Working Papers

Institute of Mathematical Economics

429

April 2010

Optimal Stopping Under Ambiguity In Continuous Time

Frank Riedel

IMW · Bielefeld University
Postfach 100131
33501 Bielefeld · Germany

email: imw@wiwi.uni-bielefeld.de
http://www.wiwi.uni-bielefeld.de/~imw/Papers/showpaper.php?429
ISSN: 0931-6558

Optimal Stopping Under Ambiguity In Continuous Time

Frank Riedel*
Institute of Mathematical Economics
Bielefeld University

April 8, 2010

Abstract

We develop a theory of optimal stopping problems under ambiguity in continuous time. Using results from (backward) stochastic calculus, we characterize the value function as the smallest (nonlinear) supermartingale dominating the payoff process. For Markovian models, we derive an adjusted Hamilton–Jacobi–Bellman equation involving a nonlinear drift term that stems from the agent's ambiguity aversion. We show how to use these general results for search problems and American Options.

Key words and phrases: Optimal Stopping, Ambiguity, Uncertainty Aversion, Robustness, Continuous-Time, Optimal Control

JEL subject classification: D81, C61, G11

*These results have been presented at seminars in Freiburg (December 2008), CENOS Workshop, Münster (May 2009), ZiF Bielefeld (September 2009), IDEISCOR Conference Risk Sharing and Finance, Toulouse (September 2009), the Workshop on Risk Measures and Robust Optimization in Finance, NUS Singapore, November 2009, Oxford University (February 2010). I am grateful for discussions with and comments from Freddy Delbaen, Hans Föllmer, and Xunyun Zhou. Financial Support through the German Research Foundation, International Graduate College "Stochastics and Real World Models", and Research Training Group EBIM, "Economic Behavior and Interaction Models", is gratefully acknowledged.

1 Introduction

This paper is a sequel to my previous analysis of optimal stopping problems under ambiguity in discrete time published in Riedel (2009). As such, the motivation, the economic examples, and the importance of the question is the same as in that paper, of course. We now pass from discrete to continuous time.

New economic insights can be obtained from continuous–time models. In our case, one real function – the *driver* of the variational expectation – will suffice to describe ambiguity aversion, a great simplification compared to using a whole set of multiple priors or a penalty function defined on the space of all probability measures. Continuous-time models also have special appeal due to their elegant solutions and the power of (stochastic) calculus.The explicit solutions in continuous time allow for comparative statics that would otherwise be difficult to find. The continuous–time solutions approximate well the discrete–time ones if the time intervals are sufficiently small (see our analysis in another paper, Cheng and Riedel (2010)), but they are frequently easier to interpret, and can be used, through their explicit formulas, as building stones for more complex models. Famous examples in the non–ambiguous setup include the theory of investment under uncertainty with sunk costs or the real options literature (Dixit and Pindyck (1994), Trigeorgis (1996), Smit and Trigeorgis (2004)), American options (Myneni (1992), Karatzas (1988), Jacka (1991)) or search models (Weitzman (1979), Sargent (1987), Nishimura and Ozaki (2004)). Optimal stopping has also important applications in the design of experiments (DeGroot (2004); Chernoff (1972)).

As a first step into the rich world of continuous–time models, we engage into the large class of diffusion models that still form the standard and benchmark in most of the literature. We thus assume that the relevant information is generated by a (d–dimensional) Brownian motion[1]. We aim to develop a general theory of optimal stopping for ambiguity–averse agents who use multiple prior, or more generally, variational preferences in such a diffusion setting.

Next to a number of economically desirable properties like ambiguity–aversion, continuity, monotonicity etc. a crucial feature of normatively appealing dynamic models is time–consistency. The concept of dynamic consis-

[1]In general, one might want to allow for jumps, or more general semimartingale models. First steps in this direction can be found in the recent work of Trevino (2008), e.g. who proves a general minimax theorem.

tency has a long history in economic thought, compare Koopmans (1960) and Duffie and Epstein (1992). On the technical side, time–consistency allows to apply the principle of dynamic programming. From the normative point of view — if we are to apply our results to the regulation of financial markets, for instance — dynamic consistency seems to me a necessary condition: inconsistent behavior (even if it admittedly occurs in reality!) is difficult to justify.

Time–consistency restricts the class of admissible models quite a bit, but allows nevertheless for a rich variety of economically interesting studies. To start with, there is the benchmark model of "κ–", or, as we prefer to call it, "drift" ambiguity by Chen and Epstein (2002), where the agent knows the volatility, but not the drift of the underlying Brownian motion. She thus considers all priors under which the Brownian motion has a drift $(\theta_t(\omega))$ with values in the interval $[-\kappa, \kappa]$ for some constant $\kappa > 0$ that measures the level of ambiguity. It is very important to allow for stochastic and time–varying drift $\theta_t(\omega)$ here because else, we loose time–consistency, as shown in Chen and Epstein (2002). Our examples below (Section 4) illustrate nicely that it is rational to change one's belief about the worst drift; this occurs, for example, when the payoff changes its monotonicity due to a stochastic event (as in Barrier Options).

κ–ambiguity is our leading example in the applications. But the class of models we consider here is much larger and covers (almost) all models of variational utility as axiomatized in Maccheroni, Marinacci, and Rustichini (2006). To see this, note that Chen and Epstein (2002) show that the conditional expected value under drift ambiguity solves a backward stochastic differential equation. Formally, if we denote by $\mathcal{P}^\kappa$ the set of priors under κ–ambiguity and let X be a suitably bounded random variable, then the conditional minimal expectation

$$Y_t = \mathcal{E}_t\left(X\right) = \min_{P \in \mathcal{P}^\kappa} E^P[X|\mathcal{F}_t]$$

satisfies

$$-\,dY_t = -\kappa|Z_t|dt - Z_t dW_t \tag{1}$$

for some volatility process Z. The "driver" $g(z) = -\kappa|z|$ in Equation (1) is concave. More generally, it has been shown in the literature that time–consistent variational preferences of the form

$$\mathcal{E}_t\left(X\right) = \min_P E^P[X|\mathcal{F}_t] + \alpha_t(P)$$

satisfy under some regularity conditions a backward stochastic differential equation of the form

$$-dY_t = g(t, Z_t)dt - Z_t dW_t$$

for some driver $g(t, z)$ that is concave in z. Such nonlinear expectations have been called g–expectations by Shige Peng who developed a rich probabilistic theory for them (see Peng (1997), Peng (1999)). The driver g and the penalty function α in the variational representation share the following relation. If we denote by f the convex dual[2] of g, then the penalty function can be written as

$$\alpha_t(P) = E^\theta \left[\int_t^T f(s, \theta_s) ds | \mathcal{F}_t \right]$$

where the process $\theta = (\theta_t)$ is the Girsanov kernel of the measure P with respect to our reference measure P_0, see Section 2.1 below.

Let us now review the contributions of the current paper. To start with, we derive the general structure of optimal stopping times and the value function. In the non–ambiguous case, it is well known that the value process is a supermartingale, and a martingale as long as it is not optimal to stop. By waiting, we keep the expected value constant if waiting is optimal, and we lose some value in expectation if waiting was the wrong decision. We show that our value process is the smallest rightcontinuous *g–supermartingale* that dominates the payoff process. Note that we have to replace the concept of supermartingale by the (analogously defined) concept of g–supermartingale (the g–supermartingales correspond to the multiple prior supermartingales in Riedel (2009)). It is optimal to stop when the payoff is equal to the value process.

As an aside, we provide during the proof of this theorem a lemma on rightcontinuous versions of g–supermartingales. One can assume without loss of generality that a g–supermartingale V has rightcontinuous sample paths if the ex–ante (nonlinear) expectations $t \mapsto \mathcal{E}_0(V_t)$ are rightcontinuous. This task is performed in the appendix F ; it relies on a generalization of the classical up– and downcrossing inequalities for supermartingales to g–supermartingales given in Chen and Peng (2000).

In the multiple prior case, it is natural to go for a minimax theorem as we maximize over stopping times a minimal expectation. With our general

[2] i.e. $f(\theta) = \sup_{z \in \mathbb{R}^d} g(z) - z \cdot \theta$

structure theorem, it is easy to derive such a theorem. With the help of a Girsanov transformation, we identify a worst–case measure. The worst case measure assigns drift $+\kappa$ or $-\kappa$ to the underlying Brownian motion depending on which case is more unfavorable. The sign of the drift is determined by the endogenous volatility process of the value function (we come back to this point below). For the moment, let us point out that it is indeed important to allow for stochastic drift (and stochastic Girsanov kernels) when specifying the class of priors. Without this, a worst–case measure would not exist, in general (see also our Examples below, in particular on the American Straddle Options). As soon as a worst–case measure exists, the minimax theorem follows trivially. Using convex duality, we also prove the existence of a worst–case measure in the general case.

Many important models in finance and economics exhibit a Markovian structure, the most famous one being the geometric Brownian motion model used in the Samuelson model of financial markets. We thus move on to study optimal stopping problems when the payoff is a function of an underlying Markov process that solves a (forward!) stochastic differential equation. In this case, we are able to derive an analog of the classical Hamilton–Jacobi–Bellman equation for the value function.

This partial differential equation has the same structure as in the classical case, except for a nonlinear term that is generated from ambiguity aversion. In our setup, this nonlinear term is of the from $g(t, v_x(t,x)\sigma(x))$ where g is the driver of our g–expectation and v_x is the first derivative of the value function with respect to the state variable; $\sigma(x)$ is the volatility of the state variable.

For drift ambiguity, the driver $g(x) = \kappa|x|$ is a multiple of the absolute value. With the help of our Bellman equation, we can then easily solve a large number of problems where the payoff is a monotone function of the underlying state variable. As an example, think of the American Put in the Samuelson model. There, the volatility $\sigma(x) = \sigma$ of the asset price is constant. The payoff $X_t = \max\{K - S_t, 0\}\, e^{-rt}$ is a function of (time and) the asset price which is monotone decreasing in the asset price. It is thus natural to guess that the value function will be monotone decreasing in the asset price. As we have $g(x) = -\kappa x$ for negative x, we are back at a linear Bellman equation. This linear Bellman equation corresponds to the value function of a classical optimal stopping problem (with a different drift).

We prove in general, that for such monotone problems, the value function coincides with the classical value function after changing the drift to κ (if the payoff function is monotone increasing) or $-\kappa$ (else). In particular, we thus obtain the existence of a worst–case measure in this restricted setting again.

In particular, our methods cover and give a rigorous proof of the results in Nishimura and Ozaki (2007).

In the monotone case with drift uncertainty, the worst–case prior is easy to guess: the agent just presumes the most unfavorable drift. In general, the situation is quite more complex, as we show with the help of two examples. When the payoff is path–dependent or a non–monotone function of some underlying Markov process, the agent's belief about the worst–case drift changes stochastically, and this can happen quite frequently. We illustrate this effect with the help of Barrier Options and the American Straddle. In general, the worst–case drift depends on the effect of the state variable on the value function. So, if we are currently in a region where the value function is increasing in the state variable, the agent pesumes the minimal possible drift, and vice versa. Whenever the state variable thus crosses a minimum of the value function, the agent changes her belief; in particular, close to such a minimum, this happens infinitely often due to the diffusion nature of the state variable. Under uncertainty, even rational agents might appear quite panicky from the outside, at least as far as their beliefs are concerned.

The paper is set up as follows. The next section introduces the continuous–time setup and explains the relation between variational expectations and backward stochastic differential equations, and g–expectations. All required techniques from this literature are explained in this paper. I hope to convince the reader that he or she does not need to know much more than the usual probability background to understand our results. Section 3 introduces the optimal stopping problem and exposes the general theorems about optimal stopping under ambiguity. First, it contains the general structure theorem that characterizes the value function as the smallest right-continuous g–supermartingale that dominates the payoff process. We also provide the relationship to reflected backward stochastic differential equations (but this part can be skipped). This section contains also the existence of a worst–case measure and the general minimax theorem as a corollary. Section 3.2 derives the Hamilton–Jacobi–Bellman equation in Markovian models. We conclude that section with a treatment of the infinite time horizon (for Gilboa–Schmeidler–like expectations). Section 4 contains examples and applications. We start with the monotone problems described above. To show that a reduction to a classical problem is not always trivial, we also study more complex problems. For barrier options and american straddles, the agent's worst–case measure exhibits a quite complex structure. Section 5 concludes, and the appendix contains additional material and most of the proofs.

2 The Optimal Stopping Problem for ambiguity–averse agents

Our informational setup is given by a probability space $(\Omega, \mathcal{F}, P_0)$ on which $B = \{B_t, 0 \leq t \leq T\}$ is a d-dimensional standard Brownian motion. We denote by $(\mathcal{F}_t)_{t\geq 0}$ the filtration generated by B augmented by the P_0–null sets. P_0 serves the role of a reference measure[3] here; it need not describe nor the empirical nor the subjective probabilities.

We want to study the optimal decision of an agent who receives a reward X_τ upon stopping at time τ where $X = (X_t)_{0\leq t\leq T}$ is an adapted process describing the agent's gain. Our agent is ambiguity–averse; her evaluation of an uncertain payoff X is given by a functional $\mathcal{E}_0(X_\tau)$ at time 0, respectively $\mathcal{E}_t(X_\tau)$ at time $t \geq 0$. The agent chooses a stopping time $\tau \leq T$ that maximizes $\mathcal{E}_0(X_\tau)$.

Before we impose more structure on our dynamic expectations, let us fix some technical condition on the payoff process that we shall need.

Assumption 2.1 *The payoff process $X = (X_t)_{0\leq t\leq T}$ is an adapted, nonnegative[4] process that has continuous sample paths[5]. In addition, X is bounded*

[3] In particular, the (in)equalities involving random variables have, unless otherwise stated, to be read as almost–sure–statements with respect to P_0; we also use P_0 to define our space of admissible payoffs.

[4] As our variational expectations are constant–additive, nonnegativity is equivalent to assuming boundedness from below. Most economic applications of optimal stopping have a certain "option" feature so that the payoff is bounded from below.

[5] Most interesting economic payoff processes, like Call, Put, etc. have continuous sample paths. However, this assumption can be weakened. A detailed inspection of our proofs shows that it is enough to have rightcontinuous sample paths and "leftcontinuity over stopping times in g–expectation", i.e. for any sequence of stopping times (τ_n) that increase to a stopping time τ, we have $\mathcal{E}_0(X_{\tau_n}) \to \mathcal{E}_0(X_\tau)$. This will cover examples like digital options in diffusion models.

in $L^2(\Omega, \mathcal{F}, P_0)$*:*

$$E_0 \sup_{t\in[0,T]} X_t^2 < +\infty\,. \tag{2}$$

Here, and in the sequel, we denote by E_0 the expectation with respect to our reference measure P_0.

