Smooth Ambiguity Aversion Toward Small Risks and Continuous-Time Recursive Utility

Costis Skiadas*

August 20, 2008; updated October 7, 2008.

Abstract

In a continuous-time setting with Brownian and Poissonian uncertainty, this paper formulates recursive utility under two smooth certainty equivalent (CE) types that have been proposed as representations of ambiguity aversion. For a smooth CE based on the formulation of Klibanoff, Marinacci, and Mukerji (*Econometrica*, 2005), it is argued that the corresponding continuous-time recursive utility reduces to Kreps-Porteus utility (*Econometrica*, 1978), that is, recursive utility with an expected utility CE. For a smooth CE based on the divergence preferences of Maccheroni, Marinacci, and Rustichini (*Econometrica*, 2006), the following conclusions are drawn. Under only Brownian uncertainty, the corresponding continuous-time recursive utility again reduces to the Kreps-Porteus case. Under Poissonian uncertainty, the same conclusion can be drawn if and only if the divergence CE is of the entropic type. A non-entropic divergence CE results in a new class of continuous-time smooth recursive utilities that price Brownian and Poissonian risks differently.

*Kellogg School of Management, Department of Finance, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208. I am grateful to Lars Hansen, Peter Klibanoff, Jianjun Miao, Sujoy Mukerji, Dimitris Papanikolaou and Viktor Todorov for helpful discussions, and especially to Mark Schroder for his feedback. I'm responsible for any errors. The latest version of this article can be found at www.kellogg.nwu.edu/faculty/skiadas/home.htm

1 Introduction

Ambiguity aversion as commonly exemplified by the Ellsberg (1961) experiments is often associated with the non-smooth utility form of Gilboa and Schmeidler (1989). The presence of a kink in the utility may have interesting consequences[1] but is not conceptually a necessary implication of ambiguity aversion. More recently, two papers by Klibanoff, Marinacci, and Mukerji (2005) and Maccheroni, Marinacci, and Rustichini (2006a), henceforth abbreviated to KMM and MMR, respectively, have proposed decision-theoretic models of ambiguity aversion in which utilities can be smooth. There is a long-standing intuition, however, going back at least to Machina (1982), that a smooth certainty equivalent (CE) is likely to be well approximated by an expected utility CE when only small risks are involved. A main concern of this paper is whether this intuition applies to CEs representing KMM or divergence preferences. Divergence preferences are within the broader preference class axiomatized by MMR and generalize entropic preferences, which are prominent in the work of Hansen, Sargent and their coauthors (see, for example, Hansen Sargent 2001, 2007). It is well-known that entropic preferences have an expected utility representation. This paper argues that for small risks, a smooth KMM CE can be generally approximated by an expected utility CE, while the same is true of a smooth divergence CE if and only if the CE is entropic. For nonentropic divergence CEs, an interesting new class of small risk approximations arises.

We use the term "small risk" to refer to two types of risk: Brownian and Poissonian. Roughly speaking, a Brownian risk involves a high probability of a small change, while a Poissonian risk involves a small probability of a large change. Since the seminal contributions of Arrow (1965, 1970) and Pratt (1964), in decision theory the term small risk has mainly been understood in the Brownian sense, in which case the risk aversion of an expected utility CE $u^{-1}\mathbb{E}u$ is captured by the Arrow-Pratt coefficient of risk aversion $-u''/u'$. As a first step of our analysis, we will argue that for Poissonian small risks, the role of an Arrow-Pratt coefficient of risk aversion is assumed by the quantity $\Sigma - (u(U+\Sigma) - u(U))/u'(\Sigma)$, which represents a risk-aversion measure toward a small probability of a jump of size Σ relative to the reference certain outcome U. For a divergence CE, this Poissonian risk-aversion coefficient is modified to $\Sigma - \zeta(u(U+\Sigma) - u(U))/u'(\Sigma)$, where the function ζ is a simple transformation of the divergence index. Risk aversion of a divergence CE toward Brownian risk, on the other hand, is fully captured by an Arrow-Pratt coefficient that can be specified independently of ζ. The function ζ, therefore, modifies risk aversion with respect to Poissonian risk, without changing an agent's risk aversion toward Brownian risk, thus generating source-dependent risk aversion. These insights are fully captured by a simple single-period model, which will be our focus in the first part of this paper.

In applications, small risks arise naturally in settings such as well-functioning financial markets, where information arrives with high frequency over time. With this motivation, we will extend the arguments behind the single-period CE approximations to formulate continuous-time recursive

[1]For the role of such a kink in continuous-time recursive utility, see Chen and Epstein (2002), Epstein and Miao (2003), Schroder and Skiadas (2003, 2005) and Skiadas (2008).

utility with a KMM or divergence CE, under both Brownian and Poissonian sources of risk. The continuous-time model is a parsimonious expression of discrete-time small-risk approximations, with associated simplifications and gains in analytical tractability. While a rigorous version of the continuous-time theory can be quite technical, this paper's presentation is intended as an intuitive and largely heuristic explanation of the continuous-time model; only casual familiarity with stochastic analysis is assumed. The inclusion of Poissonian uncertainty reflects its increasing importance as an essential tool of economic modeling. For example, unpredictable jumps have for quite some time played an essential role in the modeling of financial time series and option pricing (see Cont and Tankov (2004) for an overview), while macroeconomic models involving small disaster probabilities have been receiving increasing attention in the work of Rietz (1988), Barro (2006), Gabaix (2008) and others. One of this paper's central contributions is to highlight the qualitative difference of risk aversion toward Brownian and Poissonian uncertainty under divergence preferences.

We will take as our benchmark the utility of Kreps and Porteus (1978), that is, recursive utility with an expected utility CE. The continuous-time version of Kreps-Porteus utility was formulated[2] by Duffie and Epstein (1992), whose analysis is extended in this paper by clarifying the risk-aversion structure with respect to Poissonian risk. Duffie and Epstein emphasized that their formulation can be thought of as the limiting case of a broader class of recursive utilities than Kreps-Porteus utility, since smooth CEs can be approximated by expected utility, with the approximation becoming exact in the continuous-time limit. We will argue that this is indeed the case for recursive utility with a smooth KMM CE, which is therefore indistinguishable from Kreps-Porteus utility in continuous-time under either Brownian or Poissonian risk. In contrast, continuous-time recursive utility with a smooth divergence CE does not always reduce to the Kreps-Porteus case, since for Poissonian risk a smooth divergence CE can be approximated by an expected utility CE in the entropic case only. For a nonentropic smooth divergence CE, recursive utility reduces to Kreps-Porteus utility in a Brownian filtration but forms a new utility class in the presence of Poissonian risk. This dichotomy reflects a form of source-dependent risk aversion that emerges as a natural consequence of the informational structure, in contrast to the discrete-time formulation of KMM and continuous-time formulation of Skiadas (2008), where source-dependent risk aversion is postulated directly.

