

If, for instance, $\alpha(s) = \exp(-rs)$ and $\beta(s) = \exp(-\bar{r}s)$, then $\alpha \prec \beta$ whenever $\bar{r} \leq r$. More generally, if $\alpha(t) = \exp\left(\int_0^t -r_\alpha(s)ds\right)$ for some discount rate function r_α, with a similar expression for β, then $\alpha \prec \beta$ whenever $r_\beta(t) \leq r_\alpha(t)$ for all t. We shall refer to an agent with discount function α as simply the α agent (or agent α) and we say that agent β is *more patient* than agent α if $\alpha \prec \beta$. The next result motivates our definition of the patience order.

Proposition 1 [5] *Two agents with discount functions α and β are choosing between receiving π_1 at time t_1 and π_2 at time t_2, where both payoffs and times can vary freely, subject to $\pi_1 > 0$, $\pi_2 > 0$, and $t_2 > t_1$. Then $\alpha \prec \beta$ if and only if, whenever agent α chooses π_2 then agent β also chooses π_2.*

Proof. If $\alpha \prec \beta$ and $\alpha(t_2)\pi_2 > \alpha(t_1)\pi_1$, then clearly $\beta(t_2)\pi_2 > \beta(t_1)\pi_1$. Thus agent β is more patient in the sense that whenever agent α is prepared to wait longer to receive π_2 then so is agent β. Conversely, suppose that $\alpha \not\prec \beta$, so there is $t_2 > t_1$ such that

$$\frac{\beta(t_2)}{\beta(t_1)} < \frac{\alpha(t_2)}{\alpha(t_1)}.$$

[5] We are very grateful to an anonymous referee for suggesting this characterization.

If we can choose π_1 to be a number strictly between the two ratios and let $\pi_2 = 1$, then the agent with discount function β will strictly prefer to receive π_1 at t_1 to π_2 at t_2 (because $\beta(t_1)\pi_1 > \beta(t_2)$) while the agent with discount function α will have the opposite preference between those alternatives (because $\alpha(t_2) > \pi_1\alpha(t_1)$). ■

Notice that this choice problem can also be understood as a problem of optimal stopping, where the flow payoffs are identically zero, and there are termination payoffs of π_1 at t_1, π_2 at time t_2, and nothing at other times. Therefore, the proposition tells us that, for $\tau(\alpha) \le \tau(\beta)$ to hold generally, it is *necessary* that $\alpha \prec \beta$. The next result tells us that it is also sufficient.

2.4 Patience and Stopping

THEOREM 1 *Suppose that G is a nonnegative process in the stopping problem (2). Then $V(\beta) \ge V(\alpha)$ and $\tau(\beta) \ge \tau(\alpha)$ a.s. whenever $\alpha \prec \beta$.*

A key condition for the result is the agent's ability to stop *at any time* in the interval $[0, \bar{t}]$. In order to get some intuition for this, suppose on the contrary that the agent cannot stop, and is thus condemned to incur all cash flows till $\bar{t}$. In this case, there is a perfect symmetry in the agent's value function corresponding to a cash flow stream and the opposite of that stream. In particular, $U(\bar{t};\beta) > U(\bar{t};\alpha)$ for some stream π if and only if $U(\bar{t};\beta) < U(\bar{t};\alpha)$ for the stream $-\pi$. Thus, if the agent is limited to either stopping immediately or incurring the entire stream of cash flows, it is easy to build examples for which only the more patient (agent β) stops immediately.

What breaks this symmetry is the endogenous interruption of cash flows at any time in $[0, \bar{t}]$: if a cash flow sequence is first positive and then negative, both a patient and impatient agent will stop after getting the positive cash-flows, and the patient agent will experience a higher discounted utility from it. By contrast, if the cash flow stream is first negative and then positive, the present value may be negative for the impatient agent and positive for the patient agent, but the reverse can never hold. A patient agent may therefore find it optimal to endure the negative cash flows in order to enjoy the positive ones in the future when an impatient decides to stop immediately, but the reverse situation can never occur.

Theorem 1 generalizes to a stochastic environment the results of Arrow and Levhari (1969), Sen (1975), and Quah and Strulovici (2009). The result, like all results of this section, has a straightforward counterpart in discrete time. The proof of this theorem proceeds by showing that, at agent α's optimal stopping time $\tau(\alpha)$ (but not at other stopping times),

$$U(\tau(\alpha);\beta) \ge U(\tau(\alpha);\alpha) = V(\alpha). \tag{5}$$

Since $V(\beta) \geq U(\tau(\alpha); \beta)$, we obtain $V(\beta) \geq V(\alpha)$. Given this result, it is intuitive that $\tau(\beta) \geq \tau(\alpha)$: at a given time $s > 0$ and state, agent β's value of the future payoff stream is higher than agent α's; thus if the latter does not stop neither will the former.

The nonnegativity of G is needed in Theorem 1 for the following reason: if the terminal value gets increasingly negative (e.g., a chore that worsens over time), a more impatient agent may postpone the chore until the last minute, while a more patient agent, more sensitive to the worsening of the chore, may decide to stop immediately. Recalling an example emphasized by Sen (1975), a more patient agent may decide to stop polluting earlier.

Lastly, we consider three agents with discount functions α, β, and γ, with $\alpha \prec \beta \prec \gamma$. The next result, which is a corollary of Theorem 1 says that agent γ prefers the stopping rule of agent β agent to that of agent α. In particular, if both α and β exhibit impatience, in the sense that they are decreasing in t, then a perfectly patient agent, i.e., an agent with discount function $\gamma(t) \equiv 1$ for all $t \in [0, \bar{t}]$, will prefer the stopping decision of the less impatient agent β to that of agent α.

PROPOSITION 2 [6] *Suppose that G is a nonnegative process in the stopping problem (2) and $\alpha \prec \beta \prec \gamma$. Then $U(\tau(\beta), \gamma) \geq U(\tau(\alpha), \gamma)$, where $\tau(\alpha)$ and $\tau(\beta)$ are the optimal stopping times of agents α and β respectively.*

2.5 Application: Costly Search

Search models play a fundamental role in economics and have numerous applications to labor and housing markets, macroeconomics and monetary economics (see, for example, Rogerson et al. (2005)). In these models, it is often assumed that that marginal cost of search is constant and the item being sought has qualities that are drawn from a distribution that is fixed over time (either because it is exogenous, or because it is an equilibrium distribution that is common knowledge). In practice, however, agents learn about this distribution as their search progresses. For example, a worker learns about labor market conditions or about the idiosyncratic marketability of his skills. In housing markets, a household's search strategy may evolve over time, by narrowing or directing its focus on specific housing locations or categories. Similarly, search cost can evolve over time: it can increase as a result of scarcer opportunities (e.g., a job candidate has to look farther for a job), or decrease as a result of experience or sunk costs (the household hears of more efficient or cheaper real estate agents along the search processes). Indeed, there are realistic scenarios where search costs and search outcomes are stochastic and correlated.

[6] We would like to thank an anonymous referee for suggesting this result.

Standard search problems are instances of a pure stopping problem where the utility flow is the search cost $c_s \geq 0$ and the termination value G_τ is the value of the item eventually chosen by the agent. Formally, the agent chooses a stochastic stopping time τ to maximize

$$E\left[-\int_0^\tau \alpha(s)c_s\,ds + \alpha(\tau)G_\tau\right] ;$$

if there is no recall, $G_t = M_t$, where $M_t \geq 0$ is the value of the item uncovered at time t and, if there is perfect recall, $G_t = \max_{s \leq t}\{M_s\}$. Theorem 1 tells us that even when the search cost is time-varying and stochastic, and even when the value *distribution* of the item being sought changes over time, *the duration of search increases with patience.*

2.6 Application: Experimentation

The most widely used model of endogenous learning is probably the experimentation model based on the multi-armed bandit framework.[7] In this section we consider how the length of experimentation in a multi-armed bandit model depends on the discount rate of the agent. The problem is formalized as follows. Let $x_t = (x_t^1, x_t^2, \ldots, x_t^n)$ denote the state at instant t, where n is the number of arms. A strategy for the agent is a stochastic process $\iota : [0, +\infty) \to \{1, \ldots, n\}$ that is right continuous and has left limits. Denoting by $\mathcal{I}$ the set of all such processes, the value function of the agent is given by

$$V(x_0) = \sup_{\iota \in \mathcal{I}} E\left[\int_0^\infty e^{-rt}\pi(x_t^{\iota_t}, \iota_t)dt\right],$$

where the instantaneous payoff π depends on the arm chosen and on the state of that arm.[8] When applied to learning models, the state and payoffs have the following structure: x_t^i corresponds to the agent's belief, at time t, about the *underlying state*, θ_i, of arm i. In this case, we have $\pi(x_t^i, i) = E[\Pi(\theta_i)|x_t^i]$, where $\Pi(\theta_i)$ is the expected payoff of arm i if its underlying state is θ_i, and x_t^i evolves according to Bayesian updating. The underlying states $\{\theta_i\}_{i\in\{1,\ldots,n\}}$ are assumed to be independently distributed, so that the arms are *uncorrelated*: pulling one arm does not tell the agent anything about the underlying state of the other arms.[9] Lastly, we assume that arm 1 is a *safe* arm in that it has a known payoff distribution,

[7] One of its earliest applications to economics dates back to Rothschild (1974), who uses a one-armed bandit to model market learning and pricing by a monopolist. See Bergemann and Välimäki (2006) for a recent review of the multi-armed bandit theory and its economic applications.

[8] This formalization excludes lump sum payoffs. At the expense of introducing more notation, the model could easily be modified to include jumps in payoffs and beliefs, as in Keller et al. (2005).

[9] In the case of learning with Brownian noise, $dx_t^{\iota_t} = \kappa^{\iota_t} x_t^{\iota_t}(1 - x_t^{\iota_t})dB_t^{\iota_t}$ and $dx_t^i = 0$ for $i \neq \iota_t$, where $\{B_t^i\}_{i,t}$ is a vector of standard Brownian motions and $\kappa = (\kappa^1, \ldots, \kappa^n)$ is a vector of constants capturing the signal-to-noise ratio of each arm (see Bolton and Harris (1999) for the derivation).

and therefore entails no learning ($x_t^1 = x_0^1$ for all $t \geq 0$); all other arms have unknown payoff distributions and will be called *risky*.

It is easy to show that if the agent chooses the safe arm at some instant, then he will play that arm forever after, because the problem is Markovian and the agent learns nothing new about the arms, implying that the state becomes constant. Therefore, the experimentation problem may be re-expressed as follows:

$$V(x_0) = \sup_{\iota \in \mathcal{I}_{-1}, \tau \in \mathcal{T}} E\left[\int_0^\tau e^{-rt}\pi(x_t^{\iota_t}, \iota_t)dt + \int_\tau^\infty e^{-rt}\pi(x_0^1, 1)dt\right],$$

where $\mathcal{I}_{-1}$ is the subset of processes in $\mathcal{I}$ taking values in $\{2, \ldots, n\}$, and $\mathcal{T}$ is the set of all stopping times. Without loss of generality, the payoff of the safe arm can be normalized to zero, because the agent's optimization problem is unchanged if all cash-flows are translated upwards or downwards by a constant amount. This leads to the following further simplification of the problem:

$$\bar{V}(x_0) = \sup_{\iota \in \mathcal{I}_{-1}, \tau} E\left[\int_0^\tau e^{-rt}\bar{\pi}(x_t^{\iota_t}, \iota_t)dt\right], \tag{6}$$

where $\bar{\pi}(x^i, i) = \pi(x^i, i) - \pi(x_0^1, 1)$ for all x^i and $i \geq 2$. Stopping experimentation now means that the agent receives a payoff stream of zero forever after stopping.

Whittle (1980)[10] shows the following key result: the optimal stopping time, τ, of problem (6) is equal to the sum of the stopping times $\{\tau_i\}_{i=2}^n$, where τ_i is the solution to (6) in the case where only arm i and (the safe) arm 1 are available. This result is intuitive: how long the agent tries a given arm should be independent of other risky arms, because the states of individual arms are uncorrelated. Therefore, when the agent retires, he must have tried each risky arm exactly as long as what he would have done if only that arm had been available (in addition to the safe arm). Now, each bandit problem with one risky and one safe arm is a pure stopping problem, and Theorem 1 says that τ_i is increasing in patience. This is true of every i, and thus of $\tau = \sum_i \tau_i$. We have shown the following result.