2.1 Variational Expectations

Let us now come to the nonlinear expectation. We are inspired by our later applications to ambiguity–averse agents who evaluate payoffs according to *Gilboa–Schmeidler* preferences like

$$\mathcal{E}_0(X_\tau) = \inf_{P\in\mathcal{P}} E^P X_\tau$$

for a suitable set of priors $\mathcal{P}$, or more generally by *variational preferences* like

$$\mathcal{E}_0(X_\tau) = \inf_{P\in\Delta} E^P X_\tau + c(P)$$

for a penalty function $c(P)$ defined on the set Δ of all[6] probability measures on our space. In the recent literature on variational preferences and convex risk measures, the structure of such functionals in diffusion models has been worked out[7].

Let us start with our benchmark example, called κ–ignorance in Chen and Epstein (2002), or, as we call it, drift ambiguity. Drift ambiguity models an agent's uncertainty about the drift of the underlying Brownian motion. For simplicity, we set the dimension $d = 1$ for the moment. Fix an ambiguity parameter $\kappa > 0$ and denote by $\mathcal{D}^\kappa$ the set of all progressively measurable processes $\theta = (\theta_t)_{t\in[0,T]}$ with $|\theta_t| \le \kappa$. Call $\mathcal{P}^\kappa$ the set of all probability measures Q that are equivalent to P_0 with density $\exp\left(\int_0^T \theta_s dB_s - \frac{1}{2}\int_0^T \theta_s^2 ds\right)$. This model describes an agent who is uncertain about the drift of the underlying Brownian motion and thus allows any drift between $-\kappa$ and $+\kappa$. We stress that it is important to allow for a stochastic and time–varying Girsanov kernels (drift change) here because we otherwise lose the property of dynamic consistency, see Chen and Epstein (2002) for details.

[6] The cost function is allowed to assume the value infinity, see xxx below.

[7] The most general result in our continuous–time setting has recently been obtained by Delbaen, Peng, and Rosazza Gianin (2009).

Chen and Epstein (2002) show that drift ambiguity expectations solve a backward stochastic differential equation in the following sense. At the terminal time, we have $\mathcal{E}_T(X) = X$, as X is supposed to be known at time T (the terminal condition); before T, the expectations satisfy the recursive relation

$$d\mathcal{E}_t(X) = -\kappa|Z_t|\,dt + Z_t\,dB_t \tag{3}$$

for some progressively measurable, square–integrable process Z. Therefore, ambiguity gives rise to a version of stochastic differential utility in the sense of Duffie and Epstein (1992)[8]. The ambiguity aversion is described here by the concave function $g(z) = -\kappa|z|$.

A more general version of stochastic differential utility allows for general concave functions g of the volatility process Z in the backward recursive equation (3). This in turn leads to a suitable version of dynamic variational expectations in our diffusion setting.

Definition 2.2 *We call a function* $g : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$ *a* standard driver for variational expectations *if it satisfies the following properties:*

1. $(g(\omega,t,z))_{t\in[0,T]}$ *is an adapted process with*

$$E\int_0^T |g(t,z)|^2\,dt < \infty$$

for all $z \in \mathbb{R}^d$*;*

2. *g is Lipschitz continuous*[9] *in z, uniformly in t and ω: there exists $\mu > 0$ such that for all $t \geq 0$ and all $z_1, z_2 \in \mathbb{R}^d$ we have* $|g(\omega,t,z_1) - g(\omega,t,z_2)| \leq \mu\|z_1 - z_2\|$*;*

3. $g(\omega,t,0) = 0$ *for all $t \geq 0$ and $\omega \in \Omega$,*

4. *g is concave in z.*

[8]For more on this relationship, see, e.g., Skiadas (2003). Let us also mention that the Bernoulli utility function u is integrated into the payoff process in our paper.

[9]By assuming Lipschitz continuity for the aggregator, we exclude some dynamic risk measures. A prominent example is the so–called *entropic risk measure*, or in our context, better: entropic expectation, which has driver $g(z) = \frac{1}{2}z^2$. However, for this expectation, the optimal stopping problem under ambiguity is equivalent to the *standard* optimal stopping problem to maximize $-\frac{1}{a}E_0 \exp(-aX_\tau)$ for some parameter $a > 0$. Entropic risk corresponds in economic terms to Expected Utility with Constant Absolute Risk Aversion.

The backward stochastic differential equation

$$dY_t = -g(Z_t)dt + Z_t dB_t\,. \tag{4}$$

with terminal condition $X \in L^2$ has then a unique solution (Y, Z) in the space of square–integrable adapted processes with terminal condition $Y_T = X$. Peng (1997) calls $Y_t = \mathcal{E}_t(X)$ the (conditional) g–expectation of X.

Every g–expectation gives rise to a variational expectation in the following sense. Let

$$f(\omega, t, \theta) = \sup_{z \in \mathbb{R}^d} g(\omega, t, z) - z \cdot \theta \qquad \left(\omega \in \Omega, t \in [0, T], \theta \in \mathbb{R}^d\right)$$

be the convex dual of g[10]. We denote by D the set of all progressively measurable processes (θ_t) such that

$$E_0 \int_0^T f(s, \theta_s)^2\, ds < \infty\,.$$

Now let $t \leq \tau \leq T$ be a stopping time. El Karoui, Peng, and Quenez (1997) establish the following dual representation for solutions of backward stochastic differential equations (or g–expectations): For an $\mathcal{F}_\tau$–measurable random variable ξ the g-expectation at time t can be written as

$$\mathcal{E}_t(\xi) = \operatorname*{ess\,inf}_{\theta \in D} E^\theta\left[\xi | \mathcal{F}_t\right] + \alpha_{t,\tau}(\theta)$$

where the penalty function is

$$\alpha_{t,\tau}(\theta) = E^\theta\left[\int_t^\tau f(s, \theta_s) ds | \mathcal{F}_t\right]\,. \tag{5}$$

Here, E^θ denotes the expectation under the measure P^θ determined by the Girsanov transformation with kernel θ. Note that we consider only measures P here that are equivalent to P_0; we identify those measures with their Girsanov kernel θ. Moreover, we restrict to Girsanov kernels with values in the domain of the convex dual f; as g is Lipschitz–continuous with constant μ, we know that θ takes values in the compact set $[-\mu, \mu]^d$. We refer to Delbaen, Peng, and Rosazza Gianin (2009) for more details.

[10] Here, and in the sequel, $z \cdot \theta$ denotes the dot or scalar product in $\mathbb{R}^d$.

2.2 Properties of g–Expectations, g–Martingales

Let us list here for later reference a set of properties of variational expectations in the above sense. We shall see that they share many of the appealing properties with standard expectations, except, of course, the linearity. $\mathcal{E}_t(X)$ is defined for all square–integrable, $\mathcal{F}_T$–measurable random variables in $L^2 = L^2(\Omega, \mathcal{F}_T, P_0)$ and is itself a square–integrable, $\mathcal{F}_t$–measurable random variable. $\mathcal{E}_t(X)$ is the agent's subjective expected value for X given her information at time t.

The variational expectation is monotone, i.e.

$$X \geq Y \Rightarrow \mathcal{E}_t(X) \geq \mathcal{E}_t(Y) \tag{6}$$

and the inequality is strict whenever $X < Y$ with positive P_0–probability.

Our variational expectation is also additive with respect to payments that are known at time t: if Z is $\mathcal{F}_t$–measurable, then we have for all $X \in L^2$

$$\mathcal{E}_t(X + Z) = \mathcal{E}_t(X) + Z. \tag{7}$$

If the agent knows that an event A has occurred at time t, then her expectation does not depend on states of the world that he knows to be impossible from now on. Technically, we have for $A \in \mathcal{F}_t$ and all $X \in L^2$

$$\mathcal{E}_t(1_A X) = 1_A \mathcal{E}_t(X)\ . \tag{8}$$

This condition is equivalent to the following seemingly stronger property: for an event $A \in \mathcal{F}_t$ and all $X, Y \in L^2$, we have[11]

$$\mathcal{E}_t(1_A X + 1_{A^c} Y) = 1_A \mathcal{E}_t(X) + 1_{A^c} \mathcal{E}_t(Y)\ . \tag{9}$$

The next property in our list is a suitable kind of continuity. We will frequently use monotonic continuity,

$$X_n \uparrow X \Rightarrow \mathcal{E}_t(X_n) \uparrow \mathcal{E}_t(X) \tag{10}$$

and a version of Lebesgue's dominated convergence theorem: if (X_n) is bounded by some square–integrable random variable $Y \in L^2$ and $X_n \to X$, then

$$\mathcal{E}_t(X_n) \to \mathcal{E}_t(X)\ . \tag{11}$$

[11] The proof is as follows (can be canceled later): let $Z = 1_A X + 1_{A^c} Y$. By applying (8) twice, we have $1_A \mathcal{E}_t(X) = \mathcal{E}_t(1_A X) = \mathcal{E}_t(1_A Z) = 1_A \mathcal{E}_t(Z)$. By analogy, we have for $B = A^c$, $1_B \mathcal{E}_t(Y) = \mathcal{E}_t(1_B Y) = \mathcal{E}_t(1_B Z) = 1_B \mathcal{E}_t(Z)$. By adding up the two equations, we get (9).

So far, we have discussed basic properties of expectations. We come now to inherently dynamic considerations. For a Bayesian agent, we have the important law of iterated expectations; this law is essential in deriving time–consistency and the principle of dynamic programming. It is of utmost importance for our dynamic techniques that we have this law here, too. For all times $0 \leq s \leq t \leq T$ and random variables $X, Y \in L^2$

$$\mathcal{E}_s(X) = \mathcal{E}_s(\mathcal{E}_t(X)) \tag{12}$$

holds true.

We note some basic regularity in the time variable t; the expectation process $Y_t = \mathcal{E}_t(X)$ is continuous in t as it solves a backward stochastic differential equation. For a stopping time $\tau \leq T$, the "expectation at time τ" is then a well–defined $\mathcal{F}_\tau$–measurable random variable $\mathcal{E}_\tau(X)$ and it coincides with the process Y stopped at τ.

We model now (un)fair games against nature, or g–(super)martingales for our agent. A process $(S_t)_{t\in[0,T]} \subset L^2$ is called a g–(super–)martingale if we have for all $0 \leq u \leq t \leq T$

$$\mathcal{E}_u(S_t) = S_u \quad (\mathcal{E}_u(S_t) \leq S_u)\ . \tag{13}$$

Rational agents understand that there is no smart way to beat an unfair game. Technically, we have a version of the *Optional Sampling Theorem*: let $(S_t)_{t\in[0,T]}$ be a g–supermartingale. Then for all stopping times $0 \leq \sigma \leq \tau \leq T$ we have

$$\mathcal{E}_\sigma(S_\tau) \leq X_\sigma\ . \tag{14}$$

The fact that our variational expectation has all these nice properties has been established by Shige Peng, 1997, see also Coquet, Hu, Mémin, and Peng (2002).

Theorem 2.3 (Peng) *The g–expectation $(\mathcal{E}_t(X))$ satisfies the conditions* (6) *to* (14).

3 The Structure of Optimal Stopping Times

We now clarify the general structure of our optimal stopping problem. As our variational expectations are time–consistent, the dynamic programming principle holds true even though we are in a non–Bayesian world. Loosely speaking, it says that our value function $V = (V_t)_{0 \le t \le T}$ satisfies

$$V_t = \max \{X_t, \mathcal{E}_t (V_{t+dt})\} \ .$$

In other words, either it is optimal to stop at time t (because of $V_t = X_t$), or the value function stays constant in expectation; in probabilistic terms, the value function is a g–supermartingale ($V_t \le \mathcal{E}_t (V_{t+dt})$), and a g–martingale as long as it is not optimal to stop. Generalizing the classical theorem about optimal stopping problems (see, e.g. El Karoui (1979), Peskir and Shiryaev (2006)), we now characterize the value function as the smallest (rightcontinuous, see below) g–supermartingale that dominates the payoff process. At the same time, we extend the discrete–time result of Riedel (2009) to continuous time.

Theorem 3.1 *Let*

$$V_t = \operatorname*{ess\,sup}_{\tau \ge t} \mathcal{E}_t (X_\tau)$$

be the value function of the optimal stopping problem. One can choose a version of V *with rightcontinuous sample paths. Moreover:*

1. (V_t) *is the smallest rightcontinuous* g*–supermartingale dominating* (X_t)*;*
2. $\tau^* = \inf \{t \ge 0 : V_t = X_t\}$ *is an optimal stopping time;*
3. *the value function stopped at* τ^*, $(V_{t \wedge \tau^*})$ *is a* g*–martingale.*

We quickly comment on an important technical detail for the proof. From its definition, it is not clear that the value function has rightcontinuous sample paths; and in fact, this need not be true per se. It is thus important to show that we can find a *version* of that process with rightcontinuous sample paths. This task is performed in the appendix F ; it relies on a generalization of the classical up– and downcrossing inequalities for supermartingales to g–supermartingales given in Chen and Peng (2000)[12]

[12]The question of rightcontinuity has also been studied in a much more general framework in independent work of Trevino (2008).

A Detour: Reflected Backward Stochastic Differential Equations

At this point, we would like to highlight the relation of our approach with the mathematical literature on reflected backward stochastic differential equations and obstacle problems. Let $X = (X_t)_{0 \le t \le T}$ be an adapted stochastic process that we interpret as an obstacle. Our aim is to control a process $V = (V_t)_{0 \le t \le T}$ in such a way that

- it stays above the obstacle,
- satisfies the stochastic dynamics

$$- dV_t = g(t, Z_t)dt - Z_t \cdot dB_t + dC_t \tag{15}$$

  for an increasing process C,
- the terminal condition $V_T = X_T$ holds true,

and we spend as less "fuel" dC_t as possible. Formally, we call a triple (Y, Z, C) of progressively measurable processes with $E_0 \sup_{0 \le t \le T} Y_t^2 < \infty$, $E_0 \int_0^T \|Z_t\|^2 dt < \infty$ and C a rightcontinuous, increasing process with $C_0 = 0$ and $E_0 C_T^2 < \infty$ a solution to the reflected backward stochastic differential equation with obstacle X, terminal condition X_T and driver g if the above conditions are satisfied, and in addition, we have

$$\int_0^T (Y_t - X_t)\, dC_t = 0\,,$$

see ElKaroui, Kapoudjian, Pardoux, Peng, and Quenez (1997) for details. These authors use a Picard–Lindelöf–type approach to prove existence and uniqueness of solutions.