Recursive utility with a nonentropic divergence CE is a promising tool in applications, leading to models of differential pricing of Brownian and Poissonian risk. The specification is formally within the abstract framework of Schroder and Skiadas (2008), who establish the form of the utility gradient

[2] The most widely used form of Kreps-Porteus utility is that of Epstein and Zin (1989), which is obtained by imposing homotheticity and a specific parameterization of the intertemporal aggregator. Existence, uniqueness and basic properties of the backward stochastic differential equation (BSDE) used to defined continuous-time Epstein-Zin utility in a Brownian filtration were established in Schroder and Skiadas (1999). The reason for the separate treatment is that the BSDE of Epstein-Zin utility violates the usual Lipschitz-growth conditions assumed by Pardoux and Peng (1990) and Duffie and Epstein (1992). Related BSDE results were also established by Kobylanski (2000), again assuming a Brownian filtration. To my knowledge, the corresponding theory has not yet been extended to include Poissonian uncertainty.

and optimality conditions for a general class of recursive preferences and informational structures with both Brownian and Poissonian risk. Schroder and Skiadas (2008) focus on a categorization of analytically tractable formulations based on translation- or scale-invariance assumptions. For example, properly normalized expected discounted exponential utility is quasilinear (with respect to an annuity cash flow), but its recursive representation under Poissonian uncertainty involves a risk-aversion penalty that, in addition to a quadratic term in common with the Brownian uncertainty case, includes cubic and higher order terms. Schroder and Skiadas (2008) make the point that if the higher than quadratic order terms are omitted, one can preserve quasilinearity and a quadratic structure, with notable tractability advantages. The present paper sheds some decision-theoretic light on the resulting utility form by showing that it is quasilinear recursive utility with a conditional CE that corresponds to what MMR call Gini preferences (quadratic divergence).

The rest of this paper is organized in three sections and an appendix with proofs. The first section introduces the CE approximations in a two-state model representing either Brownian or Poissonian uncertainty. The second section, introduces recursive utility on a discrete information tree as motivation for the continuous-time formulation, which is the topic of the third section.

2 Single-Period Small Risk Approximations

This section introduces the main insights behind this paper's certainty equivalent approximations in a simple single-period model. We begin with the definition of the three certainty equivalents we study: expected utility, Klibanoff-Marinacci-Mukerji, and divergence. We then approximate these certainty equivalents on a two-state model under two types of risk: Brownian and Poissonian.

2.1 Certainty equivalents

We postulate a finite state space Ω, with at least two elements, and a reference probability $P : 2^{\Omega} \to [0,1]$ that assigns a positive mass $P(\omega) = P(\{\omega\})$ to each state $\omega \in \Omega$. The expectation operator relative to P is denoted $\mathbb{E}$. Given any other probability Q on 2^{Ω}, the corresponding expectation operator is denoted $\mathbb{E}^{Q}$, and the *density* dQ/dP is defined as the random variable that takes the value $Q(\omega)/P(\omega)$ at state $\omega \in \Omega$. Therefore, $\mathbb{E}^{Q}x = \mathbb{E}[x dQ/dP]$ for every $x \in \mathbb{R}^{\Omega}$.

We fix, for the entire paper, a constant $\ell \in [-\infty, 1)$ serving as a lower bound on consumption, with typical values in applications being $\ell = -\infty$ or $\ell = 0$. A *certainty equivalent (CE)* is an increasing continuous function of the form $\upsilon : (\ell, \infty)^{\Omega} \to (\ell, \infty)$ with the property $\upsilon(\alpha \mathbf{1}) = \alpha$ for all $\alpha \in (\ell, \infty)$ (where $\mathbf{1}$ is the element of $(\ell, \infty)^{\Omega}$ identically equal to one).

In this paper, a *prior* is any probability Q that is equivalent to the reference probability P (meaning that Q assigns probability zero to the same events as P), which in our current setting means that Q assigns a positive mass to every state. This restriction on priors is imposed purely for expositional simplicity; it is in no way essential to the paper's arguments, and it plays no role in the CE definitions that follow. The set of all priors is denoted Π.

2.1.1 Expected utility certainty equivalent

For our purposes, a *von Neumann-Morgenstern (vNM) index* is any strictly increasing, continuous function of the form $u : (\ell, \infty) \to \infty$. Differentiability assumptions will be key in our approximations. For $n = 1, 2, \ldots$, we let

$$C^n_{\text{vNM}} = \text{ set of } n \text{ times continuously differentiable vNM indices.} \tag{1}$$

A pair of a prior Q and a vNM index u define the *expected utility CE* $\upsilon = u^{-1}\mathbb{E}^Q u$, given more explicitly as

$$\upsilon(U) = u^{-1}\left(\mathbb{E}^Q u(U)\right), \quad U \in (\ell, \infty)^\Omega.$$

Recall that for any $Q \in \Pi$, two vNM indices u and $\tilde{u}$ define the same expected utility CE with prior Q if and only if they are related by a positive affine transformation (meaning that $\tilde{u} = \alpha u + \beta$ for some $\alpha \in (0, \infty)$ and $\beta \in \mathbb{R}$).