PROPOSITION 3 *In a multi-armed bandit with uncorrelated risky arms and a safe arm, a more patient agent will, almost surely, experiment longer on each risky arm (and thus stop experimentation later).*

When there is a single risky arm whose distribution is governed by a binary ("High" or "Low") unknown parameter , Cohen and Solan (2012) have characterized the optimal cutoff

[10]See also Frostig and Weiss (1999, p. 13) for a concise exposition.

policy, from which it seems easy to derive comparative statics on the discount rate.[11] Proposition 3 provides a general result where the agent can experiment with multiple risky arms and the parameter governing each arm can take arbitrarily many values.

This application illustrates that a more patient can stop later even when he *controls* his learning process (here, by choosing which arm to pull at each instant). How general is this result? In the next section we shall examine the impact of discounting and patience in models with both control and stopping.[12]

3 A General Model of Stochastic, Managed Cash Flows

We now extend the setting of the last section to combine both stopping and control decisions. The optimization problem is now the following:

$$V(\alpha) = \sup_{\tau,\lambda} U(\tau, \lambda; \alpha) = E\left[\int_0^\tau \alpha(s)u(x^\lambda(s), \lambda(s), s)ds + \alpha(\tau)G_\tau\right], \tag{7}$$

where $\{x_t^\lambda\}$, $\{\lambda_t\}$, and τ are adapted to $\mathcal{F}$ and $\{x_t^\lambda\}_{t\geq 0}$ is controlled by λ. $\{x_t^\lambda\}_{t\in T}$ takes values in some topological space $\mathcal{X}$ and $\{\lambda_t\}_{t\in T}$ takes value in some set Λ. $\{\lambda_t\}_{t\in T}$ is assumed to be *admissible*, i.e., such that x^λ is uniquely defined and such that $E\left[\int_0^{\bar{t}} |\alpha(s)u(x(s), \lambda(s), s)|ds\right]$ is finite.[13] It is easy to see, given our earlier results, that the value of this combined control and stopping problem remains monotonic in the discount rate.

Proposition 4 *Suppose that G is a nonnegative process and that α and β are discount functions for problem (7) with $\alpha \prec \beta$. Then $V(\beta) \geq V(\alpha)$.*

Proof. Suppose $(\tau(\alpha), \lambda(\alpha))$ is jointly optimal for agent α. Then $\tau(\alpha)$ is also optimal for this agent, conditional on $\lambda(\alpha)$. Therefore, by (5), we obtain $U(\tau(\alpha), \lambda(\alpha); \beta) \geq U(\tau(\alpha), \lambda(\alpha); \alpha)$. Since $V(\beta) \geq U(\tau(\alpha), \lambda(\alpha); \beta)$, we have $V(\beta) \geq V(\alpha)$. ∎

It would be tempting to conclude that the optimal stopping time is also increasing with patience. However, that is generally false, because optimal controls may differ under α and β,

[11] Their result generalizes the characterization of optimal policies (for the one-agent case) in the Brownian setting of Bolton and Harris (1999) and the Poisson setting of Keller, Rady, and Cripps (2005).

[12] In the standard multi-armed bandit model, the agent chooses only one arm at each instant. If instead he can choose the speed of experimentation (for example, by allocating resources between risky and safe arms), the actual *time* spent experimenting is not necessarily increasing with patience. As illustrated in the more abstract setting of the counterexample in Section 3.2, a more patient decision maker may experiment at a faster rate than a less patient one, resulting in the experimentation boundary being reached faster.

[13] See Fleming and Soner (1993) for a general definition of admissible controls.

resulting in different paths for the state variables. An example of this phenomenon is provided in Section 3.2. When the state has Markov dynamics, however, the next section proves in full generality that the *continuation domain*, i.e., the set of states for which stopping is suboptimal, increases with patience. Section 3.2 will then provide an additional monotonicity condition under which the optimal stopping time increases in patience.

3.1 Continuation Domain

Suppose now that $\{x_t\}$ is a Markov process: for any t' and control $\{\lambda_s\}_{s\geq t'}$, the distribution of $\{x_s\}_{s\geq t'}$ depends on past history only through $x_{t'}$. Also suppose that the termination value process is of the form $G_t = G(x_t, t)$ for some function G of state and time. In this case, it is well known that the value function depends, at any time t, only on the current state x_t and time t: $V_t = V(t, x_t)$ for some function V.

For any discount function α, define the *continuation domain* at time t by

$$C(\alpha, t) = \{x : V(\alpha, t, x) > G(t, x)\}, \quad \text{where}$$
$$V(\alpha, t, x) = \sup_{\tau\geq t,\lambda} E\left[\int_t^\tau \frac{\alpha(s)}{\alpha(t)} u(x^\lambda(s), \lambda(s), s)ds + \frac{\alpha(\tau)}{\alpha(t)} G_\tau\right].$$

$C(\alpha, t)$ is the set of states at which stopping at time t (and, thus, getting termination value $G(t, x)$) is strictly suboptimal. In general, $C(\alpha, t)$ varies with time, but if x has time-homogeneous dynamics and $G(x, t)$ only depends on x, $C(\alpha, t)$ is a constant subset of $\mathcal{X}$. For example, if x is a one-dimensional controlled diffusion, then $C(\alpha, t)$ is typically (though not necessarily) an interval (a, b) such that it is optimal to stop exactly when x_t hits a boundary of the interval.[14] The next result says that *the continuation domain expands with patience.*

Theorem 2 *Suppose that, for problem (7), x is a Markov process and $G_t = G(x_t, t)$ is nonnegative. Then $C(\alpha, t) \subset C(\beta, t)$ for all t, whenever $\alpha \prec \beta$.*

Proof. Suppose that $x \in C(\alpha, t)$. This means that $V(\alpha, t, x) > G(t, x)$. Theorem 4, applied at time t, implies that $V(\beta, t, x) \geq V(\alpha, t, x)$, so $x \in C(\beta, t)$. ■

3.2 Patience, Control and Stopping

Theorem 2 does *not* imply that a more patient agent stops later. Indeed, agents with different patience levels apply different controls to the state, and a more patient agent may accelerate

[14]In Sections 3.2 and 3.3, the interval is of the form (a, ∞).

the collision of the state with the boundary of the continuation domain, even though that domain always expands with patience. This is illustrated in the following example.

A counterexample

Consider a Markov setting with initial state $x_0 = 0$ and with two control levels: $\Lambda = \{1, 2\}$. The utility flow is equal to $u(x, \lambda) = M$ for $x \in [1, 10]$, $u(x, \lambda) = -M$ for all $x > 10$ where M is a large positive constant, and $u(x, 1) = 1$ and $u(x, 2) = -0.01$ for $x \in [0, 1)$. Finally, suppose that

$$\frac{dx}{dt} = \lambda_t.$$

The state x_t can only go up. Moreover, it is optimal to stop at $x = 10$, and not before, since there is always a control yielding positive utility before that level. Thus, the continuation domain is $C = [0, 10]$ for all discount functions. Finally, it is optimal to spend as much time as possible in the region with payout rate M, i.e. set $\lambda(x) = 1$ for $x \in [1, 10]$. The only question, therefore, is how fast to get to $x = 1$. A very impatient agent will never use the control $\lambda(x) = 2$ for x sufficiently close to 0, because that control yields negative instantaneous utility, while the reward (M) is "far away," arriving only when $x_t = 1$. By contrast, a patient agent puts more value on future cash flows; from his perspective, the small negative cash flow 0.01 is only incurred for a short time, and leads to high cash flows sooner. Thus it seems possible that for sufficiently high values of M, the more patient agent will choose $\lambda(x) = 2$ for *all* $x \in [0, 1)$ and therefore stop sooner than a less patient agent. This intuition is proved formally in the Appendix.

Positive Results: Markovian State with Increasing Utility

To obtain a positive result relating stopping and patience, we consider the following control problem which is a special case of (7):

$$\text{maximize}\, U(\tau, \lambda, ; \alpha) = E\left[\int_0^\tau e^{-rs} u(x(s), \lambda(s), s) ds\right], \tag{8}$$

subject to

$$\text{subject to } dx_t = \mu(x_t, \lambda_t, t)dt + \sigma(x_t, t)dB_t \qquad x_0 = x, \tag{9}$$

where $\{x_t\}$ is one dimensional and B is the standard Brownian motion. We assume that σ is bounded below and away from zero and that the value function is well-defined and smooth, so that it obeys its Hamilton-Jacobi-Bellman equation.[15] As stated, the problem assumes

[15]For time-homogeneous dynamics, Strulovici and Szydlowski (2012) prove that the value function is smooth as long as the control space is compact and the drift, volatility, and return functions satisfy standard Lipschitz and growth conditions.

constant discounting; this is to simplify the exposition and the case of general discount functions is treated in the Appendix. Notice also that the termination value is set at zero; it is not hard to see that allowing it to take arbitrary positive values can lead to the more patient agent stopping earlier, which is the opposite of the result we want.[16] Last but not least, we assume that the instantaneous utility function, $u(x, \lambda, t)$ *is increasing in* x *for all* (λ, t). Since it is always possible to think of the state as $-x$ rather than x, any results obtained with this assumption will have an analog in the case where $u(x, \lambda, t)$ is decreasing in x for all (λ, t). In effect, we are excluding cases where u is not monotone in x or where it increases in x for some values (λ, t) and decreases in x for others. This assumption is satisfied in many applications, though, obviously, not in the counterexample we gave.

The Markov setting of problem (8) means that the continuation value at any time depends on the history only through the current state and so could be written as $V(x, t, r)$ and the optimal control λ can similarly be thought of as a function of just (x, t, r). Because we assume u is increasing in x, so is V. Indeed, starting from some state y, with $y > x$, one can always replicate the control applied when starting from x and get a higher payoff flow at all times before and upon stopping. The continuation domain at time t consists of the states x such that $V(x, t, r) > 0$; therefore, there exists a process $a(t)$ such that the optimal stopping time is given by $\tau = \inf\{x_t \notin (a_t, \infty)\}$.

Given the initial condition, let $X(r) \in \mathcal{C}\left([0, \bar{t}], \mathbb{R}\right)$ denote[17] the path obtained when using the optimal control at each time t. In other words, for each realization ω of the underlying Brownian motion, $X(r)$ is the trajectory of the state over time. We say that $X(r)$ is decreasing in r if, for all $r'' > r'$, we have $Pr\left(\{\omega : X(r'')_t \leq X(r')_t \text{ for all } t \in [0, \bar{t}]\}\right) = 1$.

Theorem 3 *Let* $\lambda(x, t, r)$ *and* $\tau(r)$ *be the control and stopping time that solve problem (8).*[18] *Then* $\mu(x, \lambda(x, t, r), t)$, $X(r)$, *and* $\tau(r)$ *are all decreasing in* r, *provided* $u(x, \lambda, t)$ *is increasing in* x *for all* (λ, t).

Theorem 3 says i) the optimal drift, ii) the state path, and iii) the optimal stopping time are all increasing in patience. There is an intuitive explanation for this result. The key property is that when u is increasing in x, V is submodular in (x, r); this means that, starting from a

[16] For example, if the termination value is very high at some time $\bar{T}$ if the state $x \geq \bar{x}$ and zero otherwise, and if it is costly to raise the value of x over time, a patient agent may incur the cost of raising x to $\bar{x}$ in order to get to the high termination value and stop at $\bar{T}$, while a very impatient agent, focused only on the high cost of raising x, may reject this option and stop later (or never), provided the utility flow is nonnegative.

[17] $\mathcal{C}\left([0, \bar{t}], \mathbb{R}\right)$ denotes the set of real-valued continuous functions defined on $[0, \bar{t}]$ (or $[0, \infty)$, if $\bar{t} = +\infty$).

[18] In the case of multiple solutions we refer here to the solution leading to the smallest drift and the earliest stopping time; see the Appendix for details.

common state, the agent with the lower discount rate (i.e., the more patient agent), values an increase in x more than the less patient agent. Consequently, whenever they are in the same state, the more patient agent will (using his control) drive the state higher than the less patient. Thus the more patient agent has the higher drift and his state path (which is a continuous function of time) may touch but will never cross the corresponding path of the less patient agent. Recall that the stopping boundary takes the form of a *lower* bound on the state, and by Theorem 2, that bound has to be lower for the more patient agent. These observations guarantee that the state trajectory of the more patient agent hits his stopping boundary later than the trajectory of a less patient agent.