Our approach yields another proof for the existence and uniqueness of solutions to reflected backward stochastic differential equations. The reason is that the value function of our optimal stopping problem with variational expectations with driver g is the unique solution of the reflected BSDE. The obstacle problem can be viewed as a sequence of optimal stopping problems. Our value function clearly stays above the "obstacle" X; times when it is necessary to push up the process V to stay above X correspond to optimal stopping times in our problem.

Corollary 3.2 *There exists an increasing, adapted, rightcontinuous process with left limits K and an adapted process Z with $E_0 \int_0^T \|Z_t\|^2 dt < \infty$ such that (V, Z, K) is the unique solution of the reflected backward stochastic differential equation with driver g, obstacle X, and terminal condition X_T.*

Intuitively, it is clear that the value process dominates the payoff process. The g–martingale principle entails that we have the stochastic dynamics (15) with $dC_t = 0$ in the continuation region. The value function is pushed upwards whenever it is optimal to stop, i.e. when $V_t = X_t$, and only then. This yields the minimality required for a reflected backward stochastic differential equation.

3.1 Worst–Case Priors and Duality

We now relate the solution of our optimal stopping problem under ambiguity to classical single–prior solutions. We identify a *worst–case prior* P^* such that the value function of our stopping problem coincides with the value function of stopping the payoff process X under the worst–case prior P^* in case of drift ambiguity. In general, we have to take the penalty function of the variational expectation into account. Nevertheless, we still find a Girsanov kernel θ^* that achieves the infimum in the dual representation of the g–expectation.

To illustrate our general result we start with the Gilboa–Schmeidler–type example of drift ambiguity, or κ–ignorance. When the payoff is a monotone increasing function of the underlying Brownian motion, it is plausible that the maximal negative drift is the worst case to consider by the agent. In general, the worst drift depends on the local monotonicity of the payoff; in a general, non–Markovian setting, this local term is taken over by the volatility process Z that comes from the Doob–Meyer decomposition of the value process.

The Doob–Meyer–type decomposition for rightcontinuous g–supermartingales was derived by Peng (1999), see Lemma A.1 in the Appendix. Classically, any rightcontinuous supermartingale[13] can be decomposed into the difference of a martingale and a predictable, increasing process. For g–supermartingales, there is a similar decomposition of the form

$$-dV_t = g(t, Z_t)dt + dA_t - Z_t \cdot dB_t$$

[13] of class (D), an integrability condition that is satisfied in our paper

for a suitable volatility Z and an increasing process A. In the case of κ–ignorance, $g(t, Z_t) = -\kappa|Z_t|$ and we thus have

$$-dV_t = -\kappa|Z_t|dt + dA_t - Z_t \cdot dB_t \,.$$

With the help of a Girsanov transformation, we can then find a probability measure P^* and a Brownian motion B^* such that

$$dV_t = -dA_t + Z_t \cdot dB_t^* \,.$$

(B^* has the drift $\kappa \operatorname{sgn}(Z_t)$ under P^*). Our value process is thus a classical supermartingale under P^*. Therefore, it must be greater or equal to the value process of the classical stopping problem under P^*. On the other hand, as $P^* \in \mathcal{P}$, we also have the other inequality. We thus obtain the following minimax theorem.

Theorem 3.3 (Duality for drift ambiguity) *Suppose that the agent uses κ–ignorance to evaluate her expectations, i.e.*

$$\mathcal{E}_t\left(X\right) = \operatorname*{ess\,inf}_{P \in \mathcal{P}^\kappa} E^P X$$

for some parameter $\kappa > 0$. There exists a probability measure $P^* \in \mathcal{P}^\kappa$ such that

$$V_t = \operatorname*{ess\,sup}_{\tau \geq t} \mathcal{E}_t\left(X_\tau\right) = \operatorname*{ess\,sup}_{\tau \geq t} E^*\left[X_\tau | \mathcal{F}_t\right] \,.$$

In particular, optimal stopping times and worst–case measures exist, and we have the minimax relation

$$\max_\tau \min_{P \in \mathcal{P}^\kappa} E^P X_\tau = \min_{P \in \mathcal{P}^\kappa} \max_\tau E^P X_\tau$$

The existence of a worst–case measure carries over to the general case. Let

$$f(t, \theta) = \sup_{z \in \mathbb{R}^d} g(t, z) - z \cdot \theta$$

be the convex dual of g. Recall that the penalty function can be represented as

$$\alpha_t(\theta) = E\left[\int_t^T f(s, \theta_s) ds | \mathcal{F}_t\right]$$

by (5). By convex duality, we then have

$$g(t,z) = \inf_{\theta \in D_t} f(t,z) + z \cdot \theta$$

where we denote by $D_t = \left\{\theta \in \mathbb{R}^d : f(t,\theta) < \infty\right\}$ the effective domain of f. The infimum is actually achieved because the the effective domain is included in the compact set $[-\mu,\mu]^d$ (μ–Lipschitz–continuity of g). With the help of a measurable selection theorem, we can then choose an adapted process θ^* with values in the effective domain of f that solves

$$f(t,\theta^*_t) = g(t,Z_t) - Z_t\theta^*_t\,.$$

The process θ^* is the Girsanov kernel for the worst case measure. We have the following theorem:

Theorem 3.4 (General Duality) *For variational expectations with standard aggregator g there exists a worst–case measure P with Girsanov kernel θ^* that solves*

$$f(t,\theta^*_t) = g(t,Z_t) - Z_t\theta^*_t\,.$$

We have

$$V_t = \operatorname*{ess\,sup}_{\tau \geq t} E^{\theta}\left[X_\tau + \int_t^\tau f(s,\theta_s)ds|\mathcal{F}_t\right]\,.$$

Proof: Let θ^* solve

$$f(t,\theta^*_t) = g(t,Z_t) - Z_t\theta^*_t$$

as described above. Set $V^*_t = \operatorname{ess\,sup}_{\tau\geq t} E^{\theta^*}\left[X_\tau + \int_t^\tau f\left(s,\theta^*_s\right)ds|\mathcal{F}_t\right]$. From the dual representation of variational expectations (see (5)), we have

$$V_t = \operatorname*{ess\,sup}_{\tau\geq t} \mathcal{E}_t\left(X_\tau\right) = \operatorname*{ess\,sup}_{\tau\geq t}\ \operatorname*{ess\,inf}_{\nu:\,\nu_t\in D_t} E^{\nu}\left[X_\tau + \int_t^\tau f(s,\nu_s)ds|\mathcal{F}_t\right]$$

and we get immediately $V \leq V^*$.

On the other hand, the Doob–Meyer–Peng decomposition for g–supermartingales (see our Lemma A.1 in the Appendix) yields an increasing rightcontinuous adapted process A and a square–integrable adapted process Z with

$$-dV_t = g(t,Z_t)dt + dA_t - Z_t \cdot dB_t$$

which is – by definition of θ^* – equal to

$$= \left(f\left(t, \theta_t^*\right) + Z_t \cdot \theta_t^*\right) dt + dA_t - Z_t \cdot dB_t$$
$$= f\left(t, \theta_t^*\right) dt + dA_t - Z_t \cdot dB_t^* \, .$$

Here, B^* denotes the standard Brownian motion under P^{θ^*} with $B_t^* = B_t - \int_0^t \theta_s^* ds$. Hence, $M_t = V_t + \int_0^t f\left(s, \theta_s^*\right) ds$ is a P^{θ^*}–supermartingale that dominates X (f is nonnegative because of $g(t,0) = 0$). From the classical Snell envelope theorem, we conclude that $M_t \geq V_t^* + \int_0^t f\left(s, \theta_s^*\right) ds$, or $V \geq V^*$. □

3.2 The Markov Case: Hamilton–Jacobi–Bellman Equation

In many applications, the state of the economic system is described by a diffusion S with values in $\mathbb{R}^d$ that evolves according to the (forward) stochastic differential equation

$$dS_t = b(S_t)dt + \sigma(S_t)dB_t, S_0 = x \tag{16}$$

for suitable drift

$$b : \mathbb{R}^d \to \mathbb{R}^d$$

and volatility functions

$$\sigma : \mathbb{R}^d \to \mathbb{R}^{d \times d} \, .$$

We assume that b and σ satisfy the usual regularity assumptions that ensure a unique strong solution for the stochastic differential equation (16). The standard model is still the Samuelson specification with $b(x) = \mu x$ and $\sigma(x) = \sigma x$ for some constants μ and σ and $d = 1$. We obtain the Bachelier model for $b(x) = \mu$ and $\sigma(x) = \sigma$. As in Krylov (1980), let us assume

Assumption 3.5 *The functions σ and b are Lipschitz–continuous, i.e. there exists a constant $K > 0$ with*

$$\|\sigma(x) - \sigma(y)\| + \|b(x) - b(y)\| \leq K\|x - y\| \qquad (x, y, \in \mathbb{R}^d) \, .$$

σ grows at most linearly, i.e. $\|\sigma(x)\| \leq K(1+\|x\|), x \in \mathbb{R}^d$ and it is uniformly elliptic in the sense that for some $\eta > 0$

$$x \cdot ax \geq \eta \|x\|^2 \qquad (x \in \mathbb{R}^d)$$

for the diffusion matrix a with $a_{ij}(x) = \sum_{k=1}^d \sigma_{ik}(x)\sigma_{jk}(x)$.

The payoff process $V_t = f(t, S_t)$ is often a function of this strong Markov process and time. The typical examples are, of course, the American Put and Call with $f(t,x) = e^{-rt}(K-x)^+$ and $e^{-rt}(x-K)^+$ for some interest rate $r > 0$ and a strike $K > 0$. In addition to these cases, we also treat the American Straddle ($f(t,x) = e^{-rt}|K-x|$) below.

In such a situation, the value function $v(t,x)$ of the optimal stopping problem solves a Hamilton–Jacobi–Bellman equation of the type

$$\max_{(t,x)} \{f(t,x) - v(t,x), v_t(t,x) + \mathcal{L}v(t,x)\} = 0$$

where

$$\mathcal{L} = \sum_{i=1}^{d} b_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij(x)}\frac{\partial^2}{\partial x_i \partial x_j}$$

is the infinitesimal generator of the Markov process S.

Before we discuss the ambiguous situation, let us pause to reflect on the smoothness of the value function. In general, we cannot expect the value function to be twice continuously differentiable in the space variable. For example, the value function of the American Put is not twice, but only once continuously differentiable at the exercise boundary (see, e.g., Jacka (1991), Peskir and Shiryaev (2006), usw.)[14]. On the other hand, under our assumptions, the value function is smooth enough to allow for generalized derivatives; moreover, the value function is also smooth enough for an application of Itô's formula. We refer to Krylov (1980), Chapter 4.7, Theorems 4 and 7 for the existence of the generalized derivatives. The same book contains in Chapter 2.10 a discussion of the generalized version of Itô's formula[15]. For our purposes, we will thus work with the space $\mathbb{W}^{1,2}$ that consists of continuous functions that can be suitably well approximated by functions that are continuously differentiable in time and twice continuously differentiable in the space variable[16].

Let us now come to ambiguity's effect on the Bellman equation in the Markovian case. We have already seen that ambiguity is essentially described

[14]Barrier Options of the Knock–Out type even lose the differentiability at the knock–out barrier, see, e.g., Karatzas and Wang (2000). Such options have path–dependent payoffs, however.

[15]In fact, one can even find weaker assumptions, see Föllmer, Protter, and Shiryaev (1995) for the best possible theorem in this direction.

[16]See, e.g., Krylov (1980) for the definition of this space and its use in optimal control.

by the drift term $g(t, Z_t)$ above. The volatility Z can be identified via Itô's lemma as the first derivative of the value function (or gradient, in general) with respect to the state variable multiplied by the volatility (matrix) . We thus obtain the following extension of the HJB–equation to ambiguity.

Theorem 3.6 (Verification Theorem) *Let $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be a continuous function that satisfies a growth condition: there exists $m \in \mathbb{N}$ and a constant $K > 0$ such that*

$$|f(t,x)| \leq K\left(1 + \|x\|\right)^m . \tag{17}$$

Let $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be an element of $\mathbb{W}^{1,2}$ that solves the nonlinear Hamilton–Jacobi–Bellman equation

$$\max_{(t,x)\in[0,T]\times\mathbb{R}^d} \left\{ f(t,x) - v(t,x), v_t(t,x) + \mathcal{L}v(t,x) + g(t, \sigma(x)^\top \nabla v(t,x)) \right\} = 0\,, \tag{18}$$

and the terminal condition

$$v(T,x) = f(T,x) \qquad \left(x \in \mathbb{R}^d\right) .$$

Assume also that v and its generalized first derivatives satisfy a growth condition: for some $n \in \mathbb{N}$ and $K > 0$ we have

$$|v(t,x)| + \|\nabla v(t,x)\| \leq K\left(1 + \|x\|\right)^n . \tag{19}$$

Then v is the value function of our optimal stopping problem, i.e. $V_t = v(t, X_t)$.

The most general way to work with nonsmooth solutions of the Bellman equation, and, maybe, the most elegant one, is to use the concept of viscosity solutions. In quite a large class of Markovian optimal stopping problems, the value function can be shown to be the unique viscosity solution of the associated Hamiton–Jacobi–Bellman equation. See Crandall, Ishii, and Lions (1992) and Fleming and Soner (2006). For our setup, ElKaroui, Kapoudjian, Pardoux, Peng, and Quenez (1997) have shown in the context of reflected backward stochastic differential equations that the value function of our optimal stopping problem is even the unique viscosity solution of the above Bellman equation. For applications, it is often easier to check piecewise differentiability (which is enough to apply our above result). This works in a number of concrete case studies (Pham (1997), Peskir and Shiryaev (2006), Karatzas (1988), Jacka and Lynn (1992)); this is the reason why we have opted for this less general formulation that is hopefully easier to apply.

Ambiguity introduces nonlinearity into the HJB equation. From a theoretical point of view, the nonlinearity is not too critical as it involves only the first–order term, not the leading term of the partial differential equation.

We see again here that in our diffusion setting, we are essentially dealing with ambiguity about the drift. The term $g\left(t, \sigma(x)^\top \nabla v(t,x)\right)$ in the Hamilton–Jacobi–Bellman equation reflects this ambiguity.

The Bellman equation opens a way to obtain the worst case measure in this Markovian setting. If we remove the drift term $g\left(t, \sigma(x)^\top \nabla v(t,x)\right)$ by a Girsanov transformation, we are back at a classical Hamilton–Jacobi–Bellman equation under the new measure; the value function thus coincides with the value function of the classical optimal stopping problem under the new measure, and we get the above duality result Theorem 3.4 in this setting.