2.1.2 Klibanoff-Marinacci-Mukerji certainty equivalent

The first extension of an expected-utility CE we consider corresponds to the utility form of Klibanoff, Marinacci, and Mukerji (2005), abbreviated to KMM. In the KMM formulation, we can think of the agent contemplating an expected utility CE with a given vNM index u, under each one of the possible priors

$$Q^1, \ldots, Q^S.$$

The agent is uncertain about which prior best represents reality. In this sense, there is now a new source of uncertainty, represented by the new state-space $\{1, \ldots, S\}$, on which we postulate a probability represented by the weights

$$\pi^1, \ldots, \pi^S \in (0, 1), \quad \text{where} \quad \sum\nolimits_{s=1}^S \pi^s = 1.$$

We refer to these weights as the agent's prior on priors. The *KMM CE* υ is defined in terms of a second vNM index φ by

$$\upsilon(U) = \varphi^{-1}\left(\sum_{s=1}^S \varphi\left(u^{-1}\mathbb{E}^{Q^s} u(U)\right)\pi^s\right), \quad U \in (\ell, \infty)^\Omega. \tag{2}$$

The KMM CE can be thought of as a representation of source-dependent risk aversion; risk aversion toward payoffs that are contingent on the state $s \in \{1, \ldots, S\}$ is determined by φ, while conditionally on the state s, risk-aversion toward payoffs that are contingent on the state $\omega \in \Omega$ are determined by u. If $u = \varphi$, then $\upsilon = u^{-1}\mathbb{E}^Q u$, where Q is the compound probability

$$Q = \sum_{s=1}^S Q^s \pi^s. \tag{3}$$

If $u \neq \varphi$, priors and the prior on these priors cannot be compounded.

2.1.3 Divergence certainty equivalent

The second extension of an expected utility CE we consider corresponds to divergence preferences, which are within the axiomatic setting of Maccheroni, Marinacci, and Rustichini (2006a), abbreviated to MMR. Let us use the term *divergence index* to mean any strictly convex twice differentiable function of the form $\phi : (0, \infty) \to \mathbb{R}$ such that

$$\phi(1) = \phi'(1) = 0, \quad \phi''(1) = 1 \quad \text{and} \quad \lim_{x \to \infty} \phi'(x) = \infty. \tag{4}$$

Analogously to (1), we introduce the notation

$$C^n_{\text{div}} = \text{ set of } n \text{ times continuously differentiable divergence indices.}$$

We define a (smooth) *divergence CE* in terms of a vNM index u, a positive scalar θ and a divergence index ϕ by letting

$$\upsilon(U) = \inf_{Q \in \Pi} u^{-1} \left(\mathbb{E}^Q u(U) + \theta \mathbb{E} \phi \left(\frac{dQ}{dP} \right) \right), \quad U \in (\ell, \infty)^{\Omega}. \tag{5}$$

Note that θ and ϕ enter the above definition only as the product $\theta\phi$, and therefore the normalization $\phi''(1) = 1$ in the definition of a divergence index is without loss of generality. The lower values $\theta\phi$ takes the more risk averse the CE υ becomes, since it assigns lower values to risky outcomes.

Example 1 (Entropic CE) *The divergence CE* (5) *is defined to be* entropic *if*

$$\phi(x) = x \log x - x + 1, \quad x \in (0, \infty).$$

A variational identity of Donsker and Varadhan (1975) (which is also a corollary of Proposition 8 of Section 3.4) shows that in this case

$$\upsilon = (\psi \circ u)^{-1} \mathbb{E}(\psi \circ u), \quad \textit{where} \quad \psi(u) = \theta \left(1 - \exp\left(-\frac{u}{\theta} \right) \right). \tag{6}$$

An entropic CE is therefore an expected utility[3] *CE; the divergence CE* (5) *modifies the expected utility CE* $u^{-1}\mathbb{E}u$ *by an exponential concavification of the vNM index* u*, thus increasing risk aversion. The exponential vNM index* ψ *in* (6) *has been normalized so that it approaches the identity function as* θ *approaches infinity, corresponding to the fact that if* $\theta = \infty$*, then the infimum in* (5) *is achieved for* $Q = P$*. We will show later (under a smoothness assumption) that the entropic CE is the only divergence CE that is also an expected utility CE, even in an approximate sense.*

[3] Strzalecki (2008) draws a behavioral distinction between entropic preferences and expected utility by enlarging the set of primitives to include distinct objective and ambiguous uncertainty.

2.2 Single-Period Small Risk Approximations

This section introduces the paper's main ideas with the simplest nontrivial instance of the uncertainty model, where the state space consists of only two states $\Omega = \{0, 1\}$. We use $h \in (0, 1)$ to parameterize the time length of the single period: preferences are expressed without knowledge of the state at time zero, and the state is revealed at time h. In a dynamic extension, this section's model would correspond to a single node of a binomial tree. Multiple stochastically independent binomial trees of this type, properly normalized, converge to the continuous-time model to follow, as the period length h goes to zero. We are therefore interested in approximations that are valid for small values of h. We will express such approximations in terms of the usual little-oh notation, writing $o(h)$ to represent some function $\varepsilon : (0, 1) \to \mathbb{R}$ such that $\varepsilon(h)/h \to 0$ as $h \to 0$. Every instance of little oh represents a potentially different function ε.

2.2.1 Risk source and payoff structure

We will introduce two uncertainty models, corresponding to Brownian and Poissonian risk, parameterized by the period length h. A model is specified by the state space $\Omega = \{0, 1\}$, which does not depend on h, and a pair of a probability $P : 2^{\{0,1\}} \to (0, 1)$ and a random variable $B : \{0, 1\} \to \mathbb{R}$ that can depend on h. We therefore think of the pair (P, B) as a whole family $\{(P_h, B_h) : h \in (0, \varepsilon)\}$ for a sufficiently small $\varepsilon > 0$, even though the dependence on h is notationally suppressed.

The pair (P, B) is a representation of a risk that forms the basic building block of a type of uncertainty. In the Brownian uncertainty case (P, B) can be thought of as a model of a bet on the toss of a fair coin, while in the Poissonian uncertainty case (P, B) can be thought of a bet on a light bulb burning out during the time-interval $(0, h)$. In either case, the model is normalized to satisfy

$$\mathbb{E}B = 0 \quad \text{and} \quad \mathbb{E}\left[B^2\right] = h. \tag{7}$$

Since there are only two states, any random variable x has the *canonical decomposition*

$$x = \mathbb{E}x + \sigma B, \quad \text{where} \quad \sigma = \frac{\mathbb{E}[Bx]}{\mathbb{E}[B^2]} = \frac{1}{h}\mathbb{E}[Bx].$$

We fix a reference random variable U_h, which we think of as a payoff realized at time h. In the dynamic setting to follow, U_h will represent a continuation value. We assume throughout that

$$U_h = U_0 + \mu h + \Sigma B, \tag{8}$$

where the scalars U_0, μ and Σ are given parameters, whose value does not change with h. (Adding a term $o(h)$ to the above expression for U_h does not affect the results.) Our objective is the computation of a first-order approximation of the certainty equivalent $\upsilon(U_h)$ under the three specifications of υ introduced in the last section. Toward this purpose, we introduce some notation and remarks relating to the role of priors.