Notice that while Theorem 3 says the optimal drift is decreasing in r, it does not say that the same holds for the optimal control. Of course, when the control λ is scalar and the drift function μ is increasing in λ, then the optimal control must be decreasing in r since $\mu(x, \lambda, t) = m$. The next result gives conditions under which a multi-dimensional optimal control is decreasing in r.

PROPOSITION 5 *Suppose that Λ (the domain of control λ) is a lattice, μ is increasing in λ, and u and μ are supermodular in λ. Then $\lambda(x, t, r)$ is decreasing in r.*

3.3 Application: Bankruptcy

Recent models of corporate bankruptcy, starting with Leland (1994), take the view that the timing of bankruptcy is endogenously chosen by equityholders. Until then, equityholders firm must pay a coupon rate $c(x)$ to debtholders, where c is decreasing in some performance measure x of the firm, and receive a payout rate $\delta(x)$, with δ increasing in x.[19] Following Leland (1994) and Manso et al. (2009), we assume that the performance measure $\{x_t\}$ is a time-homogeneous diffusion (for example, a geometric Brownian motion, or a mean-reverting process). The shareholder problem is thus the pure stopping problem of solving

$$V(x, r) = \sup_{\tau \in \mathcal{T}} E_x \left[\int_0^\tau e^{-rt} (\delta(x_t) - c(x_t)) dt \right]. \tag{10}$$

Given the time-homogeneous Markov structure of the problem, and given that $\delta - c$ is an increasing function of x, it is easy to show that optimal default takes the form of a hitting

[19]For standard debt, c is a constant. However, in many contracts such as performance-pricing loans or step-up bonds, c increases as some performance measure of the firm deteriorates. This measure maybe the credit rating, or directly related to the earnings (EBITDA, price-earning ratio, etc.) of the issuing firm (see Manso, et al. (2009) for examples).

time $\tau(r) = \inf\{t : x_t \le A_B(r)\}$; $A_B(r)$ is called the *default-triggering level* of the firm, and is independent of the initial asset level x. If $\delta(x) = \delta x$, $c(x) = c$, and x is the geometric Brownian motion with drift μ and volatility σ, standard computations (see Leland (1994)) imply that

$$A_B(r) = \frac{\gamma(r)}{\gamma(r)+1}\left(1 - \frac{\mu}{r}\right)\frac{c}{\delta},$$

where $\gamma(r) = (m + \sqrt{m^2 + 2r\sigma^2})/\sigma^2$ and $m = \mu - \sigma^2/2$. Since A_B increases in γ and r, and $\gamma(r)$ is also increasing in r, A_B must be increasing in r. More generally, Theorems 1 and 2 together give us the following result.

PROPOSITION 6 *For problem (11), the default-triggering level $A_B(r)$ is increasing in r while the optimal default time $\tau(r)$ and the firm's value $V(x,r)$ are both decreasing in r.*

Proposition 6 ensures that the monotonicity of optimal default time with respect to the interest rate holds for general asset processes and coupon and payout profiles. This result can be made even more general by allowing the firm to influence state x_t through an action λ_t that costs $f(\lambda_t) \ge 0$, with its effect determined by (9). The objective function is then

$$V(x,r) = \sup_{\tau \in \mathcal{T}} E_x \left[\int_0^\tau e^{-rt}(\delta(x_t) - c(x_t) - f(\lambda_t))dt \right]. \tag{11}$$

The instantaneous payoff $\delta(x_t) - c(x_t) - f(\lambda_t)$ is increasing in the state x_t, and so we can apply Theorem 3, which guarantees that all the conclusions of Proposition 6 hold even in this more general setting (with both control and stopping).[20]

3.4 Application: Optimal Growth and Capital Deepening

One of the very first and most important applications of optimal control to economics is to the optimal growth problem over an infinite horizon. Sensitivity of the optimal path with respect to discounting has been studied, in a deterministic context, by Brock and Mirman (1972), Mendelsohn and Sobel (1981), Amir, Mirman, and Perkins (1991), Amir (1996), and Quah and Strulovici (2009), amongst others. In our earlier paper (Quah and Strulovici (2009)), we showed in a deterministic, continuous-time model, that a lower discount rate leads to *capital deepening*, i.e., a higher level of capital stock at all times. The result relies on minimal assumptions; in particular, the representative agent's utility function need not

[20] Using the results of the present paper, Manso (2013, Proposition 9) obtains an equivalent of Proposition 6 for the equilibrium default boundary arising in a game between the borrower and credit-rating agencies. Proposition 6 can also be used to study when households default on their mortgages.

be concave and there are no shape restrictions on the production function, which can exhibit increasing returns to scale. We now extend this result to a stochastic context.[21]

An agent with initial capital $k_0 > 0$ manages his capital through consumption, savings, and efforts so as to maximize his expected lifetime utility. Formally, the agent solves

$$\max_{c} E\left[\int_0^\infty e^{-rs} u(c_s, k_s, s)ds\right] \tag{12}$$

$$\text{subject to } \; dk_t = H(k_t, c_t, t)dt + \sigma(k_t, t)dB_t,$$

where B is the standard Brownian motion, c is a finite dimensional control, and the drift and volatility functions H and σ are such that k is nonnegative.[22] The control c can capture not only consumption, but also leisure and effort, technological choices, and other realistic components. k_t is the capital available to the agent at time t. Note that we have stated the problem as one of pure control, without stopping (as is customary for growth models); to fit this within our framework, we assume that $u > 0$, so that stopping is never optimal. In the simplest case where utility depends only on consumption, we can write $u(k_s, c_s, s) = \bar{u}(c_s k_s)$, where $\bar{u} : \mathbb{R}_+ \to \mathbb{R}_+$ is the felicity function, so $c_s \in [0, 1]$ is consumption expressed as proportion of the capital stock and u is increasing in k_s provided $\bar{u}$ is increasing. Since the drift H is completely arbitrary (subject to the usual integrability conditions), we are allowing for arbitrary returns to scale in technology.

An application of Theorem 3 immediately yields the following result.

Proposition 7 *Suppose that $u(c, k, t)$ is increasing in k for all c, t. For any initial capital level k_0, the trajectory of capital $\{k_t\}_{t\geq 0}$ at the solution to (12) decreases in r almost surely.*

4 Stochastic Discounting

Dynamic stochastic models in macroeconomics and finance typically involve stochastic discounting, where the stochastic discount is a "state price deflator," representing some (random) intertemporal rate of substitution between (in discrete time) consecutive periods.[23]

Another, related, reason for considering stochastic discounting is to think of discounting as the borrowing rate of a firm. If there is procyclicality between the interest rate and

[21] Riedel and Su (2010) study the stochastic irreversible investment decision of single firm and obtain comparative statics of investment with respect to the cost of capital.

[22] For example, $k_t = \exp(K_t)$ for some Itô process.

[23] See, e.g., Duffie (2001).

the firm profit (or if on the contrary, its profits are countercyclical with respect to the general borrowing rate, or lending rate), the discounting rate and cash flows are correlated. Discounting may also be stochastic because the borrowing rate of the firm depends on its credit rating. In this case, there is an obvious connection between cash flows and discounting. Performance pricing loans and step up bonds are examples of debt contracts where the interest rate depends on some performance measure of the borrower.

The strongest extension of our order on discount functions to a stochastic setting is to require the ratio process $\beta(s)/\alpha(s)$ to be increasing in s path by path. However, we provide an example to show that even this is not sufficient to guarantee that the more patient agent stops later. To obtain positive results linking patience and stopping times, we need to introduce certain restrictions on cash flows; we provide two results of this type. The first theorem is based on some positive correlation condition. The second theorem assumes a single crossing condition on the utility function.

4.1 Counterexample: Discount-Cash Flow Correlation

There are three periods $\{0, 1, 2\}$ and two states of the world, "high" and "low," which are equally likely. The cash-flows in Periods 0 and 1 are -1 and -2, respectively. The cash-flow in Period 2 is 21 if the state is high, and -4 if the state is low.

Agent A uses the discount function $\alpha(t) = \left(\frac{1}{2}\right)^t$ and strictly prefers to make the investment, yielding a present value equal to $-1 + \frac{1}{2}(-2) + \frac{1}{4}\left(\frac{1}{2}(21) + \frac{1}{2}(-4)\right) > 0$.

Agent B faces the stochastic discount function $\beta(0) = 1$ and $\beta(1) = \frac{1}{2}$, and $\beta(2) = \frac{1}{4}$ if the state is high, and $\beta(2) = \frac{1}{2}$ if the state is low. Agent B is more patient than agent A: the ratio β/α is increasing over time for each realization of uncertainty, and strictly so if the state is low. Plainly it makes no sense for anyone to invest in this project and stop after periods 0 or 1. Assuming that the project is terminated after period 2, the present value to agent B is $-1 + \frac{1}{2}(-2) + \frac{1}{4}\frac{1}{2}(21) + \frac{1}{2}\frac{1}{2}(-4)$ which is still less than zero. So agent B will not invest in the project.

The key feature driving this counterexample is the negative correlation between B's patience and cash flows. It is possible to obtain a positive result once this phenomenon is excluded.

4.2 Positive Results with Stochastic Discounting

A *stochastic discount function* α is a strictly positive process adapted to the filtration $\mathcal{F}$ that is defined on $[0, \bar{t}]$, with $\alpha_0 = 1$ almost surely. An agent solves the following optimal stopping problem (using the notation of Section 2):

$$V(\alpha) = \sup_{\tau \in \mathcal{T}} E\left[\int_0^\tau \alpha(s) u_s ds.\right] \tag{13}$$

We would like to compare the optimal stopping behavior under α with that under another stochastic discount function β, also adapted to $\mathcal{F}$. Since $\mathcal{F}$ represents the information that agents receive, the fact that both are adapted to $\mathcal{F}$ prevents an agent from learning more about future cash flows through his discount process than another agent with the other discount process. Otherwise, a less patient agent could for example get more positive news about future cash flows than a more patient one, which could clearly revert the result. For simplicity, the termination value is set to zero.

Our first result gives a sufficient condition for comparative statics, under the assumption that $\alpha \prec \beta$ *almost surely*, i.e.,

$$Pr\left(\left\{\omega : \frac{\beta(s)}{\alpha(s)} \text{ is increasing on } [0, \bar{t}]\right\}\right) = 1;$$

we shall denote this simply by $\alpha \prec \beta$. We know from the counterexample that this alone is not sufficient to obtain a result, but this assumption along with the following *positive correlation condition* will suffice:

$$Corr\left(\frac{\beta(s)}{\alpha(s)}, \alpha(s) u_s 1_{s<\tau(\alpha)} \middle| \mathcal{F}_t\right) \geq 0 \quad a.s., \quad \text{for all } 0 \leq t < s < \bar{t}, \tag{14}$$

(where $\tau(\alpha)$ refers to the optimal stopping time under α). The condition (14) means that agent β is more likely to value future cash flows more relative to agent α when i) these cash-flows are higher, ii) agent α values these cash flows more, and iii) agent α is less likely to have stopped already.

THEOREM 4 *Suppose that $\alpha \prec \beta$ and that the positive correlation condition (14) holds. Then, for problem (13), $\tau(\alpha) \leq \tau(\beta)$ and $V(\alpha) \leq V(\beta)$.*

Assumption 14 involves the optimal stopping time $\tau(\alpha)$. The next result relies on assumptions directly expressed on primitives. Suppose that $u_s = u(x_s)$ where $x \in X \subset R$ is some one-dimensional state and suppose, also, that the discount factors are stochastic only through the state x, i.e., $\alpha_s = \alpha(x_s, s)$. Such a specification, where payoffs and the discount factor depend on some common underlying process arises, for instance, in affine term structure models (see Duffie (2001, Chapter 7)). In such a context the following result holds.

THEOREM 5 *Suppose that the discount functions α and β obey*

(a) $\alpha(x,\cdot) \prec \beta(x,\cdot)$ *for all* $x \in X$ *and*

(b) $\beta(x,s)/\alpha(x,s)$ *is increasing in* x *for all* $s \in [0,\bar{t}]$.