3.3 Infinite Time Horizon

Our results can be extended to infinite time horizon under suitable integrability assumptions on the payoff function. We illustrate how to perform this for the benchmark case of drift ambiguity. In this subsection, let

$$\mathcal{E}_t(X) = \operatorname*{ess\,inf}_{P \in \mathcal{P}^\kappa} E^P\left[X|\mathcal{F}_t\right]$$

be the conditional expectation for an agent who faces drift ambiguity. This expression is well–defined for any P_0–square integrable random variable X on our probability space $(\Omega, \mathcal{F}, P_0)$. Now let $(X_t)_{t\geq 0}$ be a payoff process that satisfies Assumption 2.1 for $T = \infty$.

For stopping times τ that are *universally almost surely finite*, i.e. $P[\tau < \infty] = 1$ for all $P \in \mathcal{P}^\kappa$, we can then formulate the infinite time horizon optimal stopping problem

$$V_t = \operatorname*{ess\,sup}_{\tau \geq t} \mathcal{E}_t(X_\tau) = \operatorname*{ess\,sup}_{\infty > \tau \geq t} \operatorname*{ess\,inf}_{P \in \mathcal{P}^\kappa} E^P\left[X_\tau|\mathcal{F}_t\right].$$

Let us denote by V^T the value function for the corresponding problem with a finite horizon $T < \infty$. It is clear that $V^T \leq V$ and that the value function is increasing in T, as you expand your options with a longer time horizon. Hence, the limit

$$V^\infty = \lim_{T \to \infty} V^T$$

is well defined and $V^\infty \leq V$. In the next proof, we show that V^∞ is actually equal to V.

Theorem 3.7 *The value function for the infinite time horizon is the limit of the finite horizon value functions:*

$$\lim_{T\to\infty} V^T = V\,.$$

4 Examples and Applications

In this section, we assume that the agent uses κ–ambiguity to evaluate his payoff. In particular, the nonlinear function describing his ambiguity aversion is $g(z) = -\kappa|z|$ for some $\kappa > 0$. We also go to dimension $d = 1$ and keep the Assumption 3.5 for the Markov process S with dynamics (16). We also assume a positive volatility $\sigma(x) > 0$ throughout.

4.1 Monotone Markov Problems under κ–ambiguity

In many applications, the payoff function $f(t,x)$ is monotone in the state variable. Typical examples are the American Call with $f(t,x) = e^{-rt}(x - K)^+$ for some interest rate $r > 0$ and a strike $K > 0$, or the American Put with $f(t,x) = e^{-rt}(K - x)^+$. The same payoff structure holds true in irreversible investment problems and some search problems as studied recently by Nishimura and Ozaki (2004) and Nishimura and Ozaki (2007).

In such cases, it is clear from an economic point of view that the worst thing that can happen to the agent is the lowest possible drift. Technically, under any measure P^θ with Girsanov kernel $|\theta_t(\omega)| \leq \kappa$, the dynamics of the underlying S read as

$$dS_t = (b(S_t) + \theta_t \sigma(S_t))\,dt + \sigma(S_t)dB_t\,.$$

As $\sigma(S_t) > 0$, it is suggestive that the choice $\theta^*_t(\omega) = -\kappa$ should be the worst case.

How can we verify this? Under $P^{-\kappa}$, we have a standard Markovian optimal stopping problem of a diffusion process. Classical results (Krylov (1980)) then tell us that the value function under $P^{-\kappa}$, that we denote by $v^{-\kappa}(t,x)$ satisfies the classical Bellman equation

$$\max_{(t,x)} \{f(t,x) - v(t,x), v_t(t,x) + \mathcal{L}v(t,x) - \kappa v_x(t,x)\sigma(x)\} = 0\,.$$

Note the additional term $-\kappa v_x(t,x)\sigma(x)$ that comes from changing the drift.

But the monotonicity of f entails the monotonicity of $v^{-\kappa}$; hence, we have $v_x^{-\kappa}(t,x) \geq 0$. The nonlinearity in our Bellman equation (18) then vanishes because of $|v_x^{-\kappa}| = v_x^{-\kappa}$. The "old" Bellman equation under $P^{-\kappa}$ thus coincides with our Bellman equation under g–expectations. We can apply our verification theorem[17] 3.6 to show that $v^{-\kappa}$ is also the value function of our problem.

Theorem 4.1 *Assume that the payoff function f is continuous, increasing in x, and satisfies the growth condition* (17)*. Denote by $v^{-\kappa}(t,x)$ the value function of the standard optimal stopping problem with payoff function f under the measure $P^{-\kappa}$. Assume that $v^{-\kappa}$ satisfies the growth condition* (19)*. Then the optimal stopping time under κ–ambiguity has value function $V_t = v^{-\kappa}(t,S_t)$ where $v^{-\kappa}$ is the value function of the classical optimal stopping problem under the measure $P^{-\kappa}$ with the least favorable drift for the underlying process S.*

An analogous result holds true if f is decreasing in the state variable, of course.

Irreversible Investment Nishimura and Ozaki (2007) prove the important economic result that Knightian uncertainty has opposite comparative statics than risk in investment timing problems with sunk cost. In their problem, a firm receives a (discounted) profit π_t, modeled as a geometric Brownian motion, and pays a irretrievable cost $K > 0$ upon entering a market. The authors derive the Bellman equation for the finite time horizon and assume "relations among variables in the finite-horizon case converge, as the horizon goes to infinity, to those in the infinite-horizon case", see their Proposition 2. They then identify $P^{-\kappa}$ as the worst–case measure for the infinite time horizon and obtain explicit results and comparative statics.

Our above theorem[18] 4.1 shows that $P^{-\kappa}$ is also the worst–case measure in the finite time horizon model. Moreover, our results on convergence (The-

[17]The growth conditions of Theorem 3.6 are also satisfied, see Krylov (1980), Theorem 4.7.4. and 4.7.7.

[18]Let us quickly point out that the growth condition (19) is satisfied for both the American Put and the Call Option if the underlying process is a geometric Brownian motion, i.e. $\mu(x) = \mu x$ and $\sigma(x) = \sigma x$ for constants $\mu \in \mathbb{R}, \sigma \neq 0$. For the Put, with payoff function $e^{-rt}(K-x)^+$, this is clear as both the value function and the derivative in x are bounded. For the American Call ($f(t,x) = e^{-rt}(x-K)^+$) or the problem of optimal investment as

orem 3.7) establish that the assumption in Nishimura/Ozaki's proposition 2 is satisfied. This extends their analysis to the finite–time horizon case and provides a complete proof of their main result.

4.2 Non–Monotone and Path–Dependent Payoffs

4.2.1 Barrier Options

We treat now a typical barrier option under κ–ambiguity by looking at the American version of the so-called up-and-in Put. Such an option is *knocked in* when the asset hits a level H; if this does not happen over the lifetime of the contract, the option remains worthless. After knock-in, the buyer owns a usual American Put with some strike K. We assume $H > K$ and, to avoid the trivial case, $H > S_0 = 1$.

We take the standard Samuelson–model of asset pricing under P_0 by setting

$$dS_t = \mu S_t dt + \sigma S_t dW_t, S_0 = 1\,.$$

The option's payoff is

$$X_t = e^{-rt} 1_{\{t \geq \tau_H\}} (K - S_t)^+$$

for a discount rate $r > 0$, a strike $K > 0$, and a knock–in time

$$\tau_H = \inf \{t \geq 0 : S_t \geq H\}\ .$$

After knock–in at τ_H, the barrier option is nothing but a standard American put. It is thus clear from our previous results on monotone payoffs, that the worst–case measure has drift $+\kappa$. The value at time τ_H of the American barrier option is thus

$$Y_{\tau_H} = \operatorname*{ess\,sup}_{\tau \geq \tau_H} E^{\kappa} \left[e^{-r\tau} (K - S_\tau)^+ | \mathcal{F}_{\tau_H} \right]\ .$$

treated in the real options literature (see Dixit and Pindyck (1994) or Trigeorgis (1996)), we have to assume that $\mu < r$, or else stopping until the end is optimal (and the growth conditions are satisfied anyway). The value function of the infinite time horizon is polynomial, and an upper bound for the finite horizon problem. Hence, the value function satisfies the growth condition. As far as the derivative in x is concerned, we have the following. For large values of x, immediate stopping is optimal and we have $v_x(t, x) = 1$. In the continuation region, the derivative v_x is increasing in x because the value function $v(t, x)$ is convex in x. By smooth fit, we have $v_x = 1$ at the stopping boundary. So the derivative v_x is bounded.

As S is a strong Markov process under P^κ, this value depends only on $S_{\tau_H} = H$ and the time of knock–in τ_H. Let us denote by $v^\kappa(\tau_H, H)$ the subjective value of the put under the measure P^κ, see Jacka (1991) for details.

Let $\tau^\star$ denote an optimal exercise time for the American up–and–in put. As the payoff is zero before knock–in, we clearly have $\tau^\star \geq \tau_H$. Finally, by Theorem 3.1 and optional sampling for g–martingales we obtain on the event $\{\tau_H > t\} \in \mathcal{F}_t$

$$\begin{aligned} V_t &= V_{t\wedge\tau^\star} = \operatorname*{ess\,inf}_{P\in\mathcal{P}^\kappa} E^P\left[V_{\tau_H\wedge\tau^\star}|\mathcal{F}_{t\wedge\tau^\star}\right] = \operatorname*{ess\,inf}_{P\in\mathcal{P}^\kappa} E^P\left[V_{\tau_H}|\mathcal{F}_t\right] \\ &= \operatorname*{ess\,inf}_{P\in\mathcal{P}^\kappa} E^P\left[v^\kappa(\tau_H, H)\, e^{-r(\tau_H - t)}|\mathcal{F}_t\right] . \end{aligned} \tag{20}$$

We are thus left with determining the minimal expectation for a function of the knock–in time τ_H. The American put value $v^\kappa(\tau_H, H)$ is decreasing in τ_H, as is $e^{-r(\tau_H - t)}$. We will now use stochastic dominance to show that $P^{-\kappa}$ is the worst–case measure. This is quite intuitive as the asset price has the lowest drift under $P^{-\kappa}$.

Lemma 4.2 *τ^H is stochastically largest under $P^{-\kappa}$ in the set of priors $\mathcal{P}^\kappa$ in the sense that on $\{\tau_H > t\}$ we have for all $t < u \leq T$, and for all $P \in \mathcal{P}^\kappa$*

$$P^{-\kappa}\left[\tau_H > u|\mathcal{F}_t\right] \geq P\left[\tau_H > u|\mathcal{F}_t\right] .$$

Therefore, Lemma 4.2 and the usual characterization of first–order stochastic dominance yield for any $\theta \in \mathcal{D}^\kappa$ on $\{\tau_H > t\}$

$$V_t = E^{P^{-\kappa}}\left[v^\kappa(\tau_H, H)\, e^{-r(\tau_H - t)}|\mathcal{F}_t\right] \leq E^{P^\theta}\left[v^\kappa(\tau_H, H)\, e^{-r(\tau_H - t)}|\mathcal{F}_t\right] .$$

To conclude, the ambiguity averse buyer of this option uses a worst–case prior $P^\star$ that is the pasting of P^κ after $P^{-\kappa}$ at knock–in time τ_H. As a consequence, at time $t = 0$ for instance, the value of this option under κ–ambiguity is

$$V_0 = E^{-\kappa}\left[E^\kappa\left[(K - S_{\tau^\star})^+\, e^{-r\tau^\star}|\mathcal{F}_{\tau_H}\right]\right] ,$$

where

$$\tau^\star = \inf\{t \geq \tau_H | V_t = (K - S_t)^+ e^{-rt}\}$$

denotes an optimal stopping time after knock–in.

As in the related discrete time case, the pessimistic holder thus presumes a change of drift at knock–in. Before the option becomes valuable, she uses the lowest mean return in her computations, and afterward, she uses the highest mean return.

4.2.2 American Straddle

The American Straddle has payoff $f(t,x) = e^{-rt}|x-K|$ for some strike K and interest rate r; it is used as an insurance against large movements of the underlying and plays also a role in certain bets on the volatility of an asset. When selling them short, the investor bets on stable courses[19].

To keep the technicalities minimal, let us work here with the Bachelier[20] model of financial markets where the underlying S is a Brownian motion under the reference measure P_0, i.e.

$$dS_t = dB_t, S_0 = 0\,.$$

Let the strike price be $K = 0$ (this is merely a normalization). To have a stationary problem, let the time horizon be infinite.

In the non–ambiguous case, the investor maximizes

$$E_0 e^{-r\tau} \left|B_\tau\right| \,.$$

The investor exercises the option when the asset price has moved sufficiently far away from the strike; the optimal stopping time has thus the form

$$\tau_b = \inf\left\{t \geq 0 || S_t| = b\right\}$$

for some critical number $b > 0$ which is explicitly given by

$$b = \frac{x^*}{\sqrt{2r}} \cong \frac{1.199679}{\sqrt{2r}}$$

and $x^* \cong 1.199679$ is the unique positive root of the equation $x\tanh(x) = 1$. We postpone the details of this classical stopping problem to the appendix.

Our focus is here on the role of ambiguity. So let us introduce drift ambiguity with a parameter $\kappa > 0$. The buyer of the option profits if the asset price moves away from zero. We thus guess that the worst case measure assigns maximal positive drift to the asset price if it is negative, and maximal

[19]A spectacular example involving straddles is the bankruptcy of the Barings Bank, see http://riskinstitute.ch/137560.htm.

[20]Louis Bachelier developed Brownian motion as a model of stock markets in 1900 (Bachelier (1900)) long before modern option pricing started with Samuelson, Merton, and Black and Scholes. See also the book *Louis Bachelier: Aux Origines de la Finance Mathématique*, Courtault and Kabanov (2002).

negative drift if the asset price is positive. Under the (guessed) worst–case measure P^*, the dynamics of S are therefore

$$dS_t = -\kappa \operatorname{sgn}(X_t)dt + dB_t^*$$

where B^* is a standard Brownian motion under P^*.

We show below how to find the value function for this problem and verify that the solution of this classical stopping problem (with a very non–classical dynamics for the asset price) solves indeed the stopping problem under ambiguity.

For the moment, we note that the rational investor changes his belief about the asset's drift quite frequently. Indeed, when X is close to the strike $K = 0$, the diffusion nature of the asset leads to infinitely many sign changes in any small interval. As a consequence, the investor changes his belief about the drift from $+\kappa$ to $-\kappa$ infinitely often in any time interval. Ambiguity might thus play a role in explaining the high psychological stress that grabs market participants sometimes.