A prior Q defines the scalar ρ^Q through the canonical decomposition

$$\frac{dQ}{dP} = 1 + \rho^Q B, \quad \text{where} \quad \rho^Q h = \mathbb{E}\left[B\frac{dQ}{dP}\right] = \mathbb{E}^Q B. \tag{9}$$

For fixed $Q \in \Pi$, ρ^Q generally depends on h. This is not how we calibrate the model, however. Instead, we fix a scalar ρ^Q that is common for all values of h, and we define, for every $h \in (0,1)$ such that $1 + \rho^Q B$ is positive, the probability

$$Q(F) = \mathbb{E}\left[\left(1 + \rho^Q B\right) 1_F\right], \quad \text{for every event } F. \tag{10}$$

Conditions (9) and (10) are clearly equivalent. As with the probability P, a *prior* Q should be thought of as whole family of priors $\{Q_h : h \in (0, \varepsilon)\}$ for sufficiently small $\varepsilon > 0$, such that the quantity $\rho^Q = \mathbb{E}^Q B/h$ is the same for all $h \in (0, \varepsilon)$.

Finally, given any $Q \in \Pi$, we define the notation

$$U_h = U_0 + \mu^Q h + \Sigma B^Q, \quad \mu^Q = \mu + \Sigma\rho^Q, \quad B^Q = B - \rho^Q h. \tag{11}$$

Note that $\mathbb{E}^Q B^Q = 0$ by construction. The second moment of B^Q under Q can be computed as

$$\mathbb{E}^Q\left[(B^Q)^2\right] = \mathbb{E}\left[\left(1 + \rho^Q B\right)\left(B - \rho^Q h\right)^2\right] = h - \left(\rho^Q h\right)^2 + \rho^Q \mathbb{E}\left[B^3\right].$$

In the Brownian model, $\mathbb{E}\left[B^3\right] = 0$ and therefore the second moment of B^Q under Q is approximately equal to the second moment of B under P. In the Poissonian model, $\mathbb{E}\left[B^3\right] = h + o(h)$ and the second moment of B^Q under Q reflects that fact that under the prior Q, the probability of the light bulb burning out during the interval $(0, h)$ is approximately $\left(1 + \rho^Q\right) h$.

2.2.2 Brownian risk

The basic building block of Brownian risk is

$$\begin{cases} \text{state 1:} & B(1) = \sqrt{h} & \text{with probability} & P(1) = 1/2, \\ \text{state 0:} & B(0) = -\sqrt{h} & \text{with probability} & P(0) = 1/2. \end{cases} \tag{12}$$

A random walk whose increments are i.i.d. copies of B converges to Brownian motion as h goes to zero. A rigorous version of this statement is a special case of Donsker's theorem, which can be found in Billingsley (1999) (Theorem 14.1). The formalism of the last subsection applies.

Given any scalar ρ^Q, a prior Q such that $\mathbb{E}^Q B = \rho^Q h$ is given explicitly as

$$Q(1) = \frac{1}{2}\left(1 + \rho^Q\sqrt{h}\right) \quad \text{and} \quad Q(0) = \frac{1}{2}\left(1 - \rho^Q\sqrt{h}\right).$$

Under Q, a properly normalized random walk whose increments are i.i.d. copies of $B^Q = B - \rho^Q h$ also converges to Brownian motion as h goes to zero.

The following proposition summarizes the implications of the Brownian risk assumption for smooth versions of the three CE specifications of section 2.1. These approximations are essentially the result of second-order Taylor series expansions using the moment conditions

$$\mathbb{E}^Q B^Q = 0, \quad \mathbb{E}^Q \left[(B^Q)^2 \right] = h + o(h), \quad \mathbb{E}^Q \left[(B^Q)^3 \right] = o(h). \tag{13}$$

We denote the (local) coefficient of absolute risk aversion of a vNM index u as

$$a^u = -\frac{u''}{u'}. \tag{14}$$

Proposition 2 (CE Approximations for Brownian Risk) *Assuming the risk specification* (12), *the following approximation are valid for any* $u \in C^3_{vNM}$.

(a) (Expected utility CE) Given the scalar ρ^Q, *if* $\mathbb{E}^Q B = \rho^Q h$ *for all sufficiently small* h, *then*

$$u^{-1}\mathbb{E}^Q u(U_h) = U_0 + \left(\mu^Q - \frac{1}{2} a^u (U_0) \Sigma^2 \right) h + o(h), \quad \mu^Q = \mu + \Sigma \rho^Q. \tag{15}$$

(b) (KMM CE) Suppose that υ *is the divergence CE defined in section 2.1.2 for some* $\varphi \in C^2_{vNM}$, *and that there exist scalars* $\rho^1, \ldots, \rho^S$ *such that* $\mathbb{E}^{Q^s} B = \rho^s h$ *for all sufficiently small* h. *Then*

$$\upsilon(U_h) = u^{-1}\mathbb{E}^Q u(U_h) + o(h), \quad \textit{where } Q = \sum\nolimits_s Q^s \pi^s.$$

(c) (Divergence CE) Suppose that υ *is the divergence CE defined in section 2.1.3, for some divergence index* $\phi \in C^3_{div}$ *and positive scalar* θ. *Then*

$$\upsilon(U_h) = (\psi \circ u)^{-1} \mathbb{E}(\psi \circ u)(U_h) + o(h) \quad \textit{where} \quad \psi(u) = \theta \left(1 - \exp\left(-\frac{u}{\theta} \right) \right). \tag{16}$$

Proof. See Appendix, A.1. ■

Part (a) of the proposition is the familiar Arrow-Pratt CE approximation applied to the payoff approximation (8) relative to the prior Q. (See Arrow (1965, 1970) and Pratt (1964)). By equation (11) and the fact that $\mathbb{E}^Q B^Q = 0$, approximation (15) can be restated as

$$u^{-1}\mathbb{E}^Q u(U_h) = \mathbb{E}^Q U_h - \frac{1}{2} a^u (U_0) \Sigma^2 h + o(h).$$

The term $\mathbb{E}^Q U_h$ represents a risk-neutral CE under the prior Q, and the quadratic term represents a risk-aversion adjustment to the risk-neutral CE, corresponding to the local curvature of the vNM index u near the reference level U_0.