Then for any problem (13) such that $u_s = \bar{u}(x_s)$, *where* $\bar{u} : X \to R$ *has the single crossing property,*[24] *we obtain* $\tau(\alpha) \leq \tau(\beta)$ *and* $V(\alpha) \leq V(\beta)$.

Note that conditions (a) and (b) in Theorem 5 are together equivalent to requiring $\beta(x,s)/\alpha(x,s)$ to be increasing in (x,s). Condition (a) simply says that agent β is more patient than α in every state x. Condition (b) says that agent β's valuation of future cash flows relative to agent α's is greater in the higher states, which (because of the single crossing property on $\bar{u}$) are also the states with a positive payoff.

4.3 Application: Investment in a Consumption-Based Asset Pricing Model

In consumption-based asset-pricing models, cash flows are discounted according to a stochastic discount function that is proportional to the marginal utility of the representative agent.[25] Our results enable us to connect the patience of the representative agent with his level of risk aversion.[26]

In formal terms, the present value of a project available at time 0 is

$$V(u) = \max_{\tau \in \mathcal{T}} E\left[\int_0^\tau e^{-\rho t}\frac{u'(x_t)}{u'(x_0)}\pi_t dt\right],$$

where $u(\cdot)$ is the utility of the representative agent, ρ is his discount rate, $x_t > 0$ is the aggregate consumption at time t, π_t is the payoff generated by the project at time t, and τ is the termination time of the project, which is optimally chosen by its owner. Consider a project Π that, while active, entails an operating cost of c per unit of time, with the expected profit at time t depending on the state of the economy: $\bar{\pi}(x_t) = p(x_t) - c$, where p is increasing.[27] It is clear that $\bar{\pi}$ has the single crossing property; let $\bar{x}$ denote the crossing point, i.e., $p(\bar{x}) - c = 0$. Finally, suppose that the project has an initial fixed cost $I > 0$, so it will be initiated if and only if $V(u) \geq I$; in this case we say that the project is *u-valuable*.

[24] We mean that, whenever $\bar{u}(x) \geq (>)\, 0$, then $\bar{u}(\tilde{x}) \geq (>)\, 0$.

[25] See, for continuous-time versions of this result, Breeden (1979) and Duffie and Zame (1989).

[26] Nachman (1975) studies the impact of time and risk preferences in a setting with a single stochastic lump sum.

[27] It is assumed that the project is small compared to the size of the economy and, hence, has a negligible impact on $\{x_t\}$.

Without making any assumption on the aggregate consumption process x_t, the next result shows that Π is less likely to be undertaken if the representative agent is more risk averse and, assuming the project is underway, higher risk aversion leads to an earlier stopping time.

PROPOSITION 8 *Suppose that u is more risk averse to v in the sense that u has the higher coefficient of risk aversion, i.e., $-u''/u'(x) \geq -v''/v'(x)$ for all $x > 0$. Then project Π is v-valuable if it is u-valuable and the project will be abandoned later under v than under u.*

Proof. In the case where the representative agent's utility is u, the stochastic discount factor is given by $\alpha_t = \alpha(x_t, t)$, where

$$\alpha(x,t) = e^{-\rho t} \frac{u'(x)}{u'(x_0)}.$$

Similarly, in the case where it is v, we have $\beta(x,t) = e^{-\rho t} v'(x)/v'(x_0)$. Thus

$$\frac{\beta(x,t)}{\alpha(x,t)} = \frac{u'(x_0)}{v'(x_0)} \frac{v'(x)}{u'(x)}.$$

This ratio is independent of time and increasing in x. The latter holds because u is more risk averse than v and so v'/u' is increasing in x. Conditions (a) and (b) in Theorem 5 are both satisfied and thus the claim follows from an application of that theorem. ■

Note that a straightforward extension of this proposition holds if the "u" agent, in addition to being less risk averse than than the "v" agent, also has a lower discount rate.

5 Time Inconsistency and Optimal Stopping

One of the best documented and widely researched behavioral departures from the "standard model" is time inconsistency (see Frederick et al. (2002) for a survey). The agent's time preference is now characterized by a *discount bifunction* $\alpha(s,t) > 0$, which is defined for all $t \geq s$, with $\alpha(s,s) = 1$. At time s, the agent's time-preference for future cash flows is governed by the discount function $\alpha(s, \cdot)$.[28] An agent is *time consistent* if $\alpha(s,t')/\alpha(s,t)$ is independent of s for all $s \leq \min\{t,t'\}$. If this fails to hold, the agent is *time inconsistent*; this gives rise to a game between the selves of an agent at different times, who have misaligned preferences over cash flows. A *naive* agent ignores the fact that he is time inconsistent, while a *sophisticated* agent fully takes this problem into account. It is common in models of time inconsistency to assume that the different selves have the same discount function, so

[28] We use the term 'bifunction' to emphasize the dependence on two variables and to distinguish the more complex discount function of a time-inconsistent agent from the discount function of a time consistent agent.

$\alpha(s, s+k) = \alpha(s', s'+k)$ for all s, s', and $k > 0$, though (unless otherwise stated) we shall not be making this assumption.

Definition 2 (Generalized Patience Order) *Let β and α be two discount bifunctions. We say that β exhibits greater patience than α (denoted by $\alpha \prec^{inc} \beta$) if, at each time s,*

$$\frac{\beta(s,t)}{\alpha(s,t)} \quad \textit{is increasing in } t, \textit{ for all } t \geq s.$$

This definition simply says that one agent is more patient than another if his s-self is more patient than the s-self of the other agent (in the sense given by Definition 1), for all s. This section first establishes that the results of Section 2 extend to the case of naive agents. It then shows by example that when two agents are sophisticated, the more patient agent can stop earlier than the less patient one. However, when time inconsistency takes the form of *decreasing patience*, a condition satisfied by hyperbolic and quasi-hyperbolic discounting, the equilibrium stopping time of a sophisticated agent is increasing with patience, which restores the positive result obtained in Section 2. We also show that a naive agent stops later than a sophisticated one with the same discount bifunction, and that the stopping time falls with increasing sophistication. Our analysis will focus on pure stopping problems.

5.1 Naive agents

Consider the pure stopping problem introduced in Section 2. Given a discount function $\gamma : T \to \mathbb{R}_{++}$, we denote by $V_s(\gamma)$ the value function at time s of a *time-consistent* agent with discount function γ. Now consider a (possibly) time-inconsistent and naive agent with discount bifunction $\alpha : (s,t) \mapsto \alpha(s,t) \in \mathbb{R}$ for $s \leq t$. This agent believes at time s that his future selves all have the discount function $\alpha(s,\cdot)$. Therefore, he believes that stopping at time s is optimal if and only if $V_s(\alpha(s,\cdot)) = 0$. We assume that the agent stops whenever $V_s(\alpha(s,\cdot)) = 0$ and denote the stopping time arising from this rule by $\tau^N(\alpha)$.

It is easy to show that if A and B are naive agents with discount bifunctions α and β respectively and that B is more patient than A, then $\tau^N(\beta) \geq \tau^N(\alpha)$ with probability 1. Indeed, for any $s < \tau^N(\alpha)$, we have $V_s(\alpha(s,\cdot)) > 0$. That value function is exactly the same as the one computed in Section 2. Since $\alpha(s,\cdot) \prec \beta(s,\cdot)$, Theorem 1 applied to time s and discount functions $\beta(s,\cdot)$ and $\alpha(s,\cdot)$ implies that $V_s(\beta(s,\cdot)) \geq V_s(\alpha(s,\cdot))$. Therefore, B finds it strictly optimal to continue at time s whenever A does.

5.2 Sophisticated agents: A counterexample

With sophisticated agents, the multiple selves (one for each time) of an agent play an equilibrium where, at each time, exactly one of the selves makes a single move (a formal definition is given in Section 5.3). Consider the following sequence of cash flows

$$1,\ -M,\ \frac{M}{n},\ \frac{M}{n},\ldots,\ \frac{M}{n},\ 0,\ 0,\ldots$$

where $M = 10$, say, n is a large integer, and there are exactly $n+1$ periods where the cash flow equals $\frac{M}{n}$.

Agent A is time–consistent with discount function $\alpha(t) = \frac{1}{2^t}$. As is easily checked, A optimally takes the first cash flow and stops before getting $-M$.

Agent B has the same discount function as A at period 0 but suddenly becomes perfectly patient after period 1. Formally, $\beta(0,t) = \frac{1}{2^t}$ for all $t \geq 0$ and $\beta(s,t) = 1$ for all $s > 0$ and $t \geq s$. If B does not stop at time zero, his Period-1 self optimally incur the loss $-M$ in order to get future total benefits close to $(n+1)M/n > M$. This continuation is strictly worse, seen from the perspective of Period-0 B, than stopping at time zero. The more patient agent, B, stops at Period 0, while the less patient, A, stops at Period 1.

5.3 Sophisticated agents with Decreasing Patience

We now assume that time is discrete, with final period $\bar{t}$. Cash flows $\{u(t)\}_{t\leq\bar{t}}$ follow any arbitrary (possibly non Markovian) process with finite unconditional and conditional expectations. We focus on pure-strategy equilibria: at each time s, the s-self of the agent optimally decides whether or not to interrupt the cash flows, if no previous self has already done so, given the strategy profile of his future selves. Formally, let $T = \{0, 1, \ldots, \bar{t}\}$ denote the set of periods, and $\mathcal{H}^t$ denote the set of cash-flow histories from time 0 to time t excluded.[29] The set of outcomes is then given by $\Omega = \mathcal{H}^{\bar{t}+1}$. A pure-strategy equilibrium is a binary map $e : \{(t,h) : t \in T, h \in \mathcal{H}^t\} \mapsto \{C, S\}$ that determines, for each t and $h \in \mathcal{H}^t$, whether the t-self of the agent continues (C) or stops (S), given that previous selves have continued. Optimality requires that $e(t,h) = C$ only if the value function conditional on continuing is positive. (For simplicity, we rule out here any termination value, so that stopping yields zero.) We also assume, to pin down the equilibrium, that any indifferent agent stops. With these assumptions, the equilibrium is unique and can be computed by backward induction.

[29] By convention, we assume that at time t, the agent must decide whether to continue before observing the time t cash flow. The opposite convention, where current-period uncertainty is resolved before the agent's decision to continue, does not affect the analysis.

For $s \leq t$, let $\alpha(s,t)$ denote the discount factor of the s-self at time t. Let $\mathcal{S}(\alpha)$ denote the set of pairs (t,h), with $h \in \mathcal{H}^t$, at which the agent stops, in equilibrium. This set is completely determined by backward induction, and is independent of play history up to time t (excluded): whether the t-self decides to continue or stop depends only on the behavior of his future selves, not one the decisions or time preferences of his past selves. If $\omega \in \Omega$ denotes the realized sequence of cash flows, let $\omega^t \in \mathcal{H}^t$ denotes the truncation of ω up to time t. The equilibrium stopping time τ^α of an agent with discount function α is given by

$$\tau^\alpha(\omega) = \min\{t : (t, \omega^t) \in \mathcal{S}(\alpha)\}.$$

Let also τ_t^α denote the stopping time of the agent restricted to $\tau_t^\alpha \geq t$, i.e., when all selves before time t are forced to continue. From the previous observation, we have

$$\tau_t^\alpha(\omega) = \min\{t' : (t', \omega^{t'}) \in \mathcal{S}(\alpha) \text{ and } t' \geq t\}.$$

The stopping times are stochastic because the cash-flow histories (ω^t) are stochastic.

For any $s \leq t$ and history $h \in \mathcal{H}^t$, let $\bar{V}_s^t(h)$ denote the continuation value of the s-self of the agent, from time t onwards, if all selves before time t are forced to continue, and cash-flow history h has been realized. In particular,

$$\bar{V}_s^t(h) = E\left[\sum_{t'=t}^{\tau_t^\alpha - 1} \alpha(s,t')u(t')|h\right],$$

where by convention the sum equals zero if the index set is empty, and where $\tau = \bar{t} + 1$ if the agent never stops.

Given our counterexample, a positive result on the stopping behavior of sophisticated agents must involve additional restrictions. We shall impose the following condition.