We explain now how to find the value function under P^* (for $\kappa = 0$, you obtain the solution of the classical case). The infinite time horizon renders the problem homogenous in time, so we guess that the value function is of the form $e^{-rt}v(x)$ for some continuous function $v : \mathbb{R} \to \mathbb{R}$ which we assume to be twice continuously differentiable in the continuation region. By symmetry and time–homogeneity, we guess that it is optimal to wait until S hits a critical number c or $-c$, with $c > 0$. On $(0, c)$, our Bellman equation reduces then to

$$-rv - \kappa v' + \frac{1}{2}v'' = 0 \tag{21}$$

with a general solution

$$A_1 \exp(\alpha_1 x) + A_2 \exp(\alpha_2 x)$$

where $\alpha_{1,2}$ are the two solutions of the characteristic equation

$$-r - \kappa x + \frac{1}{2}x^2 = 0\,. \tag{22}$$

The constants A_1 and A_2 can be found by using continuity and smooth fit at the exercise point c, i.e.

$$v(c) = c \tag{23}$$

$$v'(c) = 1\,. \tag{24}$$

For negative values of x, we have a different Bellman equation, namely

$$-rv + \kappa v' + \frac{1}{2} v'' = 0 \tag{25}$$

where you note the sign change in front of v' because the agent changes his belief about the drift. The general solution here is of the form

$$B_1 \exp(\beta_1 x) + B_2 \exp(\beta_2 x)$$

where $\beta_{1,2}$ are the two solutions of the new characteristic equation

$$-r + \kappa x + \frac{1}{2} x^2 = 0\,. \tag{26}$$

Again, B_1 and B_2 can be found by using continuous and smooth fit at $-c$:

$$v(-c) = c \tag{27}$$

$$v'(c) = -1\,. \tag{28}$$

The only remaining unknown is the constant c at which we exercise. A first guess might be that we can find it by pasting continuously the two parts of our value function together at $x = 0$. But this continuity is satisfied automatically here (see the appendix). Had we started with asymmetric exercise points c and $-d$, the continuity condition would tell us that $c = d$. As we have made this guess from the start, the continuity condition does not yield new information now. The right condition to determine c is

$$v'(0-) = v'(0+) = 0\,. \tag{29}$$

That is, we paste the two solutions of the different Bellman equations on the right and on the left of zero smoothly together at zero in such a way that the value function has a minimum there. Economically, this is quite natural: as the payoff function has a minimum at zero, and we are in a completely symmetric situation (also the ambiguity is modeled in a symmetric way!), it is natural to expect the value function to have a minimum at zero, too.

From a more technical perspective, we paste together the solutions of two distinct differential equations on the negative and the positive axis in zero. The value function has to be continuously differentiable in zero because we would else lose the martingale property of the process $e^{-rt}v(W_t)$ in the continuation region. If the value function was not continuously differentiable in zero, a local time term would appear in the Itô formula[21] that would create trouble. So v' continuous with $v'(0) = 0$ is necessary. We show in the appendix that this even implies that the value function is twice continuously differentiable.

Theorem 4.3 *The conditions* (21) *to* (29) *determine uniquely a convex function v that is continuously differentiable. $e^{-rt}v(x)$ is the value function of the American straddle under drift ambiguity with parameter κ. In particular, the measure P^* where the asset price satisfies the dynamics*

$$dS_t = -\kappa \,\mathrm{sgn}(S_t)dt + dB_t^*$$

for the P^*–Brownian motion B^* is the worst–case measure.

5 Conclusion

We have now a fairly complete understanding of optimal stopping under Knightian uncertainty. We used the methods of stochastic calculus to characterize the value function either as a nonlinear supermartingale or as the solution to a Bellman equation. In a number of case studies, we are able to obtain closed–form solutions, as in the classical case of pure risk. As usual, the power of calculus and the beauty of the solutions justify the use of a continuous–time model even though most economic decisions take place in discrete time.

In Markovian models, the methods of stochastic calculus allow to extend the classical Bellman equation to Knightian uncertainty. Knightian uncertainty is mirrored in a nonlinear drift term that involves the first derivative of the value function with respect to the state variable. In monotone examples like the American Call or Put, or in irreversible investment problems, the nonlinearity disappears and one can easily identify the worst–case measure as the one where the state variable has the most unfortunate drift. In more complex settings, as exemplified by Barrier Options or American Straddles, the nonlinearity in the Bellman equation persists, and the worst–case measure has a highly nontrivial structure. In the example of the straddle, we have seen that a (rational!) ambiguity–averse investor can change his worst belief

[21]See, e.g., Karatzas and Shreve (1991), Theorem 3.6.22 for Itô's formula for convex functions.

about the drift infinitely often in critical times when the state variable is in a region where the monotonicity of the value function changes. Maybe a hint that "panic" is quite a natural phenomenon under Knightian uncertainty.

One of the big advantages of the model originally proposed by Gilboa and Schmeidler to capture uncertainty aversion is that it allows for a time–consistent extension and a rich dynamic theory. Our results can now be used to tackle more complex economic problems. As an example, let us mention the capacity extension program of a firm that can continuously invest into capacity (Pindyck (1988), Bertola (1988), Riedel and Su (2009)).

A A Characterization of g–Supermartingales

As in the discrete–time case of Riedel (2009), it is useful and important to understand g–supermartingales and relate them to classical supermartingales. In the case of drift ambiguity with parameter $\kappa > 0$, we shall see below that the counterpart of Lemma 6 in Riedel (2009) holds true: a pessimistic agent with multiple priors views a game against nature as unfair (a g–supermartingale) if and only if the game is unfair (a P–supermartingale) under some prior P in his set of priors.

Before we state the lemma, let us recall some notation. Let g be a standard driver for variational expectations as in Definition 2.2. Let

$$f(\omega, t, \theta) = \sup_{z \in \mathbb{R}^d} g(\omega, t, z) - z \cdot \theta \qquad \left(\omega \in \Omega, t \in [0,T], \theta \in \mathbb{R}^d\right)$$

be the convex dual of g. We denote by D the set of all progressively measurable processes (θ_t) such that

$$E_0 \int_0^T f\left(s, \theta_s\right)^2 ds < \infty\,.$$

Lemma A.1 *Let V be a process satisfying Assumption 2.1. Let g be a standard aggregator for variational expectations with Lipschitz constant $\kappa \geq 0$. Then the following assertions are equivalent:*

1. *V is a g–supermartingale,*
2. *the process $S_t = V_t + \int_0^t f\left(s, \theta_s\right) ds$ is a P^θ–supermartingale for some $\theta \in D$,*

3. V has a unique Doob–Meyer–Peng *decomposition of the form*

$$V_t = V_T + \int_t^T g(s, Z_s)ds - \int_t^T Z_s dB_s + A_T - A_t \qquad (30)$$

for a càdlàg increasing process A with $A_0 = 0$ and $E_0 A_T^2 < \infty$ and an adapted, square–integrable process Z.

V is a g–martingale if and only if the process A in the decomposition (30) *is zero.*

Before we prove this lemma, we also provide the generalization of the characterization of multiple prior super– and submartingales in Riedel (2009), Lemma 6.

Lemma A.2 *Let V be a process satisfying Assumption 2.1. Let $g(z) = -\kappa|z|$ be the driver for κ–ambiguity. Then V is a g–supermartingale if and only if it is a P–supermartingale fpr some $P \in \mathcal{P}^\kappa$. V is a g–submartingale if and only if it is a P–submartingale for all $P \in \mathcal{P}^\kappa$.*

The submartingale part of the above lemma is immediate from the definitions, see also Riedel (2009), Lemma 6. The supermartingale part follows from Lemma A.1 and the fact that the convex dual f takes the value 0 for $\theta \in \mathcal{D}^\kappa$, the set of all Girsanov kernels bounded by κ.

Let us now come to the proof of Lemma A.1.

Proof: The Doob–Meyer–Peng decomposition for g–supermartingales, that is, the equivalence of 1. and 3. above is due to Peng, Peng (1999), Theorem 3.3. and Corollary 3.12. In this original paper, the decomposition is formulated for strong g–supermartingales. The Optional Sampling Theorem of Chen and Peng (2000) show that strong and weak g–supermartingales coincide. For our version, see also Coquet, Hu, Mémin, and Peng (2002), Theorem 6.3.

If V is a g–martingale, then, by definition, $V_t = \mathcal{E}_t(V_T)$; so, there exists a square–integrable process Z such that (V, Z) is the unique solution of the backward stochastic differential equation with driver g and terminal value V, i.e.

$$V_t = V_T + \int_t^T g(s, Z_s)\, ds - \int_t^T Z_s dB_s$$

which shows that A is zero. On the other hand if $A = 0$ in (30), then (V, Z) is the unique solution of the backward stochastic differential equation with driver g and terminal value V, and this entails $V_t = \mathcal{E}_t(V_T)$, the g–martingale property.

Now we prove the implication 3. to 2. Choose a process $\theta^* \in D$ that solves

$$f(t, \theta_t^*) = g(t, Z_t) - Z_t \theta_t^* .$$

See Footnote 11 in El Karoui, Peng, and Quenez (1997) on existence and measurability of θ^*. There, it is also explained that $|\theta_t^*| \leq \kappa$ (because g is Lipschitz with constant κ). So, θ^* is bounded. Hence, there exists a measure P^{θ^*} equivalent to P_0 with Girsanov kernel θ^*. By Girsanov's theorem, $B_t^* = B_t - \int_0^t \theta_s^* ds$ is a standard Brownian motion under P^{θ^*}. We then get

$$-dS_t = g(t, Z_t)dt + dA_t - Z_t \cdot dB_t$$

which is – by definition of θ^* – equal to

$$\begin{aligned} &= \left(f\left(t, \theta_t^*\right) + Z_t \cdot \theta_t^*\right) dt + dA_t - Z_t \cdot dB_t \\ &= f\left(t, \theta_t^*\right) dt + dA_t - Z_t \cdot dB_t^* . \end{aligned}$$

. Hence, $S_t = V_t + \int_0^t f\left(s, \theta_s^*\right) ds$ is a P^{θ^*}–supermartingale.

To see the implication from 2. to 3., recall the dual representation of g–expectations from (5): we have for $t \leq \ t + u \leq T$

$$\mathcal{E}_t\left(V_{t+u}\right) = \operatorname*{ess\,inf}_{\theta \in D} E^{\theta} \left[V_{t+u} + \int_t^{t+u} f(s, \theta_s) ds | \mathcal{F}_t\right] .$$

From 3. and $\theta^* \in D$, we then get immediately

$$\begin{aligned} \mathcal{E}_t\left(V_{t+u}\right) &\leq E^{\theta^*} \left[V_{t+u} + \int_t^{t+u} f(s, \theta_s) ds | \mathcal{F}_t\right] \\ &= E^{\theta^*} \left[S_{t+u} | \mathcal{F}_t\right] - \int_0^t f(s, \theta_s) ds \end{aligned}$$

which – S is a P^{θ^*}–supermartingale – is less or equal to

$$\leq S_t - \int_0^t f(s, \theta_s) ds = V_t .$$

This concludes the proof. □

B The Structure of Optimal Stopping Times: Proof of Theorem 3.1

Let us first note that V_t is well–defined because for every stopping time $\tau \leq T$, the payoff X_τ is bounded by the square–integrable random variable $\sup_{t\in[0,T]} X_t$ by Assumption 2.1. From now on, all stopping times appearing in this proof are assumed to be bounded by T.

It is important that we can approximate the value V_t by choosing an appropriate sequence of stopping times $\tau_n \geq t$ with $\mathcal{E}_t\left(X_{\tau_n}\right) \uparrow V_t$. To this end, we need the following lemma.

Lemma B.1 *For all $t \geq 0$, the family*

$$\left\{\mathcal{E}_t\left(X_\sigma\right) : \sigma \geq t\right\}$$

is upwards directed.

Proof: Let $\sigma, \tau \geq t$ be two stopping times. Let

$$\xi_1 = \mathcal{E}_t\left(V_\sigma\right), \xi_2 = \mathcal{E}_t\left(V_\tau\right) .$$

It is sufficient to find a stopping time $\nu \geq t$ with

$$\max\left\{\xi_1, \xi_2\right\} = \mathcal{E}_t\left(V_\nu\right) .$$

Set

$$A := \left\{\xi_1 \geq \xi_2\right\}$$

and note that $A \in \mathcal{F}_t$. Our desired stopping time will be

$$\nu = \sigma 1_A + \tau 1_{A^c} ,$$

because we have

$$\begin{aligned} \max\left\{\xi_1, \xi_2\right\} &= \xi_1 1_A + \xi_2 1_{A^c} \\ &= \left(\mathcal{E}_t\left(V_\sigma\right)\right) 1_A + \left(\mathcal{E}_t\left(V_\tau\right)\right) 1_{A^c} \end{aligned}$$

which, by property (8) is equal to

$$= \mathcal{E}_t\left(V_\sigma 1_A + V_\tau 1_{A^c}\right) = \mathcal{E}_t\left(V_\nu\right) .$$

□

For every $t \geq 0$, the last lemma allows us to choose a sequence $(\tau_n(t))$ of stopping times greater or equal t with

$$\mathcal{E}_t\left(X_{\tau_n(t)}\right) \uparrow V_t\,.$$

By monotone continuity (property (10)), and by the law of iterated expectations (12), we obtain for $0 \leq s < t \leq T$

$$\begin{aligned}\mathcal{E}_s\left(V_t\right) &= \mathcal{E}_s\left(\operatorname*{ess\,sup}_{\tau\geq t}\mathcal{E}_t\left(X_\tau\right)\right)\\ &= \mathcal{E}_s\left(\lim_{n\to\infty}\mathcal{E}_t\left(X_{\tau_n(t)}\right)\right)\\ &= \lim_{n\to\infty}\mathcal{E}_s\left(\mathcal{E}_t\left(X_{\tau_n(t)}\right)\right)\\ &= \lim_{n\to\infty}\mathcal{E}_s\left(X_{\tau_n(t)}\right)\\ &\leq \operatorname*{ess\,sup}_{\tau\geq s}\mathcal{E}_t\left(X_\tau\right) = V_s\end{aligned}$$

and we conclude that the value process (V_t) is a g–supermartingale.