Part (b) of the preceding proposition is essentially a corollary of the first part. Equation (15) applied to each Q^s implies that to first order, we can approximate the KMM CE by using a linear approximation of φ, in which case the terms φ^{-1} and φ cancel out. Since φ becomes irrelevant to this approximation, we can set it equal to u, allowing priors and the prior on priors to be compounded.

This shows that the KMM CE is, to first order, approximately equal to the corresponding expected utility CE obtained by setting $\varphi = u$ and compounding priors.

Part (c) of Proposition 2 can be viewed as a corollary of Example 1, stating that for an entropic CE, approximation (16) holds as an exact relationship (that is, without the $o(h)$ term). To see why, we use restrictions (4) on ϕ in a second-order Taylor expansion, to compute

$$\mathbb{E}\phi\left(\frac{dQ}{dP}\right) = \mathbb{E}\phi\left(1 + \rho^Q B\right) = \frac{1}{2}\left(\rho^Q\right)^2 h + o(h). \tag{17}$$

To first order, any smooth divergence index ϕ results in approximately the same divergence CE value $\upsilon(U_h)$, which must therefore be approximately equal to its value for the entropic case. The proof in the appendix gives an alternative constructive proof that does not rely on the result for the entropic case.

A comparison of the approximations of parts (a) and (c), and the fact that $a^{\psi \circ u} = a^u + \theta^{-1} u'$ indicate that approximation (16) can be alternatively stated as

$$\upsilon(U_h) = U_0 + \left(\mu - \frac{1}{2} a^{\psi \circ u}(U_0)\,\Sigma^2\right) h + o(h) = u^{-1}\mathbb{E}u(U_h) - \frac{u'(U_0)}{2\theta}\Sigma^2 + o(h). \tag{18}$$

We have already noted that decreasing $\theta\phi$ generally increases the CE's risk aversion. What is interesting here is that the divergence index ϕ does not affect risk aversion toward small Brownian risks (given the normalization $\phi''(1) = 1$). Decreasing θ increases risk aversion by penalizing the CE for exposure to the risk B more heavily through the last quadratic term in (18).

Finally, it is worth noting that Proposition 2(c) remains valid for what MMR call weighted divergence preferences, corresponding to the replacement of the term $\mathbb{E}\phi(dQ/dP)$ in the divergence CE definition by the term $\mathbb{E}^W\phi(dQ/dP)$ for a new prior W. This follows easily by verifying that approximation (17) remains valid if the expectation is computed relative to the new prior W.

2.2.3 Poissonian risk

The basic building block of Poissonian risk is

$$\begin{cases} \text{state 1:} \quad B(1) = 1 - h - \varepsilon(h) & \text{with probability} \quad P(1) = h + \varepsilon(h), \\ \text{state 0:} \quad B(0) = -h - \varepsilon(h) & \text{with probability} \quad P(0) = 1 - h - \varepsilon(h), \end{cases} \tag{19}$$

where $\varepsilon(h) = o(h)$. While any such choice of $\varepsilon(h)$ yields the same results, for concreteness, we set

$$\varepsilon(h) = \frac{1}{2} - h - \sqrt{\frac{1}{4} - h},$$

which implies that $\mathbb{E}\left[B^2\right] = h$, and therefore that the formalism of subsection 2.2.1 applies exactly. A random walk whose increments are i.i.d. copies of B converges to a compensated Poisson process with unit arrival rate as h goes to zero. A rigorous version of this statement can be found in

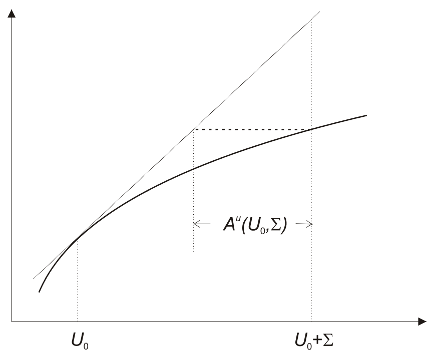

Figure 1: Graphical representation of $A^u(U_0, \Sigma)$. The curved solid line is part of the graph of u, while the slanted straight line is tangent to the graph at U_0.

Billingsley (1999) (Example 12.3). The assumption of a unit arrival rate can be thought of as the definition of the unit of time, and therefore entails no loss of generality.

Given any scalar ρ^Q, a prior Q such that $\mathbb{E}^Q B = \rho^Q h$ is given by

$$Q(1) = 1 - Q(0) = \lambda^Q h + \varepsilon(h), \quad \text{where} \quad \lambda^Q = 1 + \rho^Q. \tag{20}$$

The distribution of $B^Q = B - \rho^Q h$ under Q is

$$B^Q = \begin{cases} 1 - \lambda^Q h - \varepsilon(h), & \text{with } Q\text{-probability } \lambda^Q h + \varepsilon(h), \\ -\lambda^Q h - \varepsilon(h), & \text{with } Q\text{-probability } 1 - \lambda^Q h - \varepsilon(h). \end{cases}$$

Under Q, a random walk whose increments are i.i.d. copies of B^Q converges to a compensated Poisson process with arrival rate λ^Q as h goes to zero. The effect of a change of prior in this context is, therefore, approximately the same as a change in the unit of time.

Proposition 3 below summarizes the implications of Poissonian risk for the three CE specifications of section 2.1. The stated approximations are essentially the result of first-order Taylor series expansions, utilizing the moment conditions

$$\mathbb{E}^Q B^Q = 0 \quad \text{and} \quad \mathbb{E}^Q\left[(B^Q)^2\right] = \lambda^Q h + o(h). \tag{21}$$

In order to state the proposition, we first introduce several pieces of new notation.