Definition 3 (Decreasing Patience) *The discount bifunction α exhibits* decreasing patience *if, for any $s < s'$, $\alpha(s',\cdot) \prec \alpha(s,\cdot)$ on the time interval $T \cap [s', \bar{t}]$:*

$$\frac{\alpha(s,t)}{\alpha(s',t)} \text{ is increasing in } t \text{ for } t \geq s'.$$

It is easy to see that the property hold for the two most popular ways of modeling time-inconsistency.[30]

[30] Prelec (2004) introduces a similar condition which he calls "decreasing *im*patience." The difference in terminology comes from our different viewpoints: the agent exhibits more impatience between two dates as he gets closer to the earlier of the two dates, according to our patience ordering (that is what we call decreasing patience). However, fixing the time at which the agent sits, he exhibits less impatience between any two dates separated by a fixed duration, as these dates get further away in time. The conditions are consistent whenever $\alpha(s,t) = \tilde{\alpha}(t-s)$ for some function $\tilde{\alpha} : T \to \mathbb{R}_+$.

Example 1 (Hyperbolic Discounting) Decreasing patience holds for hyperbolic discounting, defined as

$$\alpha(s,t) = 1/(1 + k(t - s))$$

for $t > s$, where k is some positive parameter.

Example 2 (Quasi-Hyperbolic Discounting) Decreasing patience holds for quasi-hyperbolic discounting (also referred to as β-δ discounting), defined as

$$\begin{aligned} \alpha(s,t) &= 1 \ \textit{ for } t = s \textit{ and} \\ \alpha(s,t) &= \beta\, \delta^{t-s} \ \textit{ for } t > s, \end{aligned}$$

where $\delta \in (0,1)$ and $\beta \in (0,1)$.

To clarify the notation, we let $\bar{V}_s^t(h;\alpha)$ denote the value function $\bar{V}_s^t(h)$ for agent α.

Theorem 6 *Suppose that α and β exhibit decreasing patience and $\beta \succ^{inc} \alpha$. Then,*

$$\mathcal{S}(\beta) \subset \mathcal{S}(\alpha)$$

(thus $\tau^\beta \geq \tau^\alpha$) and, for all (s,h) with $h \in \mathcal{H}^s$, $\bar{V}_s^s(h;\beta) \geq \bar{V}_s^s(h;\alpha)$.

5.4 Comparing naive, sophisticated, and partially sophisticated agents

It is easy to see that between a sophisticated and a naive agent with the same discount bifunction α, *the sophisticated agent always stops before the naive agent*. This is because the naive agent thinks that he is time consistent. Equivalently, he believes that his future selves will do exactly what he wants them to do, in contrast to the sophisticated agent, who realizes that he has to compromise with his future selves. Therefore, the naive agent's *perceived* value function must be weakly higher than the sophisticated agent's, because it is optimal from his perspective. In particular, whenever the sophisticated agent has a positive continuation value, so will the naive agent. So he only wants to stop if the sophisticated agent wants to stop.[31] The result below formalizes this observation; $\tau^N(\alpha)$ ($\tau^S(\alpha)$) denote the optimal (equilibrium) stopping time of a naive (sophisticated) agent with discount function α.

[31] This argument is reminiscent of the control-sharing effects in Strulovici (2010): an individual incentives to pursue experimentation are always higher if he is a dictator than if he shares control with other members of society (see Theorems 1 and 8 in that paper). The relevant interpretation here is that the naive agent erroneously believes that he is a dictator over his future selves.

PROPOSITION 9 *Given any discount bifunction $\alpha(s,t)$, $\tau^S(\alpha) \leq \tau^N(\alpha)$ almost surely.*[32]

Using Theorem 6, we could extend this result to partial sophistication. Suppose that a time-inconsistent agent has a β-δ discount function. O'Donoghue and Rabin (2001) propose a model of *partial sophistication* in which the agent believes, at any time s, that his future self at time $s' > s$ has the discount function $\hat{\alpha}(s',t) = 1$ for $t = s'$ and $\hat{\alpha}(s',t) = \hat{\beta}\delta^{t-s'}$ for $t > s'$ and some $\hat{\beta} \in [\beta, 1]$. The case $\hat{\beta} = 1$ corresponds to a naive agent, while $\hat{\beta} = \beta$ (the true value) corresponds to a fully sophisticated agent. Thus, sophistication varies inversely with $\hat{\beta}$. We denote by $\tau_{\hat{\beta}}$ the stopping time of an agent whose sophistication coefficient is $\hat{\beta}$.

PROPOSITION 10 *If two agents have the same quasi-hyperbolic discount function α but different levels of sophistication, the more sophisticated agent will stop earlier, i.e., if $\hat{\beta}' > \hat{\beta}''$, then $\tau_{\hat{\beta}'} \geq \tau_{\hat{\beta}''}$.*

Proof. At any time s, the agent with sophistication coefficient $\hat{\beta} \in [\beta, 1]$ perceives the value of the stopping process in the same way as a (fully) sophisticated agent with the discount function $\gamma_{\hat{\beta}}$, where $\gamma_{\hat{\beta}}(s,t) = \alpha(s,t)$ for all $t \geq s$, and $\gamma_{\hat{\beta}}(s',t) = 1$ for $t = s' > s$ and $\gamma_{\hat{\beta}}(s',t) = \hat{\beta}\delta^{t-s'}$ for $t > s'$. Note that $\gamma_{\hat{\beta}}$ exhibits decreasing patience in the domain $\{(s',t) : t \geq s' \geq s\}$ and that $\gamma_{\hat{\beta}'} \succ^{inc} \gamma_{\hat{\beta}''}$ on the same domain. Therefore, by Theorem 6, if $\bar{V}_s^s(h; \gamma_{\hat{\beta}''}) > 0$ then $\bar{V}_s^s(h; \gamma_{\hat{\beta}'}) > 0$. So the less sophisticated will continue if the more sophisticated agent decides to continue. ■

5.5 Application: Social Experimentation and Altruism

It is well known in the dynamic social choice literature that pure altruism (i.e., a generation takes into account the fact that later generations care about even later generations) results in time inconsistency.[33] For example, consider a dynastic model where each generation lives for one period, and let x_t denote the vector of goods consumed by the time-t generation. This vector could include, among other items, the pollution level in period t, and hence embed trade-offs between the production of perishable goods and the production of long-lasting pollution. Suppose that each generation cares about the utility of future generations at a decaying exponential rate: using U_t to denote the total utility of the time-t generation, and

[32] In both this result and in Proposition 10, we assume that the payoffs received over time do not vary with the sophistication of the agents. In particular, we assume that the termination values equal zero in both cases. When stopping involves the agent moving on to some other activity, and if that activity is also subject to time consistency issues, then we have a situation where the termination value of a project is itself a dependent on the agent's sophistication; in such a situation the result can be overturned.

[33] See, e.g., Andreoni (1989) and Saez-Marti and Weibull (2005).

$u(\cdot)$ the immediate utility that it receives from immediate consumption of x_t, we have

$$U_t = u(x_t) + \sum_{t'>t} \gamma^{t'-t} U_{t'}$$

for some discount factor γ. This formula, which holds for each t, implies a time-inconsistent β-δ discount function in terms of the consumption stream, given by

$$U_t = u(x_t) + \frac{1}{2} \sum_{t'>t} (2\gamma)^{t'-t} u(x_{t'}).$$

(See Saez-Marti and Weibull (2005) for the derivation.) This utility function is well-defined provided $\gamma < 1/2$. Therefore, a decision maker who, at each period, represents the interests of the current generation, will make decisions taking into account the interests of future generations: the situation is formally equivalent to dynamic decision making by a time-inconsistent social planner.

Social discounting is of primary importance in evaluating the optimal level of experimentation in social/public policy. An application of particular interest concerns policies towards global warming (see, e.g., Karp (2005)). What is the equilibrium amount of social experimentation when the social planner exhibits time inconsistency, and how does it vary with patience? As we pointed out in Section 2.6, experimentation is a stochastic pure stopping problem when it is modeled as a bandit problem with one risky arm and one save arm; thus Theorem 6 provides a general answer: if the social planner exhibits decreasing patience, such as β-δ discounting, then increasing his overall patience level will result in an increase in social experimentation.[34]

6 Conclusion

Whether they are financial, economic, or psychological, time preferences play a central role in economic decisions. This paper provides a new methodology to systematically analyze the impact of discount rates on agents' values and decisions, yielding insights into the validity of common intuition on this issue. We provide examples with counter-intuitive outcomes arising in some plausible economic environments, as well as new conditions guaranteeing

[34] Another widely studied source of time inconsistency, in dynamic political models, comes from the fact that elected representatives and pivotal voters typically change over time (see Persson and Svensson (1989) and Tabellini and Alesina (1990)). This time inconsistency is particularly important in the context of social experimentation, as noted by Strulovici (2010) and Callander and Hummel (2013). Theorem 6 shows that as all decision makers become more patient, social experimentation lasts longer, provided that the sequence of discount functions used by the successive decision makers in office exhibits decreasing patience.

positive results in other, at least as natural, environments. Therefore, the main lesson of the paper is not simply that "lower discount rates imply higher valuation and later stopping." Instead, we provide a road map for economists that identify environments under which such intuitive statements hold, as well as key mechanisms that could lead to its violation.

Our analysis has been mostly limited to a single agent, although the arguments used for time inconsistent agents illustrate how our methodology can be employed to analyze multi-agent settings. In particular, our techniques provide a way of analyzing supermodular games in which the timing decisions of all agents exhibit complementarities, with or without cash-flow management. One promising application concerns the analysis of partnerships in stochastic matching models. For example, McAdams' (2011) description of how discounting affects the length of partnerships is consistent with our results, while Fujiwara-Greve and Yasuda (2011) obtain similar comparative statics results with respect to outside options in dynamic partnerships.[35] More generally, our techniques could be applied to analyze models of multi-agent search and experimentation; exploring those connections is a natural and promising area for future research.

7 Appendix

Proof of Theorem 1

Zero Termination Value We first prove Theorem 1 when the termination value is zero. The proof is based on several lemmas.

Lemma 1 *Let τ be any optimal stopping time for the stopping problem* (2). *Then*

$$P\left\{\omega \in \Omega : t \le \tau \text{ and } E\left[\int_t^\tau \alpha(s) u_s ds | \mathcal{F}_t\right] < 0\right\} = 0$$

Proof. Let $A_\varepsilon = \{\omega \in \Omega : t \le \tau \text{ and } E\left[\int_t^\tau \alpha(s) u_s ds | \mathcal{F}_t\right] \le -\varepsilon\}$ for $\varepsilon \ge 0$. We need to show that $P(A_0) = 0$. It suffices to show that $P(A_\varepsilon) = 0$ for all $\varepsilon > 0$ since this implies that $P(A_0) = P(\cup_{n=1}^\infty A_{1/n}) \le \sum_{n=1}^\infty P(A_{1/n}) = 0$. Suppose on the contrary that $P(A_\varepsilon) > 0$ for some $\varepsilon > 0$, and let $\tau^* = t 1_{\omega \in A_\varepsilon} + \tau 1_{\omega \notin A_\varepsilon}$. Since A_ε is $\mathcal{F}_t$ measurable, τ^* is a stopping time.

[35]The authors also observe that the comparative statics of the present paper holds in their setting.

Denoting $\Omega \setminus A_\varepsilon$ by A_ε^c, we have

$$
\begin{aligned}
E\left[\int_0^{\tau^*} \alpha(s)u_s ds\right] &= P(A_\varepsilon^c)E\left[\int_0^{\tau^*} \alpha(s)u_s ds|A_\varepsilon^c\right] + P(A_\varepsilon)E\left[\int_0^{\tau^*} \alpha(s)u_s ds|A_\varepsilon\right] \\
&= P(A_\varepsilon^c)E\left[\int_0^{\tau} \alpha(s)u_s ds|A_\varepsilon^c\right] + P(A_\varepsilon)E\left[\int_0^{\tau^*} \alpha(s)u_s ds|A_\varepsilon\right] \\
&\geq P(A_\varepsilon^c)E\left[\int_0^{\tau} \alpha(s)u_s ds|A_\varepsilon^c\right] + P(A_\varepsilon)(E\left[\int_0^{\tau} \alpha(s)u_s ds|A_\varepsilon\right] + \varepsilon) \\
&= E\left[\int_0^{\tau} \alpha(s)u_s ds\right] + P(A_\varepsilon)\varepsilon,
\end{aligned}
$$

which contradicts the optimality of τ. ■

Given an optimal stopping time τ for problem (2), we define $v : \mathbb{R}_+ \to \mathbb{R}$ by

$$v(s) = E[u_s 1_{s<\tau}]; \tag{15}$$

$v(s)$ is the expected payoff rate at time s, where the payoff is zero in the event that one has stopped before s (i.e., $s \geq \tau$). By the optimality of τ, Fubini's theorem applied to $V(\alpha) = E\left[\int_0^\tau \alpha(s)u_s ds\right]$ implies that

$$V(\alpha) = U(\tau(\alpha);\alpha) = \int_0^{\bar{t}} \alpha(s)v(s)ds. \tag{16}$$

The next result is a simple consequence of the fact that, at every point in time, the expected payoff of an optimizing agent looking forward must be non-negative.