Our next task is to find a version of the value process (V_t) with rightcontinuous sample paths. We are going to use our generalization (Lemma F.1) of a classical result in martingale theory. All we need to show is that the ex ante expectations $t \mapsto \mathcal{E}_0(V_t)$ are rightcontinuous. So let $t_n \downarrow t$. First, we note that by the supermartingale property, $\mathcal{E}_0(V_{t_n})$ is increasing and less or equal $\mathcal{E}_0(V_t)$. Hence, the limit exists and $\lim_{n\to\infty}\mathcal{E}_0(V_{t_n}) \leq \mathcal{E}_0(V_t)$. Now fix $\epsilon > 0$. There exists a number $m \in \mathbb{N}$ with $\mathcal{E}_0\left(X_{\tau_m(t)}\right) \geq \mathcal{E}_0(V_t) - \epsilon$ (for the sequence $(\tau_k(t))_{k\in\mathbb{N}}$ chosen above, after Lemma B.1). Define a new sequence of stopping times via

$$\nu_n = \max\{\tau_m(t), t_n\}\,.$$

Clearly, $\nu_n \geq t_n$ and $\nu_n \downarrow \tau_m(t)$. By rightcontinuity of the payoff process X, we have $X_{\nu_n} \to X_{\tau_m(t)}$, and by dominated convergence (property (11)), we also have $\mathcal{E}_0(X_{\nu_n}) \to \mathcal{E}_0\left(X_{\tau_m(t)}\right)$. We thus obtain

$$\mathcal{E}_0(V_t) \leq \mathcal{E}_0\left(X_{\tau_m(t)}\right) + \epsilon = \lim_{n\to\infty}\mathcal{E}_0\left(X_{\nu_n}\right) + \epsilon\,.$$

The law of iterated expectations (property (12)), the definition of the value function, and monotonicity (property (6)) allow us to conclude

$$\lim_{n\to\infty}\mathcal{E}_0\left(X_{\nu_n}\right) = \lim_{n\to\infty}\mathcal{E}_0\left(\mathcal{E}_{t_n}\left(X_{\nu_n}\right)\right) \leq \lim_{n\to\infty}\mathcal{E}_0\left(V_{t_n}\right)\,.$$

Putting all these things together, we obtain

$$\lim_{n\to\infty} \mathcal{E}_0\left(V_{t_n}\right) \le \mathcal{E}_0\left(V_t\right) \le \lim_{n\to\infty} \mathcal{E}_0\left(V_{t_n}\right) + \epsilon\,.$$

As $\epsilon > 0$ was arbitrary, we get the desired rightcontinuity. Hence, by Lemma F.1, we can assume that (V_t) has rightcontinuous paths.

Now let (S_t) be another rightcontinuous g–supermartingale with $S_t \ge X_t$ for all $t \in [0,T]$. Monotonicity (property (6)) and the Optional Sampling Theorem (property (14)) yield immediately

$$V_t = \lim_{n\to\infty} \mathcal{E}_t\left(X_{\tau_n(t)}\right) \le \lim_{n\to\infty} \mathcal{E}_t\left(S_{\tau_n(t)}\right) \le S_t\,.$$

Hence, (V_t) is the smallest rightcontinuous g–supermartingale dominating X.

We continue our proof of Theorem 3.1 with a characterization of optimal stopping times.

Lemma B.2 *A stopping time $\tau \le T$ is optimal if and only if we have $X_\tau = V_\tau$ and $M := (M_t)_{t\in[0,T]} := \left(V_{t\wedge\tau}\right)_{t\in[0,T]}$ is a g–martingale.*

PROOF: If we have $X_\tau = V_\tau$ and the g–martingale property for the value process stopped at τ, then we obtain

$$\mathcal{E}_0\left(X_\tau\right) = \mathcal{E}_0\left(V_\tau\right) = V_0$$

and τ is optimal. Conversely, if τ is optimal, then

$$V_0 = \mathcal{E}_0\left(X_\tau\right) \le \mathcal{E}_0\left(V_\tau\right) \le V_0\,, \tag{31}$$

where we have used the fact that V dominates X and is a g–supermartingale (and the optional sampling theorem (14)). Strict monotonicity ((6) and the remark thereafter) of our variational expectation then implies $X_\tau = V_\tau$. We also get that the stopped process $\left(V_{t\wedge\tau}\right)_{t\in[0,T]}$ is a g–martingale because else the last inequality in (31) would be strict. □

We can now continue the proof of Theorem 3.1. As $V_T = X_T$, the infimum in the definition of τ^* is well–defined. By rightcontinuity of X and V, we get $X_{\tau^*} = V_{\tau^*}$[22]. It remains to show that the stopped process $\left(V_{t\wedge\tau^*}\right)_{t\in[0,T]}$ is a g–martingale.

[22]Let us remark here that the preceding two sentences have to be changed in the case of $T = \infty$. For infinite horizon, we have to *assume* that the infimum is finite a.s. to conclude $X_{\tau^*} = V_{\tau^*}$.

The idea is to approximate τ^* from the left[23] with the help of the stopping times

$$\tau^\lambda := \inf\{t \geq 0 : X_t \geq \lambda V_t\}$$

for $0 < \lambda < 1$.

Note that the family is increasing in λ, and bounded by τ^*. Hence, the stopping time $\bar{\tau} = \lim_{\lambda \uparrow 1} \tau^\lambda$ exists and is less or equal τ^*.

We show in Lemma B.3 below that we have

$$V_0 = \mathcal{E}_0\left(V_{\tau^\lambda}\right) .$$

The definition of τ^λ, leftcontinuity of X and dominated convergence allow to conclude

$$V_0 = \mathcal{E}_0\left(V_{\tau^\lambda}\right) \leq \frac{1}{\lambda}\mathcal{E}_0\left(X_{\tau^\lambda}\right) \to_{\lambda \uparrow 1} \mathcal{E}_0\left(X_{\bar{\tau}}\right) \leq V_0 .$$

Hence, $\bar{\tau}$ is an optimal stopping time. By Lemma B.2, we have that $X_{\bar{\tau}} = V_{\bar{\tau}}$. By definition of τ^*, we thus have $\tau^* \leq \bar{\tau}$; as we already know the other inequality, $\tau^* = \bar{\tau}$ follows. Invoking again Lemma B.2, we see that $(V_{t \wedge \tau^*})$ is a g–martingale. This concludes the proof of Theorem 3.1.

We still have to check the claim used in the above proof:

Lemma B.3 *With the notation introduced above,*

$$V_0 = \mathcal{E}_0\left(V_{\tau^\lambda}\right) .$$

Proof: For this lemma, we introduce more generally for all $t \geq 0$ and $0 < \lambda < 1$ the family of stopping times

$$\tau^\lambda(t) = \inf\{u \geq t : X_u \geq \lambda V_u\} .$$

As above, we have

$$\tau^\lambda(t) \uparrow_{\lambda \uparrow 1} \bar{\tau}(t)$$

for some stopping time $\bar{\tau}(t) \leq \tau^*(t) := \inf\{u \geq t : X_u = V_u\}$.

Introduce the process

$$W_t := \mathcal{E}_t\left(V_{\tau^\lambda(t)}\right) .$$

[23] Here is the point where we use leftcontinuity of X; actually, as you see from the proof, "leftcontinuity in g–expectation over stopping times" would be enough.

We claim that W is a g–supermartingale that admits rightcontinuous sample paths. For $0 \le t \le t+u \le T$ we have

$$\begin{aligned} \mathcal{E}_t\left(W_{t+u}\right) &= \mathcal{E}_t\left(\mathcal{E}_{t+u}\left(V_{\tau^\lambda(t+u)}\right)\right) \\ \text{iterated expectations} &= \mathcal{E}_t\left(V_{\tau^\lambda(t+u)}\right) \\ V\ g\text{–supermartingale} &\le \mathcal{E}_t\left(V_{\tau^\lambda(t)}\right) = W_t\,. \end{aligned}$$

So, W is a g–supermartingale. For rightcontinuity, use again Lemma F.1. Let $t_n \downarrow t$. Then $\tau^\lambda(t_n) \downarrow \tau^\lambda(t)$ by rightcontinity of X and V. Then also $V_{\tau^\lambda(t_n)} \to V_{\tau^\lambda(t)}$ almost surely and in g–expectation (Assumption 2.1). So

$$\mathcal{E}_0\left(W_{t_n}\right) = \mathcal{E}_0\left(\mathcal{E}_{t_n}\left(V_{\tau^\lambda(t_n)}\right)\right) = \mathcal{E}_0\left(V_{\tau^\lambda(t_n)}\right) \to_{n\to\infty} \mathcal{E}_0\left(V_{\tau^\lambda(t)}\right)\ .$$

Now look at the rightcontinuous g–supermartingale

$$Y_t = \lambda V_t + (1-\lambda) W_t\,.$$

We claim that W dominates X. For $X_t \ge \lambda V_t$, we have $\tau^\lambda(t) = t$, hence $W_t = V_t$ and $Y_t = V_t \ge X_t$. If $X_t < \lambda V_t$, we have $W_t \ge 0$ as $X \ge 0$ (Assumption 2.1), so

$$Y_t \ge \lambda V_t \ge X_t\,.$$

From Theorem E.1, we get $Y \ge V$. This is equivalent to $W \ge V$. On the other hand, by definition of W and the g–supermatingale property of V: $W \le V$. So we conclude $W = V$. In other words, we finally get

$$V_t = W_t = \mathcal{E}_t\left(V_{\tau^\lambda(t)}\right)\ .$$

This proves the lemma (set $t=0$). □

C Reflected Backward Stochastic Differential Equations

Proof of Corollary 3.2

V is a rightcontinuous g–supermartingale; by Assumption 2.1, we have $V_t \le \sup_{t\in[0,T]} X_t \in L^2(\Omega, \mathcal{F}, P_0)$. We can thus apply the Doob–Meyer–Peng decomposition of g–supermartingales from Lemma A.1 to get an increasing, adapted, rightcontinuous process with left limits K and an adapted process Z with $E_0 \int_0^T \|Z_t\|^2 dt < \infty$ such that

$$V_t = X_T + \int_t^T g(s, Z_s) ds - \int_t^T Z_s dB_s + K_T - K_t \, .$$

It remains to show $\int_0^T (V_u - X_u) du = 0$ which is equivalent to

$$K_{\nu(t)} = K_t$$

for all $0 \leq t < T$ and

$$\nu(t) = \inf \{u \geq t : V_u = X_u\} \ .$$

Note that $\nu(t)$ is an optimal stopping time if we start at time t; by our characterization of optimal stopping times, $\left(V_{u \wedge \nu(t)}\right)_{t \leq u \leq T}$ is then a g–martingale. From Lemma A.1 (applied with the starting time t), we see that the increasing process in the Doob–Meyer–Peng decomposition must be zero, that is, $K_{\nu(t)} = K_t$. This proves minimality of K. Hence, (V, Z, K) is indeed a solution of the reflected backward stochastic differential equation with obstacle X, terminal condition X_T and driver g .

To see uniqueness, let (W, ζ, A) be another solution of the reflected backward stochastic differential equation with obstacle X, terminal condition X_T and driver g. Then W is a g–supermartingale that dominates X, and from Theorem 3.1, we get $W \geq V$. Now let

$$\nu(t) = \inf \{u \geq t : W_u = X_u\} \ .$$

(The infimum is taken over a nonempty set because of the terminal condition $W_T = X_T$.) By rightcontinuity, we have $W_{\nu(t)} = X_{\nu(t)}$. As the increasing process A is minimal, we also have $A_{\nu(t)} = A_t$. We conclude that $\left(W_{u \wedge \nu(t)}\right)_{t \leq u \leq T}$ is a g–martingale. It follows

$$V_t \geq \mathcal{E}_t \left(X_{\nu(t)}\right) = \mathcal{E}_t \left(W_{\nu(t)}\right) = W_t \, .$$

Therefore, $V = W$. From the uniqueness of the Doob–Meyer–Peng decomposition in Lemma A.1, we obtain $Z = \zeta$ and $A = C$.

D Markovian Models: Proof of Theorem 3.6

To start with, note that the family $\sup_{t\in[0,T]}|v(t,X_t)+f(t,X_t)| \in L^2(\Omega,\mathcal{F},P_0)$. This is due to our growth estimate (19) on v (resp. (17) for f) and the fact that $\sup_{t\in[0,T]}\|X_t\|$ is in any L^q–space, see, e.g. Jacka and Lynn (1992), Lemma 2.1. (and its proof). Hence, we can take the g–conditional expectation when required below.

By a similar argument, one shows that the process

$$g(u,\sigma(S_u)^\top\nabla v(u,S_u)), 0\le u\le T$$

is square–integrable: this is due to Assumption 3.5, the growth condition (19) on the first derivative of the value function, the fact that $\sup_{t\in[0,T]}\|X_t\|$ is in any L^q–space, and last not least the Lipschitz–continuity of the driver g. We shall need this below to apply our Lemma A.1.

As v is in $\mathbb{W}^{1,2}$, we can apply Krylov's (Krylov (1980), Theorem 1.10.1) generalized version of Itô's lemma and get with the help of the Hamilton–Jacobi–Bellman Equation (18) for any stopping time $\tau\ge t$

$$\begin{aligned}
v(t,S_t) &= v(\tau,S_\tau)-\int_t^\tau \left(v_t(u,S_u)+\mathcal{L}v(u,S_u)\right)du-\int_t^\tau \nabla v(u,S_u)^\top\sigma(S_u)dB_u\\
&= v(\tau,S_\tau)+\int_t^\tau g(u,\sigma(S_u)^\top\nabla v(u,S_u))du-\int_t^\tau \nabla v(u,S_u)^\top\sigma(S_u)dB_u\\
&\quad -\int_t^\tau \big(v_t(u,S_u)+\mathcal{L}v(u,S_u)+g(u,\sigma(S_u)^\top\nabla v(u,S_u))\big)du\\
&= v(\tau,S_\tau)+\int_t^\tau g(u,\sigma(S_u)^\top\nabla v(u,S_u))du-\int_t^\tau \nabla v(u,S_u)^\top\sigma(S_u)dB_u\\
&\quad +A_\tau-A_t\,, \qquad (32)
\end{aligned}$$

where

$$A_t=-\int_0^t \big(v_t(u,S_u)+\mathcal{L}v(u,S_u)+g(u,\sigma(S_u)^\top\nabla v(u,S_u))\big)du$$

is an increasing process due to our Bellman Equation (18). By Lemma A.1, $(v(t,S_t))_{t\in[0,T]}$ is a g–supermartingale. Our general Theorem 3.1 implies that $v(t,S_t)\ge V_t$ for all $t\in[0,T]$.

For the stopping time $\nu(t)=\inf\{u\ge t|v(t,S_u)=f(t,S_u)\}$, the Hamilton–Jacobi–Bellman equation (18) yields $A_{\nu(t)}-A_t=0$. As a consequence, $(v(u\wedge\nu(t),S_{u\wedge\nu(t)}))_{t\le u\le T}$ is a g–martingale, again by Lemma A.1.