In the context of Poissonian risk, the role of a coefficient of absolute risk aversion of a vNM index u is assumed by the quantity

$$A^u(U_0, \Sigma) = \Sigma - \frac{u(U_0 + \Sigma) - u(U_0)}{u'(U_0)}, \tag{22}$$

which is represented graphically in Figure 1. While $a^u(U_0)$ is a local measure of risk aversion toward risks taking values near U_0, $A^u(U_0, \Sigma)$ is a measure of risk aversion toward risks that take the value

$U_0 + \Sigma$ with a small probability, and the value U_0 otherwise. Both a^u and A^u are invariant to positive affine transformations of u. The inequality $A^u \geq 0$ is equivalent to the gradient inequality for u, and therefore concavity of u is equivalent to the nonnegativity of A^u. It is also worth noting parenthetically that if $u \in C^3_{\mathrm{vNM}}$, then

$$A^u (U_0, \Sigma) = \frac{1}{2} a^u (U_0) \Sigma^2 + o\left(\Sigma^2\right),$$

as can be seen by taking a second-order Taylor series expansion of u in (22). In our current context, however, Σ is not assumed to be small, and u need not have a third derivative.

In analyzing a divergence CE, we will focus on the case in which the minimization problem defining a divergence CE has an interior solution. We will see, as part of the proof of Proposition 3, that the latter condition is equivalent to membership of the reference (U_0, Σ) to the set

$$D = \left\{(U_0, \Sigma) \in (\ell, \infty) \times \mathbb{R} : U_0 + \Sigma > \ell \text{ and } u(U_0 + \Sigma) - u(U_0) < -\theta \phi'(0+)\right\}. \tag{23}$$

Given any function $\zeta : \left(-\infty, -\theta\phi'(0+)\right) \to \mathbb{R}$, we further define the notation

$$A^u_\zeta (U_0, \Sigma) = \Sigma - \frac{\zeta\left(u(U_0 + \Sigma) - u(U_0)\right)}{u'(U_0)}, \tag{24}$$

which extends the notation in (22), since $A^u = A^u_{\mathrm{identity}}$. In fact, the last identity can be viewed as a limiting case (for $\theta = \infty$) of the more interesting identity

$$A^{\psi \circ u} = A^u_\psi, \quad \text{where} \quad \psi(u) = \theta\left(1 - \exp\left(-\frac{u}{\theta}\right)\right).$$

Finally, we use the convex conjugate notation (see Rockafellar (1970))

$$\phi^*(x^*) = \sup_{x \in (0,\infty)} \left\{x^* x - \phi(x)\right\}. \tag{25}$$

Proposition 3 (CE Approximations for Poissonian Risk) *Assuming the risk specification* (19), *the following approximation are valid for any* $u \in C^2_{vNM}$.

(a) (Expected utility CE) Given the scalar ρ^Q, *if* $\mathbb{E}^Q B = \rho^Q h$ *for all sufficiently small* h, *then*

$$u^{-1}\mathbb{E}^Q u(U) = U_0 + \left(\mu^Q - A^u(U_0, \Sigma)\lambda^Q\right) h + o(h), \quad \mu^Q = \mu + \Sigma\rho^Q, \quad \lambda^Q = 1 + \rho^Q.$$

(b) (KMM CE) The statement of part (b) of Proposition 2 is valid in the current context, too.

(c) (Divergence CE) Suppose that υ *is the divergence CE defined in section 2.1.3, for some divergence index* $\phi \in C^2_{div}$ *and positive scalar* θ. *If* $(U_0, \Sigma) \in D$, *then*

$$\upsilon(U_h) = U_0 + \left(\mu - A^u_\zeta(U_0, \Sigma)\right) h + o(h), \quad \text{where} \quad \zeta(\delta) = -\theta\phi^*\left(-\frac{\delta}{\theta}\right). \tag{26}$$

Finally, there exists some prior W *and* $w \in C^2_{vNM}$ *such that*

$$\upsilon(U_h) = w^{-1}\mathbb{E}^W w(U_h) + o(h) \quad \text{for all } (U_0, \Sigma) \in D \tag{27}$$

if and only if $W = P$ *and* $\phi(x) = x\log x - x + 1$, *in which case identity* (6) *holds and* $\zeta = \psi$.

Proof. See Appendix, A.2. ■

Remark 4 *Expression* (26) *is specific to the normalization of the reference model* (P, B)*, corresponding to a unit arrival rate in the Poisson limit. The functional form with a reference model corresponding to an arbitrary arrival rate* $\lambda \in (0, \infty)$ *can be obtained by a change of the time unit, resulting in the CE approximation*

$$\upsilon(U_h) = U_0 + \left(\mu - A^u_\zeta(U_0, \Sigma)\lambda\right)h + o(h).$$

In the continuous-time extension, the presence of multiple Poissonian sources of risk means that a single change of time units cannot normalize all arrival rates at once. Instead, we achieve such a normalization by a change of the underlying probability measure.

The expected utility CE approximation of part (a) takes the same form (15) of the Brownian risk case, but with the risk-aversion penalty $a^u(U_0)\Sigma^2/2$ replaced with $A^u(U_0, \Sigma)\lambda^Q$. Given part (a), the argument that gives part (b) is essentially the same as in the Brownian case. A smooth KMM CE is therefore approximately an expected utility CE given a small risk, whether the risk is Brownian or Poissonian.

Part (c) of Proposition 3 gives this paper's first instance of a CE approximation that is *not* consistent with expected utility. In Example 1, we saw that an entropic divergence CE is an expected utility CE. Part (c) of the above proposition gives a strong converse: if a smooth divergence CE can be approximated by a smooth expected utility CE, then it must be entropic.

Unlike the Brownian case, both θ and ϕ affect the risk aversion of the divergence CE toward Poissonian risks. We have already noted that the lower values $\theta\phi$ takes the more risk averse the CE υ is. Consistent with this observation, the risk-aversion penalty $A^u_\zeta(U_0, \Sigma)h$ is higher the lower values $\theta\phi$ takes, a fact that is immediate clear from the definition of A^u_ζ and the observation that the function ζ defined in (26) is also given by

$$\zeta(\delta) = \inf_{x \in (0,\infty)} \{\theta\phi(x) + \delta x\}, \qquad \delta \in \left(-\infty, -\theta\phi'(0+)\right).$$