Lemma 2 *For all t in $[0,\bar{t})$, $\int_t^{\bar{t}} \alpha(s)v(s)ds \geq 0$.*

Proof. By definition of v,

$$
\begin{aligned}
\int_t^{\bar{t}} \alpha(s)v(s)ds &= E\left[\int_t^{\bar{t}} \alpha(s)u_s 1_{s<\tau} ds\right] \\
&= E\left[E\left[\int_t^{\bar{t}} \alpha(s)u_s 1_{s<\tau} 1_{t<\tau} ds|\mathcal{F}_t\right]\right] \\
&= E\left[E\left[\int_t^{\tau} \alpha(s)u_s ds|\mathcal{F}_t\right] 1_{t<\tau}\right].
\end{aligned}
$$

Optimality of τ and Lemma 1 imply that the inner expectation is almost surely nonnegative if $t < \tau$. Therefore, the random variable $E[\int_t^{\bar{t}} \alpha(s)u_s 1_{s<\tau} ds|\mathcal{F}_t]1_{t<\tau}$ is always nonnegative, and so is its expectation. ■

The next lemma is proved in Quah and Strulovici (2009).

Lemma 3 *Suppose that γ and h are integrable real-valued functions defined on some compact interval $[x', x'']$ of $\mathbb{R}$, with γ increasing. If $\int_x^{x''} h(s)ds \geq 0$ for all x in $[x', x'']$, then*

$$\int_{x'}^{x''} \gamma(s)h(s)ds \geq \gamma(x') \int_{x'}^{x''} h(s)ds. \tag{17}$$

We now conclude the proof of Theorem 1 for the case of a zero termination value. First, note that $V(\beta) \geq V(\alpha)$ since $V(\beta) \geq U(\tau(\alpha);\beta)$, $V(\alpha) = U(\tau(\alpha);\alpha)$ and

$$U(\tau(\alpha);\beta) = \int_0^{\bar{t}} \beta(s)v(s)ds = \int_0^{\bar{t}} \frac{\beta(s)}{\alpha(s)} \alpha(s)v(s)ds \geq \int_0^{\bar{t}} \alpha(s)v(s)ds = U(\tau(\alpha);\alpha). \tag{18}$$

The inequality follows from Lemmas 2 and 3 and the last equality from (16).

It remains for us to establish that the smallest optimal stopping times are ranked, i.e., $\tau(\beta) \geq \tau(\alpha)$ almost surely. Suppose on the contrary that the set $\Omega' = \{\omega : \tau(\beta) < \tau(\alpha)\}$ has strictly positive probability. By mimicking at time $\tau(\beta)$ the argument that gave us (18), we have on Ω' (which is $\mathcal{F}_{\tau(\beta)}$-measurable)

$$E\left[\int_{\tau(\beta)}^{\tau(\alpha)} \beta(s)u_s ds \,\middle|\, \mathcal{F}_{\tau(\beta)}\right] \geq \frac{\beta(\tau(\beta))}{\alpha(\tau(\beta))} E\left[\int_{\tau(\beta)}^{\tau(\alpha)} \alpha(s)u_s ds \,\middle|\, \mathcal{F}_{\tau(\beta)}\right] \quad \text{a.s.}$$

This gives a lower bound for the continuation value with discount β at time $\tau(\beta)$, whenever $\tau(\beta) < \tau(\alpha)$. Since $\tau(\alpha)$ is the essential infimum of the optimal stopping times for the discount function α, $E[\int_{\tau(\beta)}^{\tau(\alpha)} \alpha(s)u_s|\mathcal{F}_{\tau(\beta)}] > 0$ almost surely on Ω'. Otherwise the agent would be weakly better off stopping at $\tau(\beta)$ when the event Ω' occurs, on a set of positive probability. Since the ratio β_t/α_t is also strictly positive for all t, this implies that

$$E\left[\int_{\tau(\beta)}^{\tau(\alpha)} \beta(s)u_s ds \, 1_{\Omega'}\right] > 0. \tag{19}$$

Therefore, the stopping time $\max\{\tau(\alpha), \tau(\beta)\}$ strictly dominates $\tau(\beta)$ for the discount function β, contradicting optimality of $\tau(\beta)$.

General Case Here we consider the case of an arbitrary nonnegative termination value process G_t. By absolute continuity of $\alpha(t)G_t$, the objective can be reexpressed as

$$\max_{\tau} E\left[\int_0^{\tau} \alpha(s)\left(u_s ds + dG_s + G_s \frac{d\alpha_s}{\alpha_s}\right)\right],$$

where the last two integrals are Lebesgue-Stieltjes integral.[36] This objective is similar to the case of a zero termination value, except that the new utility flow now directly depends on α

[36] See, e.g., Royden (2007) for a definition. The integrals are well defined whenever the integrating process has bounded variation and the integrand is measurable and bounded.

through the third term $G_s \frac{d\alpha_s}{\alpha_s}$. Optimality of τ implies that

$$E\left[\int_t^\tau \alpha(s)\left(u_s ds + dG_s + G_s \frac{d\alpha_s}{\alpha_s}\right)\right] \geq 0$$

for all $t < \tau$. We use the following generalization of Lemma 3:

Lemma 4 *Let H be a stochastic process of bounded variation such that $E[H_t] \leq E[H_{\bar{t}}]$ for all $t \in [0, \bar{t})$.*[37] *Then, for any positive, increasing, deterministic process γ, we have*

$$E\left[\int_0^{\bar{t}} \gamma_s dH_s\right] \geq \gamma(0)E[H(\bar{t}) - H(0)].$$

The proof is by integration by parts: we have

$$\int_0^{\bar{t}} \gamma_s dH_s = \gamma(0)(H(\bar{t}) - H(0)) + \int_0^{\bar{t}} (H_{\bar{t}} - H_s)d\gamma_s.$$

Taking expectations proves the lemma. Applying the lemma to $\gamma = \beta/\alpha$ and to the process H defined (up to a constant) by

$$dH_s = \alpha(s)\left(u_s ds + dG_s + G_s \frac{d\alpha_s}{\alpha_s}\right) 1_{s<\tau},$$

we obtain

$$E\left[\int_0^{\bar{t}} \beta(s)\left(u_s ds + dG_s + G_s \frac{d\alpha_s}{\alpha_s}\right)\right] \geq E\left[\int_0^{\bar{t}} \alpha(s)\left(u_s ds + dG_s + G_s \frac{d\alpha_s}{\alpha_s}\right)\right]. \tag{20}$$

Since β/α is increasing, $\frac{d\beta_s}{\beta_s} \geq \frac{d\alpha_s}{\alpha_s}$ for all s (the inequality holds even when β and α are discontinuous). This, combined with (20) and the fact that G is nonnegative, implies that

$$E\left[\int_0^\tau \beta(s)\left(u_s ds + dG_s + G_s \frac{d\beta_s}{\beta_s}\right)\right] \geq E\left[\int_0^\tau \alpha(s)\left(u_s ds + dG_s + G_s \frac{d\alpha_s}{\alpha_s}\right)\right] \geq 0.$$

Therefore, $V(\beta) \geq V(\alpha)$. The rest of the proof is similar to the case of a zero termination value: on the set Ω', we can show that continuing until at least $\tau(\alpha)$ is strictly beneficial, using an inequality similar to (20) for continuation values evaluated at time $\tau(\beta)$ and contradict optimality of $\tau(\beta)$. ■

Proof of Proposition 2

[37] If $\bar{t}$ is infinite, $H_{\bar{t}}$ is the limit of H_t as $t \to \infty$, and is assumed to exist a.s.

Suppose that $\alpha \prec \beta$. Theorem 1 implies that $\tau(\alpha) \leq \tau(\beta)$ with probability 1. This means that, at time $\tau(\alpha)$, continuing until $\tau(\beta)$ is optimal for agent β and in particular yields positive value. Since $\beta \prec \gamma$, by an argument analogous to that leading to (18), we know that the value of continuing until $\tau(\beta)$ is positive for the γ agent, and thus preferred to stopping at $\tau(\alpha)$. ■

Counterexample of Section 3.2

Suppose that the agent discounts future cash flows at rate r. The Bellman equation for $x \in [0, 1]$ is given by

$$0 = \max\{-rV(x) + 1 + V'(x); -rV(x) - 0.01 + 2V'(x)\},$$

where the first term corresponds to the control $\lambda = 1$ and the second term to $\lambda = 2$. Thus it is optimal to choose $\lambda = 2$ if and only if

$$V'(x) \geq 1.01. \tag{21}$$

We will check that for r small enough, the solution to the HJB equation is maximized by the second term, which corresponds to control $\lambda = 2$. The general solution to

$$-rV(x) - 0.01 + 2V'(x) = 0$$

is

$$V(x) = \frac{-0.01}{r} + c \exp\left(\frac{rx}{2}\right). \tag{22}$$

The boundary condition is $V(1) = M[1 - \exp(-rT)]/r$, where T is the time it takes for x to go from 1 to 10 under control $\lambda = 1$ (so $T = 9$). This implies that

$$c = \frac{e^{-r/2}}{r}\left[M\left((1 - e^{-rT}\right) + 0.01\right], \tag{23}$$

and hence that

$$V'(x) = \frac{\exp(r(x-1)/2)}{2}\left[M\left(1 - e^{-rT}\right) + 0.01\right] \geq \frac{\exp(-r/2)}{2}\left[M\left(1 - e^{-rT}\right) + 0.01\right], \tag{24}$$

which is uniformly, arbitrarily large for M large and r fixed. In particular, at a given discount rate $\tilde{r}$, there is $\tilde{M}$ such that (21) is satisfied. This shows that the function V defined by (22), with c defined in (23), and the control $\lambda = 2$ solve the HJB equation, and hence setting $\lambda(x) = 2$ for all $x \in [0, 1)$ is optimal for this agent.[38]

[38]From straightforward inspection of the righthand side of (24), one can show the following, sharper result: for all $M > 1$, there exists $\bar{r}$ such that for all $r' < \bar{r}$, it is optimal to set $\lambda = 2$ for all $x \leq 1$.

On the other hand, for a fixed $\hat{x} \in (0,1)$ and with $M = \tilde{M}$, it is clear from (24) that (21) will be violated if r is sufficiently high above $\tilde{r}$. Such an agent's optimal control will involve choosing $\lambda(x) = 1$ for at least some $x \in [0,1)$; consequently, she will stop *later* than the agent with discount rate r. ■

Proof of Theorem 3

Recall that $V(x,r)$ is the continuation value at time 0 when the state is x and the discount rate is r. Our first objective is to prove the following result.

LEMMA 5 *$V(x,r)$ is submodular in (x,r).*

Proof. Given the discount rate r, we denote the optimal control at state $\hat{x}$ and time t by $\lambda(\hat{x},t)$ and let $\tau(x)$ denote the optimal stopping time with *initial* state x. We claim that if $y > x$, then

$$E\left[\int_0^{\tau(y)} se^{-rs}u(y_s,\lambda(y_s,s),s)ds\right] \geq E\left[\int_0^{\tau(x)} se^{-rs}u(x_s,\lambda(x_s,s),s)ds\right] \tag{25}$$

To see this, define

$$h_s = E\left[u(y_s,\lambda(y_s,s),s)1_{\tau(y)\geq s} - u(x_s,\lambda(x_s,s),s)1_{\tau(x)\geq s}\right], \tag{26}$$

where the expectation is taken with respect to the Wiener measure (the probability measure generated by the paths of Brownian Motion). Notice that *for all t*, $\int_t^\infty e^{-rs}h_s ds \geq 0$. This is because (proceeding as in the proof of Lemma 2) we can write

$$\int_t^\infty e^{-rs}h_s ds = E\left[V(y_t,t,r)1_{t\leq\tau(y)} - V(x_t,t,r)1_{t\leq\tau(x)}\right];$$

since $V(y_t,t,r) \geq V(x_t,t,r) \geq 0$ and $1_{t\leq\tau(y)} \geq 1_{t\leq\tau(x)}$ (the latter because it takes more time to hit the lower boundary when starting from a higher level), the difference inside the expectation is nonnegative almost surely and, therefore, so is the expectation. Using Lemma 3, this time with the function $\gamma(s) = s$, we conclude that

$$\int_0^\infty se^{-rs}h_s ds \geq 0.$$

This inequality can in turn be rewritten as (25).