By continuity of v and f, and continuity of the sample paths of S, we also have $v(\nu(t), S_{\nu(t)}) = f(\nu(t), S_{\nu(t)})$. It follows that

$$V_t \geq \mathcal{E}_t \left(f(\nu(t), S_{\nu(t)}) \right) = \mathcal{E}_t \left(v(\nu(t), S_{\nu(t)}) \right) = v(t, S_t)$$

and we obtain finally $V_t = v(t, S_t)$.

E Infinite Time Horizon: Proof of Theorem 3.7

We start by noting that Theorem 3.1 holds true for $T = \infty$ if the payoff process X satisfies Assumption 2.1 — the proof goes through without changes, except that — as in the case of pure risk — we have to assume that the candidate for an optimal stopping time

$$\tau^* = \inf \{t \geq 0 : V_t = X_t\}$$

remains finite a.s. (see also the footnote 22).

We also need to extend the concept of g–(super)martingale to the infinite time horizon. We call a process $(S_t)_{t\geq 0}$ that satisfies Assumption 2.1 for $T = \infty$ a g–supermartingale if it is a g–supermartingale for all finite horizons $T < \infty$.

Theorem E.1 *Let*

$$V_t = \operatorname*{ess\,sup}_{\tau \geq t} \mathcal{E}_t \left(X_\tau \right)$$

be the value function of the optimal stopping problem. Assume that

$$\tau^* = \inf \{t \geq 0 : V_t = X_t\}$$

is universally finite (i.e. $P[\tau < \infty] = 1$ for all $P \in \mathcal{P}^\kappa$). Then one can choose a version of V with rightcontinuous sample paths. Moreover,

1. *(V_t) is the smallest rightcontinuous g–supermartingale dominating (X_t);*
2. *$\tau^* = \inf \{t \geq 0 : V_t = X_t\}$ is an optimal stopping time;*
3. *the value function stopped at τ^*, $(V_{t\wedge\tau^*})$ is a g–martingale.*

Let us now come to the proof of Theorem 3.7. Assumption 2.1 ensures that the value process

$$V_t = \operatorname*{ess\,sup}_{\infty>\tau\geq t} \mathcal{E}_t\left(X_\tau\right) = \operatorname*{ess\,sup}_{\infty>\tau\geq t} \operatorname*{ess\,inf}_{P\in\mathcal{P}^\kappa} E^P\left[X_\tau|\mathcal{F}_t\right]$$

is well-defined. Without loss of generality, we can and do choose rightcontinuous versions for our value functions V^T and V.

As one increases one's options with a longer time horizon, we also have $V^T \leq V$ and obviously $V^\infty \leq V$ as well. It is then enough to show $V^\infty \geq V$.

The g–expectation is continuous from below (property 10), so the g–supermartingale property of V^T yields for all $t, u \geq 0$

$$\begin{aligned} V_t^\infty &= \lim_{T\uparrow\infty} V_t^T \\ &\leq \lim_{T\uparrow\infty} \mathcal{E}_t\left(V_{t+u}^T\right) \\ &= \mathcal{E}_t\left(\lim_{T\uparrow\infty} V_{t+u}^T\right) \\ &= \mathcal{E}_t\left(V_{t+u}^\infty\right) \,. \end{aligned}$$

Hence, V^∞ is a g–supermartingale.

Let us show next that V^∞ admits a rightcontinuous version. By Lemma F.1, it is enough to show that the mapping

$$t \mapsto \mathcal{E}_0\left(V_t^\infty\right)$$

is rightcontinuous. Fix $t_0 \geq 0$ and a sequence (t_n) with $t_n \downarrow t_0$. The mapping $n \mapsto \mathcal{E}_0\left(V_{t_n}^\infty\right)$ is increasing as V^∞ is a g–supermartingale. Also, $T \mapsto \mathcal{E}_0\left(V_t^T\right)$ is increasing for all $T \geq t \geq 0$. We can thus interchange the limits in the following computation and obtain the desired result:

$$\begin{aligned} \lim_{n\to\infty} \mathcal{E}_0\left(V_{t_n}^\infty\right) &= \lim_{n\to\infty} \mathcal{E}_0\left(\lim_{T\uparrow\infty} V_{t_n}^T\right) \\ \text{monotone continuity} &= \lim_{n\to\infty}\lim_{T\uparrow\infty} \mathcal{E}_0\left(V_{t_n}^T\right) \\ \text{remark above} &= \lim_{T\uparrow\infty}\lim_{n\to\infty} \mathcal{E}_0\left(V_{t_n}^T\right) \\ \text{monotone continuity} &= \lim_{T\uparrow\infty} \mathcal{E}_0\left(\lim_{n\to\infty} V_{t_n}^T\right) \\ \text{rightcontinuity of } V^T &= \lim_{T\uparrow\infty} \mathcal{E}_0\left(V_{t_0}^T\right) \\ \text{monotone continuity and definition of } V^\infty &= \mathcal{E}_0\left(V_{t_0}^\infty\right) \,. \end{aligned}$$

Choose a rightcontinuous version of V^∞, that we denote again by V^∞. By Theorem E.1, we must have $V^\infty \geq V$.

F Right–Continuous Versions of g–Supermartingales

As in the classical martingale theory, we need sometimes an argument that allows us to work with right–continuous sample paths. We prove here the corresponding version of the classical lemma[24].

Lemma F.1 *Let $(X_t)_{t\in[0,T]}$ be a g–supermartingale with*

$$E[\sup_{t\in[0,T]} X_t^2] < +\infty \,.$$

Assume that the mapping

$$t \in [0,T] \mapsto \mathcal{E}_0\left(X_t\right)$$

is continuous. Then there exists an event $\Omega^* \subset \Omega$ with $P\left(\Omega^*\right) = 1$ on which the (right–continuous!) process

$$Y_t = X_{t+} = \lim_{s\downarrow t, s\in\mathbb{Q}} X_s \qquad (t \in [0,T]) \tag{33}$$

is well–defined. We have $Y_t = X_t$ a.s. for all $t \in [0,T]$, and $(Y_t)_{t\in[0,T]}$ is also a g–supermartingale.

Proof: Chen and Peng (2000) prove a downcrossing inequality for g–supermartingales[25]. As in the classical case, it follows that on some event Ω^* with probability 1 the limit in (33) exists. Let us define $Y_t = 0$ outside the set Ω^*. By the usual conditions, Y is adapted, so $Y_t = \mathcal{E}_t\left(Y_t\right)$. Because of the g–supermartingale property and dominated convergence, we have

$$\mathcal{E}_t\left(Y_t\right) = \mathcal{E}_t\left(\lim_{s\downarrow t, s\in\mathbb{Q}} X_s\right) = \lim_{s\downarrow t, s\in\mathbb{Q}} \mathcal{E}_t\left(X_s\right) \leq X_t \,.$$

[24]Coquet, Hu, Mémin, and Peng (2002) mention in their Remark 2.2 the possibility to prove this lemma. As the proof has not been spelt out anywhere to our knowledge, we provide it here.Trevino (2008) proves rightcontinuity of value functions under multiple priors in a general semimartingale setting.

[25]The assumption of positivity is not necessary here as explained in Coquet, Hu, Mémin, and Peng (2002), Remark 2.1.

We thus have $Y_t \le X_t$ a.s. for all $t \in [0,T]$. On the other hand, we have assumed that $t \in [0,T] \mapsto \mathcal{E}_0(X_t)$ is continuous. From dominated convergence, we then get

$$\mathcal{E}_0(Y_t) = \lim_{s \downarrow t, s \in \mathbb{Q}} \mathcal{E}_0(X_s) = \mathcal{E}_0(X_t) \ .$$

The comparison theorem implies $Y_t = X_t$ a.s. □

G Barrier Options

We provide here the proof of Lemma 4.2.

Fix a progressively measurable process θ with $|\theta_t| \le \kappa$ for all $t \in [0,T]$. Under P^θ, $B_t^\theta = B_t - \int_0^t \theta_s ds$ is a standard Brownian motion. By definition, we have

$$\begin{aligned} S_u &= S_t \exp\left(\sigma(B_u - B_t) + (\mu - \sigma^2/2)(u-t)\right) \\ &= S_t \exp\left(\sigma(B_u^\theta - B_t^\theta) + (\mu - \sigma^2/2)(u-t) + \int_t^u \theta_s ds\right) \ . \end{aligned}$$

Given $\tau_H > t$, the event that we reach the level H later than u can be written as

$$\begin{aligned} \{\tau_H > u\} &= \{\forall v \in [t,u] : S_v \le H\} \\ &= \left\{\forall v \in [t,u] : B_v^\theta - B_t^\theta + \int_t^v \theta_s ds \le L\right\} \ , \end{aligned}$$

where we write

$$L = \frac{1}{\sigma}\left(\log(H/S_t) - (\mu - \sigma^2/2)(v-t)\right) \ .$$

Given that $\theta_s \ge -\kappa$, this set is included in

$$\left\{\forall v \in [t,u] : B_v^\theta - B_t^\theta - \kappa t \le L\right\}$$

As B^θ is a standard Brownian motion under P^θ and similarly, $B_t^{-\kappa} = B_t - \kappa t$ a standard Brownian motion under $P^{-\kappa}$, we have

$$\begin{aligned} &P^\theta\left[\forall v \in [t,u] : B_v^\theta - B_t^\theta - \kappa(v-t) \le L | \mathcal{F}_t\right] \\ &= P^{-\kappa}\left[\forall v \in [t,u] : B_v^{-\kappa} - B_t^{-\kappa} - \kappa(v-t) \le L | \mathcal{F}_t\right] \ . \end{aligned}$$

But $B_v^{-\kappa} - B_t^{-\kappa} - \kappa(v-t) = B_v - B_t$; we thus get

$$
\begin{aligned}
P^\theta\left[\tau_H > u|\mathcal{F}_t\right] = P^\theta\left[\forall v \in [t,u] : B_v^\theta - B_t^\theta + \int_t^v \theta_s ds \le L|\mathcal{F}_t\right] \\
\le P^\theta\left[\forall v \in [t,u] : B_v^\theta - B_t^\theta - \kappa(v-t) \le L|\mathcal{F}_t\right] \\
= P^{-\kappa}\left[\forall v \in [t,u] : B_v^{-\kappa} - B_t^{-\kappa} - \kappa(v-t) \le L|\mathcal{F}_t\right] \\
= P^{-\kappa}\left[\forall v \in [t,u] : B_v - B_t \le L|\mathcal{F}_t\right] \\
= P^{-\kappa}\left[\tau_H > u|\mathcal{F}_t\right] .
\end{aligned}
$$

H Details for the American Straddle

Let us start with the pure risk case, when $\kappa = 0$. As we are in the stationary infinite horizon case, the value function will have the form $e^{-rt}v(x)$. The payoff function being symmetric, it is natural to guess that we will exercise when the absolute value of the Brownian motion without drift B hits a critical value b. In the continuation region $C = (-b, b)$, the process $e^{-rt}v(B_t)$ is a martingale, which leads to the ordinary differential equation

$$-rv + 1/2v'' = 0$$

in C. The solutions of this differential equation are of the form

$$v(x) = A_1 \exp(\alpha_1 x) + A_2 \exp(\alpha_2 x)$$

for $\alpha_{1,2} = \pm\sqrt{2r}$ and some constants A_1, A_2. The standardway to determine these constants is to use continuity and smooth fit at the boundary b which yields the two linear equations (in $A_{1,2}$)

$$
\begin{aligned}
A_1 \exp(\alpha_1 b) + A_2 \exp(\alpha_2 b) = b \\
A_1\alpha_1 \exp(\alpha_1 b) + A_2\alpha_2 \exp(\alpha_2 b) = 1
\end{aligned}
$$

with determinant $-2\sqrt{2r}$ (note that $\alpha_1 + \alpha_2 = 0$) and solution (in terms of the still unknown b)

$$
\begin{aligned}
A_1 &= \frac{\sqrt{2r}b + 1}{2\sqrt{2r}} \exp(-\sqrt{2r}b) \\
A_2 &= \frac{\sqrt{2r}b - 1}{2\sqrt{2r}} \exp(\sqrt{2r}b) .
\end{aligned}
$$

Continuity and smooth fit in $-b$ yield the two equations

$$A_1 \exp(-\alpha_1 b) + A_2 \exp(-\alpha_2 b) = b$$
$$A_1 \alpha_1 \exp(-\alpha_1 b) + A_2 \alpha_2 \exp(-\alpha_2 b) = -1$$

which yields

$$A_1 = \frac{\sqrt{2r}b - 1}{2\sqrt{2r}} \exp(\sqrt{2r}b)$$
$$A_2 = \frac{\sqrt{2r}b + 1}{2\sqrt{2r}} \exp(-\sqrt{2r}b)\,.$$

We thus obtain two equations for b which we can luckily satisfy in a consistent way. b has to satisfy

$$\frac{\sqrt{2r}b - 1}{2\sqrt{2r}} \exp(\sqrt{2r}b) = \frac{\sqrt{2r}b + 1}{2\sqrt{2r}} \exp(-\sqrt{2r}b)$$

or, if we write $x = \sqrt{2r}b$, we need

$$x(\exp(x) - \exp(-x)) = \exp(x) + \exp(-x)\,,$$

or

$$x \tanh(x) = 1$$

which has a unique positive solution $x^* \cong 1.199679$.

We now come to Knightian uncertainty.