The different ways in which the parameters $\theta\phi$ affect risk aversion toward Brownian and Poissonian risks is more interesting within a model that allows both types of risk, resulting in source-dependent risk aversion. We illustrate the basic idea with a minimal model with this feature. Let (P_1, B_1) be the Brownian risk model of the last subsection, let (P_2, B_2) be the Poissonian risk model of the current subsection, and consider the product model $(P_1 \times P_2, B)$, where $B = (B_1, B_2)'$, defined on the product state space $\Omega = \{0, 1\} \times \{0, 1\}$. The formalism of section 2.2.1 has a straightforward extension to this setting. Let υ be the divergence CE of section 2.1.3, for some divergence index $\phi \in C^2_{\text{div}}$ and positive scalar θ, and consider the risk $U_h = U_0 + \mu h + \Sigma' B$, where the parameters $\mu \in \mathbb{R}$ and $\Sigma \in \mathbb{R}^2$ do not depend on h. Since Taylor series approximations are additive, we have the approximation

$$\upsilon(U_h) = U_0 + \left(\mu - \frac{1}{2}a^{\psi \circ u}(U_0)(\Sigma_1)^2 - A^u_\zeta(U_0, \Sigma_2)\right)h + o(h).$$

The CE value $\upsilon(U_h)$ is to first-order approximately equal to the risk-neutral CE $\mathbb{E}U_h = U_0 + \mu h$ minus two terms reflecting risk adjustment toward Brownian and Poissonian risk, respectively. For given divergence index ϕ, decreasing θ increases both $a^{\psi\circ u}$ and A_ζ^u, and hence increases risk aversion toward both types of risk. For fixed θ, on the other hand, decreasing the values that the divergence index ϕ takes does not change $a^{\psi\circ u}$ while increasing A_ζ^u, thus increasing risk aversion toward Poissonian risk alone.

3 Discrete Recursive Utility

In order to motivate and clarify the continuous-time formulation to follow, in this section we give rigorous discrete-time definitions of the three dynamic utility classes that are this paper's focus: Kreps-Porteus utility, recursive utility with a KMM conditional CE, and finally recursive utility with a divergence CE. In the context of the latter, we solve for the minimizing prior, in a result that contrasts with the simplifications resulting from the assumption of Brownian and Poissonian uncertainty in the continuous-time formulation.

Although we will not introduce axioms here, the ordinal axiomatic basis for the utility forms we consider is essentially already available in the literature. Kreps and Porteus (1978) emphasized the role of preferences for the timing of resolution of uncertainty, an issue that is cast more broadly in Skiadas (1998) as preferences over pairs of information trees (filtrations) and consumption plans. In our context, the information tree is given and fixed, and as a consequence the axioms that lead to recursive utility with an arbitrary conditional CE are quite simple, and can be found in Chapter 6 of Skiadas (forthcoming). The key ingredients are dynamic consistency, the irrelevance of past or unrealized consumption, and the irrelevance of current consumption for risk aversion toward one-period-ahead uncertainty. The additional axioms required to specify the conditional CE can be established by embedding the corresponding single-period theory, such as that of KMM or MMR, in the setting of the information tree. This is achieved by identifying, at each nonterminal node, a payoff of the single-period theory with a consumption plan whose value at each immediate successor node remains constant over time.

In addition to the above cited papers, the foundations of various forms of recursive utility have been discussed in Selden (1978), Johnsen and Donaldson (1985), Chew and Epstein (1989), Epstein and Zin (1989), Wang (2003), Hayashi (2005), and Klibanoff and Ozdenoren (2007). Finally, Klibanoff, Marinacci, and Mukerji (2007) and Maccheroni, Marinacci, and Rustichini (2006b) have formulated dynamic utilities extending, respectively, KMM and MMR single-period uncertainty, but with more restrictive assumptions on intertemporal choice than we will impose here.

3.1 Stochastic Setting and Recursive Utility

We consider a finite state space Ω, and a time-set $\{0, h, 2h, \ldots, Nh\}$, for some $h \in (0,1)$, and positive integer N. The terminal time is $T = Nh$. Information is represented by the filtration $\{\mathcal{F}_t : t = 0, h, \ldots, Nh\}$, where $\mathcal{F}_0 = \{\Omega, \emptyset\}$ and $\mathcal{F}_T$ is the set of all subsets of Ω. $L(\mathcal{F}_t)$ denotes the set of every $\mathcal{F}_t$-measurable random variable. For every time t, the algebra $\mathcal{F}_t$ is generated by a partition of Ω that we denote $\mathcal{F}_t^0$. A time-t (informational) *spot* is a pair of the form (F, t), where $F \in \mathcal{F}_t^0$. One can arrange these spots as nodes of an information tree with the obvious succession rule. A (stochastic) process $x : \Omega \times \{0, h, \ldots, Nh\} \to \mathbb{R}$ is *adapted* if $x_t = x(\cdot, t) \in L(\mathcal{F}_t)$ for every time t. If x is an adapted process and (F, t) is any spot, then $x(F, t)$ denotes the common value of the random variable x_t over the event F; that is, $x(F, t) = x(\omega, t)$ for every $\omega \in F \in \mathcal{F}_t^0$. We write $\mathbf{1}$ for a random variable or process that is identically equal to one, the meaning being clear from the context.

Consumption is assumed to be valued in the interval (ℓ, ∞), for some fixed $\ell \in [-\infty, 1)$. A *consumption plan* is any (ℓ, ∞)-valued adapted process c, with $c(F, t)$ representing a given agent's contingent consumption at spot (F, t). The set of all consumption plans is denoted C. For our purposes, a *utility function* is any strictly increasing and continuous function of the form $U_0 : C \to \mathbb{R}$. The utility function U_0 is *normalized* if $U_0(s\mathbf{1}) = s$ for every $s \in (\ell, \infty)$. For any utility function $\tilde{U}_0 : C \to \mathbb{R}$, the utility function

$$U_0 = f^{-1} \circ \tilde{U}_0, \quad \text{where} \quad f(s) = \tilde{U}_0(s\mathbf{1}),$$

is normalized and ordinally equivalent to $\tilde{U}_0$. We will work with normalized utilities throughout this paper.

A *conditional certainty equivalent (CE)* is a mapping υ that assigns to every time $t < T$ an increasing continuous function of the form $\upsilon_t : L(\mathcal{F}_{t+h}) \to L(\mathcal{F}_t)$ such that

- For every $x, y \in L(\mathcal{F}_{t+h})$ and $F \in \mathcal{F}_t^0$, $x1_F = y1_F$ implies $\upsilon_t(x) 1_F = \upsilon_t(y) 1_F$.
- For every $s \in (\ell, \infty)$, $\upsilon_t(s\mathbf{1}) = s$.