By a generalized envelope theorem (see Milgrom and Segal, 2002),[39]

$$V_r(x,r) = \frac{\partial}{\partial r}E\left[\int_0^\tau e^{-rt}u(x_t,\lambda_t,t)dt\right],$$

[39] The parameter r does not affect the distribution of the underlying process, which justifies the application of the theorem.

evaluated at the optimal controls λ and τ. Computing the derivative explicitly,

$$V_r(x,r) = E\left[\int_0^\tau (-t)e^{-rt}u(x_t,\lambda_t,t)dt\right].$$

This implies that for $y > x$,

$$V_r(y,r) - V_r(x,r) = -E\left[\int_0^{\tau(y)} se^{-rs}u(y_s,\lambda(y_s,s))ds\right] + E\left[\int_0^{\tau(x)} se^{-rs}u(x_s,\lambda(x_s,s))ds\right],$$

which is less than zero from (25). ■

We can now conclude the proof of Theorem 3.

Proof. The smoothness of the value function V implies that it satisfies the following Hamilton-Jacobi-Bellman (HJB) equation.

$$0 = \sup_\lambda \left\{u(x,\lambda,t) + \mu(x,\lambda,t)V_x(x,t,r) + V_t(x,t,r) + \frac{1}{2}\sigma^2(x)V_{xx}(x,t,r) - rV(x,t,r)\right\} \tag{27}$$

for all $x > a(t)$. We first make a change of variable, replacing the initial control variable by a direct determination of the drift. Let $\upsilon(x,m,t) = \sup\{u(x,\lambda,t) : \mu(x,\lambda,t) = m\}$, with $\upsilon(x,m,t) = -\infty$ if the defining set is empty. Thus, $\upsilon(x,m,t)$ is the maximal utility one can achieve at (x,t) while providing a drift m to the state. We assume that this maximum is achieved whenever υ is finite, which is easy to guarantee by simple assumptions on the primitives.[40] The optimal control problem can be re-expressed in terms of m:

$$0 = \sup_m \left\{\upsilon(x,m,t) + mV_x(x,t,r) + V_t(x,t,r) + \frac{1}{2}\sigma^2(x)V_{xx}(x,t,r) - rV(x,t,r)\right\}. \tag{28}$$

Any drift-control process $\{m_t\}$ that achieves at all times the supremum in (28) is an optimal control. Reciprocally, a control process that fails to achieve the supremum in (28) on a set of positive measure is suboptimal.[41] We assume that the set of maximizers of (28) is nonempty and compact for all (x,t).[42] Let $m(x,t,r)$ be the smallest maximizer of (28) for state x, time t, and discount r. By Lemma 5, $V(x,t,r)$ is submodular in (x,r), so the objective function in (28) is submodular in (m,r). By standard monotone comparative statics arguments (see Milgrom and Shannon (1994)), we know that $m(x,t,r)$ is decreasing in r.

[40] For example, if μ and u are continuous in λ and λ lies in a compact set, then the set $\{\lambda : \mu(x,\lambda,t) = m\}$ is closed and therefore compact, and $\lambda \mapsto u(x,\lambda,t)$ achieves its maximum on that set.

[41] Precisely, one may show that any HJB-maximizing control strictly dominates any control that does not solve (28) for a time set of positive measure.

[42] This is guaranteed if υ is continuous in m and m lies in a compact set for each (x,t). This, in turn, is guaranteed, by the maximum theorem, if u and μ are continuous in λ and λ lies in a compact set.

We now claim that $X(r)$, the path obtained when using the control m, is decreasing in r. Given an initial state x, let x'_t be some realized optimal trajectory for rate r' (and x_t the corresponding optimal path for r). For $r < r'$ and for all $t > 0$, we know that $m(x'_t, t, r) \geq m(x'_t, t, r')$. Given this, the paths of x_t and x'_t cannot cross because, given that they are continuous, they must first touch, and the moment they touch (i.e. have the same value of the state variable at the same time), the path of the lower discount rate receives a higher drift.[43] Thus $X(r)$ is decreasing in r. Lastly, we know from Theorem 2 that the boundary $a(r)$ is increasing in r. The previous observations imply that $x_t(r)$ hits $a(r)$ later than $x_t(r')$ hits $a(r')$, path by path, so $\tau(r) \geq \tau(r')$. ■

Proof of Proposition 5

We shall prove a bit more than what is stated in the proposition. We shall identify an optimal control function $\lambda(x, t, r)$ that is decreasing in r and also (in order to properly link up this result with Theorem 3) we shall also show that $\mu(x, \lambda(x, t, r), t) = m(x, t, r)$, where m is defined in the proof of Theorem 3.

It is straightforward to check that the objective function in (27) is supermodular in λ and has decreasing differences in (λ, r), given our assumptions and using the submodularity of V in (x, r) (Lemma 5). Therefore, the set of maximizers of (27) form a lattice and are decreasing in the strong set order (see Topkis (1978) or Milgrom and Shannon (1994)). Let $\lambda(x, t, r)$ be the smallest element in the set of maximizers; $\lambda(x, t, r)$ is then decreasing in r.

To see that $\mu(x, \lambda(x, t, r), t) = m(x, t, r)$, note that, given (x, t), there must be $\tilde{\lambda}$ such that $\mu(x, \tilde{\lambda}, t) = m(x, t, r)$ and $\tilde{\lambda}$ must also maximize the objective function in (27). Therefore, $\tilde{\lambda} \geq \lambda(x, t, r)$, by the definition of $\lambda(x, t, r)$. Since μ is increasing, $m(x, t, r) = \mu(x, \tilde{\lambda}, t) \geq \mu(x, \lambda(x, t, r), t)$. On the other hand, because $\lambda(x, t, r)$ maximizes the objective in (27), $\mu(x, \lambda(x, t, r), t)$ must maximize the objective in (28), so $\mu(x, \lambda(x, t, r), t) \geq m(x, t, r)$ by the definition of m. ■

Extension of Section 3.2 to General Discount Functions

Suppose that α and β have absolutely continuous logarithms, so $\alpha(t) = \alpha(0) \exp\left(-\int_0^t r_\alpha(s) ds\right)$, with a similar expression for β. In this case, $\alpha \prec \beta$ is equivalent to $r_\beta(t) \leq r_\alpha(t)$ for all $t \in T$. Let

$$\gamma_\nu(t) = \alpha(t) \left(\frac{\beta(t)}{\alpha(t)}\right)^\nu,$$

[43] Continuity plays an important role in this argument. In discrete time, or if jumps were allowed, the paths could cross each other.

for $\nu \in [0,1]$. It is easy to check that $\gamma_0 = \alpha$, $\gamma_1 = \beta$, and $\gamma_\nu \prec \gamma_{\nu'}$ for all $\nu \leq \nu'$.

Lemma 5 can then modified as follows: *let $V(x,\nu)$ denote the value function of an agent with discount function γ_ν, starting from state x; then, $V(x,\nu)$ is submodular.* Indeed, the envelope theorem implies that

$$\frac{\partial V(x,\nu)}{\partial \nu} = E\left[\int_0^{\tau_\nu}\left(-\log\left(\frac{\beta(t)}{\alpha(t)}\right)\right)\gamma_\nu(t)u(x_t,\lambda_\nu(t))dt\right].$$

To show that $V_\nu(y,\nu) - V_\nu(x,\nu) \leq 0$ for $y \geq x$, we use the same argument as in Lemmas 5, with the difference that, instead of $\int_0^\infty se^{-rs}h_s ds \geq 0$, we show that

$$\int_0^\infty \log\left(\frac{\beta(s)}{\alpha(s)}\right)\gamma_\nu(s)h_s ds \geq 0.$$

The result is obtained by an application of Lemma 3 to the increasing function $s \mapsto \log\left(\beta(s)/\alpha(s)\right)$.

The HJB equations are modified only by replacing r with the time varying rate $r_\nu(t) = r_\alpha(t) + \nu(r_\beta(t) - r_\alpha(t))$. The claims in Theorem 3 and Proposition 5 can then be extended to the case of general discount functions, with the proof proceeding in the same way as the proof we have already given for the case of constant discount rate. ■

Proof of Theorem 4

We first show that if it is strictly optimal for A to continue beyond $t = 0$, then it is also strictly optimal for B to do so. Let $V_t = E[\int_t^{\tau(\alpha)}\alpha(s)u_s ds|\mathcal{F}_t]$ denote the value function of agent α at time $t < \tau(\alpha)$. Optimality of $\tau(\alpha)$ implies that $E[V_t 1_{t<\tau(\alpha)}] \geq 0$ and, hence (see Lemma 2)

$$\int_t^{\bar{t}} E[\alpha(s)u_s 1_{s<\tau(\alpha)}]ds \geq 0$$

for all $t \leq \bar{t}$. Since Agent B can always choose the stopping time $\tau(\alpha)$, we have (using, again, the same manipulation as in the proof of Theorem 1),

$$V(\beta) \geq \int_0^{\bar{t}} E\left[\frac{\beta(s)}{\alpha(s)}\alpha(s)u_s 1_{s<\tau(\alpha)}\right]ds \geq \int_0^{\bar{t}} E\left[\frac{\beta(s)}{\alpha(s)}\right] \times E[\alpha(s)u_s 1_{s<\tau(\alpha)}]ds,$$

where the second inequality comes from condition (14) evaluated at $t = 0$. The function $E[\beta(s)/\alpha(s)]$ is increasing in s, since $\beta(s)/\alpha(s)$ is increasing path by path. Applying Lemma 3, we conclude that $V(\beta) \geq V(\alpha)$. In particular, if $V(\alpha)$ is strictly positive, then so is $V(\beta)$; thus the β agent will not stop at $t = 0$ if agent α does not stop at $t = 0$. For any time $\tilde{t} > 0$ and realization such that $\tau(\alpha) > \tilde{t}$, one shows similarly that $\tau(\beta) > \tilde{t}$, using again (14), this time evaluated at $t = \tilde{t}$, and the fact that $E[\beta(s)/\alpha(s)|\mathcal{F}_{\tilde{t}}]$ is increasing in s. We conclude that $\tau(\beta) \geq \tau(\alpha)$ almost surely. ■

Proof of Theorem 5

The proof is identical to that of Theorem 4, except for the following. Single-crossing of the utility function u and monotonicity of the ratio β/α with respect to x (condition (b)) imply that

$$E\left[\frac{\beta(x_s,s)}{\alpha(x_s,s)}\alpha(x_s,s)\bar{u}(x_s)1_{s<\tau}\right] \geq \frac{\beta(\bar{x},s)}{\alpha(\bar{x},s)} \times E[\alpha(x_s,s)\bar{u}(x_s)1_{s<\tau}],$$

where $\bar{x}$ is a crossing point of u. Condition (a) then implies that we can apply Lemma 3 as in the proof of Theorem 4, and reach the same conclusion. ■

Proof of Theorem 6

We start with the following lemma.