Lemma H.1 *Let*

$$\alpha_{1,2} = \kappa \pm \sqrt{2r + \kappa^2}$$
$$A_1 = \frac{\alpha_2 c - 1}{\alpha_2 - \alpha_1} \exp((\alpha_2 - 2\kappa)c)$$
$$A_2 = \frac{1 - \alpha_1 c}{\alpha_2 - \alpha_1} \exp((\alpha_1 - 2\kappa)c)$$
$$\beta_{1,2} = -\kappa \pm \sqrt{2r + \kappa^2}$$
$$B_1 = \frac{1 + \beta_2 c}{\beta_2 - \beta_1} \exp((-\beta_2 - 2\kappa)c)$$
$$B_2 = \frac{-1 - \beta_1 c}{\beta_2 - \beta_1} \exp((-\beta_1 - 2\kappa)c)\,.$$

Let $c > 0$ be the unique positive solution of the equation

$$\frac{c - \frac{1}{\alpha_1}}{c - \frac{1}{\alpha_2}} = \exp(-2\sqrt{2r + \kappa^2}\, c)\,.$$

The function

$$v(x) = \begin{cases} x & \text{if } x \geq c \\ A_1 \exp(\alpha_1 x) + A_2 \exp(\alpha_2 x) & \text{for } 0 \leq c \\ B_1 \exp(\beta_1 x) + B_2 \exp(\beta_2 x) & \text{for } -c < x < 0 \\ -x & \text{if } x \leq -c \end{cases}$$

is the unique solution of (21) *to* (29)*. v is convex, continuously differentiable everywhere, and twice continuously differentiable except at the critical points $\{-c, c\}$ (where the Straddle is exercised).*

Proof: The solutions of the quadratic equation (22) are

$$\alpha_{1,2} = \kappa \pm \sqrt{2r + \kappa^2}\,.$$

Note that $\alpha_1 + \alpha_2 = 2\kappa$. The ansatz

$$v(x) = A_1 \exp(\alpha_1 x) + A_2 \exp(\alpha_2 x)$$

for $x > 0$ leads to the following continuity and smooth fit condition at the (yet to be determined) exercise point $c > 0$:

$$\begin{aligned} A_1 \exp(\alpha_1 c) + A_2 \exp(\alpha_2 c) &= c \\ A_1 \alpha_1 \exp(\alpha_1 c) + A_2 \alpha_2 \exp(\alpha_2 c) &= 1\,. \end{aligned}$$

This system is linear in $A_{1,2}$ with determinant

$$det = \exp((\alpha_1 + \alpha_2)\kappa c)(\alpha_2 - \alpha_1) = \exp(2\kappa c)(\alpha_2 - \alpha_1) \neq 0$$

and unique solution (in terms of c)

$$\begin{aligned} A_1 &= \frac{\alpha_2 c - 1}{\alpha_2 - \alpha_1} \exp((\alpha_2 - 2\kappa)c) \\ A_2 &= \frac{1 - \alpha_1 c}{\alpha_2 - \alpha_1} \exp((\alpha_1 - 2\kappa)c)\,. \end{aligned}$$

On the negative axis, the solutions of the equation (26) are

$$\beta_{1,2} = -\kappa \pm \sqrt{2r + \kappa^2}\,.$$

Note that we have $\beta_1 = -\alpha_2$ and $\beta_2 = -\alpha_1$. The ansatz

$$v(x) = B_1 \exp(\beta_1 x) + B_2 \exp(\beta_2 x)$$

for $x < 0$ leads to the following continuity and smooth fit condition at the exercise point $-c$:

$$\begin{aligned} B_1 \exp(-\beta_1 c) + B_2 \exp(-\beta_2 c) &= c \\ B_1\beta_1 \exp(-\beta_1 c) + B_2\beta_2 \exp(-\beta_2 c) &= -1\,. \end{aligned}$$

As above, this leads to

$$\begin{aligned} B_1 &= \frac{1 + \beta_2 c}{\beta_2 - \beta_1} \exp((-\beta_2 - 2\kappa)c) \\ B_2 &= \frac{-1 - \beta_1 c}{\beta_2 - \beta_1} \exp((\alpha_1 - 2\kappa)c)\,. \end{aligned}$$

Note at this point that we have

$$B_1 = A_2, B_2 = A_1\,.$$

From this, we immediately get $A_1 + A_2 = B_1 + B_2$, i.e., the two solutions paste continuously together at 0; we have $v(0-) = v(0+)$. We now choose the constant c in such a way that v is even continuously differentiable in zero. The equation $v'(0-) = v'(0+)$ is equivalent to

$$A_1\alpha_1 + A_2\alpha_2 = B_1\beta_1 + B_2\beta_2$$

which we can transform to

$$A_1\alpha_1 + A_2\alpha_2 = -A_2\alpha_2 - A_1\alpha_1$$

in light of the above relations between the parameters. This equation can only be satisfied if both sides are equal to zero. We are thus led to consider

$$0 = A_1\alpha_1 + A_2\alpha_2$$

which is equivalent to

$$0 = \alpha_1(\alpha_2 c - 1)\exp(\alpha_2 c) + \alpha_2(1 - \alpha_2 c)\exp(\alpha_1 c)$$

or

$$0 = \alpha_1(\alpha_2 c - 1)\exp((\alpha_2 - \alpha_1)c) + \alpha_2 - \alpha_1\alpha_2 c$$

what we can rewrite as

$$\frac{c - \frac{1}{\alpha_1}}{c - \frac{1}{\alpha_2}} = \exp(-2\sqrt{2r + \kappa^2}\, c)\,.$$

On $[0, \infty[$, the left side is continuously increasing from $\alpha_2/\alpha_1 < 0$ to 1 and the right side is continuously decreasing from 1 to 0. Hence, there exists a unique number $c > 0$ that makes our guessed function continuously differentiable in zero with $v'(0) = 0$.

We remark that we then get a continuous second derivative in zero as well. By the Bellman equations (21) and (25) we have

$$\begin{aligned} v''(0+) - v''(0-) = 2\left(rv(0+) + \kappa v'(0+) - (rv(0-) - \kappa v'(0-))\right) \\ = 2\kappa v'(0) = 0\,. \end{aligned}$$

□

With the help of the preceding lemma and Itô's formula (which can be applied in the standard form because the function v is convex and in $\mathbb{W}^2$, or in other words, the generalized second derivative does not have a mass point in $\pm c$), is is easy to see that $e^{-rt}v(S_t)$ is a P^*–supermartingale, and a P^*–martingale on $(-c, c)$, compare Equation (32). To conclude the proof that $e^{-rt}v(S_t)$ is the value function under P^*, we just have to show that the candidate optimal stopping times

$$\nu(t) = \inf\{u \geq t | v(S_u) = |S_u|\}$$

remain finite a.s. (as we are now in the infinite horizon case). To see this, we introduce the scale function

$$s(x) = \begin{cases} \frac{e^{2\kappa x} - 1}{2\kappa} & \text{if } x \geq 0 \\ \frac{1 - e^{-2\kappa x}}{2\kappa} & \text{else} \end{cases}\,.$$

Then $s(S_t)$ is a P^*–martingale (again, Itô's formula can be applied because the second derivative of s exists as a measurable function) and by the usual time–change argument, $\tilde{B}_t = s(S_{T(t)})$ is a P^*–Brownian motion for some new time scale $(T(t))_{t\geq 0}$, see, e.g. Rogers and Williams (1987), V.46–47. We know that a P^*–Brownian motion (in fact, any regular diffusion) leaves $(-c, c)$ almost surely in finite time; as a consequence, so does S under P^* and we get $\nu(t) < \infty$ P^*-a.s.

We now have to verify that the function $e^{-rt}v(x)$ is indeed the value function under drift ambiguity. We know already that that $e^{-rt}v(S_t)$ is a P^*–supermartingale, and a P^*–martingale on $(-c, c)$. Now take any $P \in \mathcal{P}^\kappa$ with Girsanov kernel θ bounded by κ and apply Itô's formula (which can be applied in the standard form because the function v is convex and in $\mathbb{W}^2$, or in other words, the generalized second derivative does not have a mass point in $\pm c$), to see that for $t \leq \tau^*$

$$\begin{aligned} e^{-rt}v(S_t) &= v(S_0) + \int_0^t e^{-ru} \left(-rv(S_u) + \frac{1}{2}v''(S_u) \right) du + \int_0^t e^{-ru}v'(S_u)dB_u \\ &= v(S_0) + \int_0^t e^{-ru}\kappa \operatorname{sgn}(X_u)v'(X_u)du + \int_0^t e^{-ru}v'(S_u)dB_u \end{aligned}$$

by the Bellman equations (21) and (25); we continue to write

$$= v(S_0) + \int_0^t e^{-ru}v'(X_u) \left(\kappa \operatorname{sgn}(X_u) + \theta_u\right) du + \int_0^t e^{-ru}v'(S_u)dB_u^\theta$$

where B^θ is the Brownian motion under P^θ. As we have $\operatorname{sgn} v'(x) = \operatorname{sgn}(x)$ and $\operatorname{sgn}(\kappa \operatorname{sgn}(X_u) + \theta_u) = \operatorname{sgn}(X_u)$, we see that $e^{-rt}v(S_t)$ is a P–submartingale. We therefore have $v(S_0) = \inf_{P\in\mathcal{P}^\kappa} E^P e^{-r\tau^*}v(c) = \inf_{P\in\mathcal{P}^\kappa} E^P e^{-r\tau^*}c \leq V_0$. The other inequality being obvious, the proof is complete.

References

BACHELIER, L. (1900): "Théorie de la Spéculation," *Annales Scientifiques de l'École Normale Supérieure*, pp. 21–86.

BERTOLA, G. (1988): "Irreversible Investment," Working Paper, Princeton University.

CHEN, Z., AND L. EPSTEIN (2002): "Ambiguity, Risk and Asset Returns in Continuous Time," *Econometrica*, 70, 1403–1443.

CHEN, Z., AND S. PENG (2000): "A General Downcrossing Inequality for g–Martingales," *Statistics and Probability Letters*, 46, 169–175.

CHENG, X., AND F. RIEDEL (2010): "Approximation of Continuous–Time Optimal Stopping Problems," Discussion paper, Institute of Mathematical Economics.

CHERNOFF, H. (1972): *Sequential Analysis and Optimal Design*. SIAM.

COQUET, F., Y. HU, J. MÉMIN, AND S. PENG (2002): "Filtration–consistent Nonlinear Expectations and Related g–Expectations," *Probability Theory and Related Fields*, 123, 1–27.

COURTAULT, J.-M., AND Y. KABANOV (2002): *Louis Bachelier: Aux Origines de la Finance Mathématique*. Presses universitaires de Franche-Comté.

CRANDALL, M., H. ISHII, AND P. LIONS (1992): "User's Guide to Viscosity Solutions of Second–Order Partial Differential Equations," *Bulletin of the American Mathematical Society*, 27(1), 1–67.

DEGROOT, M. H. (2004): *Optimal Statistical Decisions*. Wiley.

DELBAEN, F., S. PENG, AND E. ROSAZZA GIANIN (2009): "Representation of the Penalty Term of Dynamic Concave Utilities," Working Paper.

DIXIT, A., AND R. PINDYCK (1994): *Investment under Uncertainty*. Princeton University Press, Princeton.

DUFFIE, D., AND L. EPSTEIN (1992): "Stochastic Differential Utility," *Econometrica*, 60, 353–394.

El Karoui, N. (1979): "Les aspects probabilistes du controle stochastique," in *Ecole dEte de Probabilites de Saint-Flour*, ed. by C. Dellacherie, and P. Meyer, pp. 74–238. Springer.

El Karoui, N., S. Peng, and M. Quenez (1997): "Backward Stochastic Differential Equations in Finance," *Mathematical Finance*, 7(1), 1–71.

ElKaroui, N., C. Kapoudjian, E. Pardoux, S. Peng, and M. Quenez (1997): "Reflected Solutions of Backward SDE's, and Related Obstacle Problems for PDE's," *Annals of Applied Probability*, 25(2), 702–737.

Fleming, W., and H. Soner (2006): *Controlled Markov Processes and Viscosity Solutions*. Springer.

Föllmer, H., P. Protter, and A. Shiryaev (1995): "Quadratic Covariation and an Extension of Itô's Formula," *Bernoulli*, 1, 149–169.

Jacka, S. (1991): "Optimal Stopping and the American Put," *Mathematical Finance*, 1, 1–14.

Jacka, S., and J. Lynn (1992): "Finite-horizon Optimal Stopping, Obstacle Problems and the Shape of the Continuation Region," *Stochastics and Stochastics Reports*, 39, 25–42.

Karatzas, I. (1988): "On the Pricing of American Options," *Applied Mathematics and Optimization*, 17, 37–60.

Karatzas, I., and S. E. Shreve (1991): *Brownian Motion and Stochastic Calculus*. Springer, New York.

Karatzas, I., and H. Wang (2000): "A Barrier Option of American Type," *Applied Mathematics and Optimization*, 42, 259–279.

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

Krylov, N. (1980): *Controlled Diffusion Processes*. Springer.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Dynamic Variational Preferences," *Journal of Economic Theory*, 128, 4–44.

MYNENI, R. (1992): "The Pricing of the American Option," *Annals of Applied Probability*, 2, 1–23.

NISHIMURA, K. G., AND H. OZAKI (2004): "Search and Knightian Uncertainty," *Journal of Economic Theory*, 119, 299–333.

——— (2007): "Irreversible Investment and Knightian Uncertainty," *Journal of Economic Theory*, 136, 668–694.

PENG, S. (1997): "BSDE and Related g–Expectation," in *Backward Stochastic Differential Equation*, ed. by N. El Karoui, and L. Mazliak, no. 364 in Pitman Research Notes in Mathematics Series. Addison Wesley Longman Ltd.

PENG, S. (1999): "Monotonic Limit Theorem of BSDE and Nonlinear Decomposition Theorem of Doob–Meyer's Type," *Probability Theory and Related Fields*, 113, 473–499.

PESKIR, G., AND A. SHIRYAEV (2006): *Optimal Stopping and Free Boundary Problems*, Lectures in Mathematics, ETH Zurich. Birkhauser.

PHAM, H. (1997): "Optimal Stopping, Free Boundary, and American Option in a Jump-Diffusion Model," *Applied Mathematics and Optimization*, 35, 145–167.

PINDYCK, R. S. (1988): "Irreversible Investment, Capacity Choice, and the Value of the Firm," *American Economic Review*, 78, 969–985.

RIEDEL, F. (2009): "Optimal Stopping with Multiple Priors," *Econometrica*.

RIEDEL, F., AND X. SU (2009): "On Irreversible Investment," *Finance and Stochastics*.

ROGERS, L., AND D. WILLIAMS (1987): *Diffusions, Markov Processes, and Martingales*, vol. 2 of *Series in Probability and Mathematical Statistics*. Wiley, Chichester.

SARGENT, T. (1987): *Dynamic Macroeconomic Theory*. Harvard University Press.

SKIADAS, C. (2003): "Robust Control and Recursive Utility," *Finance and Stochastics*, 7, 745–789.

Smit, H., and L. Trigeorgis (2004): *Strategic Investment: Real Options and Games*. Princeton University Press.

Trevino, E. (2008): "American Options in Incomplete Markets: Upper and Lower Snell Envelopes and Robust Partial Hedging," Ph.D. thesis, Humboldt University at Berlin.

Trigeorgis, L. (1996): *Real Options: Managerial Flexibility and Strategy in Resource Allocation*. MIT Press.

Weitzman, M. L. (1979): "Optimal Search for the Best Alternative," *Econometrica*, 47, 641–654.