Lemma 6 *Suppose that α exhibits decreasing patience. Then for any $s \leq t$ and history $h \in \mathcal{H}^t$, we have $\bar{V}_s^t(h) \geq 0$.*

Proof. We show this by backward induction on s that for any $t \geq s$ and $h \in \mathcal{H}^t$. For $s = \bar{t}$ (and thus, $t = s$), the final self can always avoid the last cash flow and get 0, which proves the condition, since $\tau_{\bar{t}}^{\alpha}$ is chosen by the last self. Suppose now that the induction hypothesis holds for all $s' > s$. We will show that it holds for s. For $t = s$, the inequality holds because the s self can always interrupt cash flows immediately and get 0, and would indeed do so if his present value, conditional on continuing, were negative. Now suppose that $t > s$ and $h \in \mathcal{H}^t$. We have

$$\bar{V}_s^t(h) = \sum_{t'=t}^{\bar{t}} \alpha(s,t')v(t'),$$

where $v(t') = E\left[u(t')1_{t'<\tau_t^{\alpha}} \,\middle|\, h\right]$. We now show that for all $\tilde{t} \geq t$,

$$\sum_{t'=\tilde{t}}^{\bar{t}} \alpha(t,t')v(t') \geq 0. \tag{29}$$

We can write this expression as

$$E\left[1_{\tilde{t}\leq\tau_t^{\alpha}} \sum_{t'=\tilde{t}}^{\tau_t^{\alpha}-1} \alpha(t,t')u(t')\middle|\, h\right] = E\left[1_{\tilde{t}\leq\tau_t^{\alpha}} E\left[\sum_{t'=\tilde{t}}^{\tau_{\tilde{t}}^{\alpha}-1} \alpha(t,t')u(t')\middle|\, \tilde{h}\right]\middle|\, h\right],$$

where $\tilde{h}$ is the history at time $\tilde{t}$, and we use the fact that $\tau_t^\alpha = \tau_{\tilde{t}}^\alpha$ whenever $\tilde{t} \le \tau_t^\alpha$ (see observation above). The inner expectation is equal to $\bar{V}_t^{\tilde{t}}(\tilde{h})$ and is nonnegative, by the induction hypothesis. This shows (29).

Applying the discrete version of Lemma 3 (see Quah and Strulovici (2009)), and using that $\alpha(s, \cdot) \succ \alpha(t, \cdot)$, we conclude that

$$\sum_{t'=t}^{\bar{t}} \alpha(s,t')v(t') \ge \alpha(s,t) \sum_{t'=t}^{\bar{t}} \alpha(t,t')v(t') \ge 0,$$

which proves the induction step and the lemma. ■

We can now prove the theorem.

Proof. Let $\mathcal{S}^s(\alpha)$ denote the set of pairs (t,h) in $\mathcal{S}(\alpha)$ such that $t \ge s$, i.e., the pairs of time-histories, after time s, at which the agent stops. We show by backward induction on s that $\mathcal{S}^s(\beta) \subset \mathcal{S}^s(\alpha)$. The inclusion holds as an equality for $s = \bar{t}$, since both agents continue at $\bar{t}$ if and only if the expected utility of that period is positive, and get that expected utility, multiplied by their discount factor. Suppose that the inclusion holds for all $s' > s$. We will show that it also holds for s. Consider any $h \in \mathcal{H}^s$. If A (the α agent) stops at (s,h), then there is nothing to show. Suppose that A continues at (s,h). We need to show that B (the β agent) also continues. Since A continues, we have, conditional on h, $\tau_s^\alpha = \tau_{s+1}^\alpha$, and

$$E\left[\sum_{t=s}^{\tau_s^\alpha - 1} \alpha(s,t)u(t) \middle| h\right] > 0.$$

If B continues, he gets

$$E\left[\sum_{t=s}^{\tau_{s+1}^\beta - 1} \beta(s,t)u(t) \middle| h\right]. \tag{30}$$

By the induction hypothesis, $\tau_{s+1}^\beta \ge \tau_{s+1}^\alpha = \tau_s^\alpha$. Therefore, (30) may be reexpressed as

$$E\left[\sum_{t=s}^{\tau_s^\alpha - 1} \beta(s,t)u(t) + \bar{V}_s^{\tau_s^\alpha}(h(\tau_s^\alpha);\beta) \middle| h\right],$$

where $h(\tau_s^\alpha)$ is the (random) realized history up until time τ_s^α. We show that the first term is strictly positive and that the second term is nonnegative. The first term equals $\sum_{t=s}^{\bar{t}} \beta(s,t)v(t)$, where $v(t) = E[u(t)1_{t<\tau_s^\alpha}|h]$. We now show that for all $s' \ge s$,

$$\sum_{t=s'}^{\bar{t}} \alpha(s,t)v(t) \ge 0. \tag{31}$$

Proceeding as in Lemma 6, that expression is equal to

$$E\left[1_{s'\leq\tau_s^\alpha}E\left[\sum_{t=s'}^{\tau_{s'}^\alpha-1}\alpha(s,t)u(t)\middle|h'\right]\middle|h\right],$$

where h' is the history at time s'. The inner expectation is equal to $\bar{V}_s^{s'}(h';\alpha)$, and is nonnegative by Lemma 6. This shows (31). An application of the discrete version of Lemma 3, along with the fact that $\beta(s,\cdot)\succ\alpha(s,\cdot)$, then shows that

$$\sum_{t=s}^{\bar{t}}\beta(s,t)v(t)\geq\sum_{t=s}^{\bar{t}}\alpha(s,t)v(t) \tag{32}$$

and, therefore, that the first term is strictly positive. The second term is nonnegative, because $\bar{V}_s^t(h;\beta)$ is always nonnegative, from Lemma 6.[44] Therefore, it is strictly optimal for B to continue at (s,h), which shows the induction step. Finally, (32) implies that $V_s^s(h;\beta)\geq V_s^s(h,\alpha)$, which shows the second part of the theorem. ■

References

Akerlof, G. (1991) "Procrastination and Obedience," *American Economic Review*, Vol. 81, No. 2, pp. 1–19.

Amir, R., Mirman, L., and Perkins, W. (1991) "One-Sector Nonclassical Optimal Growth: Optimality Conditions and Comparative Dynamics," *International Economic Review*, Vol. 32, pp. 625–644.

Amir, R. (1996) "Sensitivity Analysis of Multisector Optimal Economic Dynamics," *Journal of Mathematical Economics*, Vol. 25, pp. 123–141.

Andreoni, J. (1989) "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, Vol. 97, pp. 1447–1458.

Arrow, K., Levhari, D. (1969) "Uniqueness of the Internal Rate of Return with Variable Life of Investment," *Economic Journal*, Vol. 79, pp. 560-566.

Bergemann, D., Välimäki, J. (2006) "Bandit Problems," Discussion Paper No. 1551, Cowles Foundation, Yale University.

[44]Recall that nonnegativity of $\bar{V}_s^t$ was obtained by backward induction, and holds for all histories leading up to time t.

Bolton, P., Harris, C. (1999) "Strategic Experimentation," *Econometrica*, Vol. 67, No. 2, pp. 349–374.

Breeden, D. (1979) "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, Vol. 7, pp. 265–296.

Brock, W., Mirman, W. (1972) "Optimal Economic Growth and Uncertainty: The Discounted Case," *Journal of Economic Theory*, Vol. 4, pp. 479–513.

Callander, S., Hummel, P. (2013) "Preemptive Experimentation under Alternating Political Power," Working Paper, Stanford University.

Cochrane, J. (2011) "Presidential Address: Discount Rates," *Journal of Finance*, Vol. 46, No. 4, pp. 1047–1108.

Cohen, A., Solan, E. (2012) "Bandit Problems with Lévy Payoff Processes," forthcoming in *Mathematics of Operations Research.*

Duffie, D (2001) *Dynamic Asset Pricing Theory*, Third Edition. Princeton University Press.

Duffie, D., Zame, W. (1989) "The Consumption-Based Capital Asset Pricing Model," Vol. 57, No. 6, pp. 1279–1297.

Durrett, R. (2004) *Probability: Theory and Examples*, Third Edition. Duxbury Press.

Fleming, W. H., Soner, H.M. (1993) *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York.

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

Frostig, E., Weiss, G (1999) "Four proofs of Gittins multiarmed bandit theorem, *Applied Probability Trust.*

Fujiwara-Greve, T., Yasuda, Y. (2011) "Repeated Cooperation with Outside Options," Working Paper, Keio University.

Karp, L. (2005) "Global Warming and Hyperbolic Discounting," *Journal of Public Economics*, Vol. 89, pp. 261–282.

KELLER, G., RADY, S., AND CRIPPS, M. (2005) "Strategic Experimentation with Exponential Bandits," *Econometrica*, Vol. 73, No. 1, pp. 39–68.

LAMBERTON, D., LAPEYRE, B. (2007) *Introduction to Stochastic Calculus Applied to Finance*, Second Edition. Chapman and Hall.

LELAND, H. (1994) "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *Journal of Finance*, Vol. 49, pp. 1213-1252.

MANSO, G. (2013) "Feedback Effects of Credit Ratings," *forthcoming in the Journal of Financial Economics.*

MANSO, G., STRULOVICI, B., AND TCHISTYI, A. (2010) "Performance-Sensitive Debt," *Review of Financial Studies*, Vol. 23, No. 5, pp. 1819–1854.

MCADAMS, D. (2011) "Performance and Turnover in a Stochastic Partnership," *American Economic Journal: Economics*, Vol. 3, No. 4, pp. 107–142.

MENDELSOHN, R., SOBEL, M. (1981) "Capital Accumulation and the Optimization of Renewable Resource Models," *Journal of Economic Theory*, Vol. 23, pp. 243–260.

MILGROM, P., SEGAL, I. (2002) "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, Vol. 70, No. 2, pp. 583–601.

MILGROM, P., SHANNON, C. (1994) "Monotone Comparative Statics," *Econometrica*, Vol. 62, No. 1, pp. 157–180.

NACHMAN, D. (1975) "Risk Aversion, Impatience, and Optimal Timing Decisions," *Journal of Economic Theory*, Vol. 11, pp. 1960–246.

O'DONOGHUE, T., RABIN, M. (1999) "Doing It Now or Later," *American Economic Review*, Vol. 89, No. 1, pp. 103–124.

O'DONOGHUE, T., RABIN, M. (2001) "Choice and Procrastination," *Quarterly Journal of Economics*, Vol. 116, No. 1, pp. 121–160.

PERSSON, T., SVENSSON, L. (1989) "Why a Stubborn Conservative would Run a Deficit: Policy with Time-Inconsistent Preferences," *Quarterly Journal of Economics*, Vol. 104, pp. 325–345.

PRELEC, D. (2004) "Decreasing Impatience: A Criterion for Non-Stationary Time Preference and "Hyperbolic" Discounting," *Scandinavian Journal of Economics*, Vol. 106, No. 3, pp. 511–532.

QUAH, J., STRULOVICI, B. (2009) "Comparative Statics, Informativeness, and the Interval Dominance Order," *Econometrica*, Vol. 77. No. 6, pp. 1949–1992.

RIEDEL, F., SU, X. (2010) "On Irreversible Investment," forthcoming, *Finance and Stochastics*.

ROGERSON, R., SHIMER, R., AND WRIGHT, R. (2005) "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, Vol. 18, pp. 959–988.

ROTHSCHILD, M. (1974) "A Two-Armed Bandit Theory of Market Pricing" *Journal of Economic Theory*, Vol. 9, pp. 185–202.

ROYDEN, H. (2007) *Real Analysis*, 4^{th} Edition, Prentice Hall.

SAEZ-MARTI, M, WEIBULL, J. (2005) "Discounting and Altruism to Future Decision-Makers," *Journal of Economic Theory*, Vol. 122, pp. 254–266.

SAMUELSON, P. (1937) "Some Aspects of the Pure Theory of Capital," *Quarterly Journal of Economics*, pp. 469–496.

SEN, A. (1975) "Minimal conditions for monotonicity of capital value," *Journal of Economic Theory*, Vol. 11, pp. 340-355.

STRULOVICI, B. (2010) "Learning While Voting: Determinants of Collective Experimentation," *Econometrica*, Vol. 78, No. 3, pp. 933–971.

STRULOVICI, B., SZYDLOWSKI, M. (2012) "On the Smoothness of Value Functions and the Existence of Optimal Strategies," Working Paper, Northwestern University.

TABELLINI, G., ALESINA, A. (1990) "Voting on the Budget Deficit," *American Economic Review*, Vol. 80, pp. 37–49.

TOPKIS, D. (1978) "Minimizing a Submodular Function on a lattice," *Operations Research*, Vol. 26, pp. 305–321.

WHITTLE, P. (1980) "Multi-armed bandits and the Gittins index," *Journal of the Royal Statistical Society*, Series B, Vol. 42, pp. 143–149.