A (normalized, intertemporal) *aggregator* is a function of the form[4] $\Phi : (0, \infty) \times (\ell, \infty) \times (\ell, \infty) \to \mathbb{R}$ such that for every $h \in (0, \infty)$, the section $\Phi(h, \cdot, \cdot)$ is an increasing and continuous function satisfying $\Phi(h, 1, 1) = 1$. The first argument of Φ will represent the period length (and is fixed for the given time set), the second argument represents consumption, and the third a certainty equivalent of next period's utility.

The normalized utility function $U_0 : C \to \mathbb{R}$ is *recursive utility* if there is a conditional CE υ and an aggregator Φ such that for any $c \in C$, $U_0(c)$ is the initial value of the adapted process $U(c)$

[4] The entire analysis that follows applies with a spot-dependent aggregator Φ. One can simply reinterpret every occurence of Φ as depending on the state and time arguments (ω, t). Of course, Φ has to be adapted, meaning that $\Phi(\omega, t, \cdot) = \Phi(\omega', t, \cdot)$ if $\omega, \omega' \in F \in \mathcal{F}_t^0$.

specified by the backward (in time) recursion

$$U_t(c) = \Phi(h, c_t, \upsilon_t(U_{t+h}(c))), \quad t < T; \quad U_T(c) = c_T. \tag{28}$$

Any pair of a conditional CE υ and aggregator Φ defines a recursive utility by the above recursion. We refer to $U(c)$ as the *utility process* of the consumption plan c.

For notational convenience, we fix a reference probability P on $\mathcal{F}$ that assigns a positive value to every nonempty event. The corresponding expectation operator is denoted $\mathbb{E}$, while $\mathbb{E}_t$ is short for the conditional expectation operator $\mathbb{E}[\cdot \mid \mathcal{F}_t]$. A *prior* is any probability Q on $\mathcal{F}$ that, like P, assigns zero probability only to the empty set. The corresponding expectation operator is denoted $\mathbb{E}^Q$, and $\mathbb{E}_t^Q$ is an abbreviation of $\mathbb{E}^Q[\cdot \mid \mathcal{F}_t]$. As before, the set of all priors is denoted Π.

3.2 Kreps-Porteus Utility

An *expected-utility (EU) conditional CE* is a conditional CE of the form

$$\upsilon_t = u^{-1}\mathbb{E}_t^Q u, \tag{29}$$

where Q is a prior and u is a vNM index (defined as in section 2.1.1). *Kreps-Porteus utility* is recursive utility with an EU conditional CE. The corresponding utility recursion (28) is, therefore, of the form

$$U_t(c) = \Phi\left(h, c_t, u^{-1}\mathbb{E}_t^Q u(U_{t+1}(c))\right), \quad U_T(c) = c_T. \tag{30}$$

Example 5 (Expected discounted utility) *Consider the non-normalized utility function*

$$\tilde{U}_0(c) = \mathbb{E}^Q\left[\sum_{n=0}^{N-1} e^{-\beta nh} u(c_{nh}) + e^{-\beta Nh}\frac{u(c_T)}{1-e^{-\beta h}}\right], \quad c \in C, \tag{31}$$

where $Q \in \Pi$, $\beta \in (0,\infty)$ *and* u *is a vNM index. The value* $\tilde{U}_0(c)$ *can be computed as the initial value of the process* $\tilde{U}(c)$ *that solves the recursion*

$$\tilde{U}_t(c) = u_t(c) + e^{-\beta h}\mathbb{E}_t^Q \tilde{U}_{t+h}(c), \quad \tilde{U}(c_T) = \frac{u(c_T)}{1-e^{-\beta h}}.$$

The normalized utility process

$$U_t(c) = u^{-1}\left(\left(1-e^{-\beta h}\right)\tilde{U}_t(c)\right) \tag{32}$$

solves recursion (30) *with aggregator*

$$\Phi(h, c, \upsilon) = u^{-1}\left(\left(1-e^{-\beta h}\right)u(c) + e^{-\beta h}u(\upsilon)\right). \tag{33}$$

3.3 Klibanoff-Marinacci-Mukerji Conditional CE

To define a conditional version of the KMM CE introduced in section 2.1.2, we again take as primitive the priors $Q^1, \ldots, Q^S$ and corresponding weights

$$\pi_0^1, \ldots, \pi_0^S \in (0,1), \quad \text{where} \quad \sum_{s=1}^{S} \pi_0^s = 1.$$

We interpret π_0^s as the agent's time-zero prior probability mass placed on the prior Q^s. For each $s \in \{1, \ldots, S\}$, we define the process

$$\pi_t^s = \frac{\xi_t^s \pi_0^s}{\sum_{r=1}^{S} \xi_t^r \pi_0^r}, \quad \text{where} \quad \xi_t^s = \mathbb{E}_t \left[\frac{dQ^s}{dP} \right]. \tag{34}$$

For any spot (F, t), it can easily be verified that $\xi^s(F,t) = Q^s(F)/P(F)$ and therefore $\pi^s(F,t)$ is the time-t Bayesian update of the prior π_0 given the realization of the event F. Finally, we define a *KMM conditional CE* in terms of the vNM indices u and φ by

$$\upsilon_t(\cdot) = \varphi^{-1}\left(\sum_{s=1}^{S} \varphi\left(u^{-1} \mathbb{E}_t^{Q^s} u(\cdot) \right) \pi_t^s \right). \tag{35}$$

As in the single-period case,

$$\varphi = u \quad \text{implies} \quad \upsilon_t = u^{-1} \mathbb{E}_t^{Q} u, \quad \text{where} \quad Q = \sum_{s=1}^{S} Q^s \pi_0^s.$$

In the following example, we note that a finite-horizon version of the dynamic utility of Klibanoff, Marinacci, and Mukerji (2007) is recursive utility with a KMM conditional CE and an intertemporal aggregator that takes the same form (33) as for normalized expected discounted utility.

Example 6 (Klibanoff-Marinacci-Mukerji Recursive Utility) *Suppose that the non-normalized utility function $\tilde{U}_0 : C \to \mathbb{R}$ is defined by the recursion*

$$\tilde{U}_t(c) = u(c_t) + e^{-\beta h} \phi^{-1}\left(\sum_{s=1}^{S} \phi\left(\mathbb{E}_t^{Q^s} \tilde{U}_{t+h}(c) \right) \pi_t^s \right), \quad \tilde{U}(c_T) = \frac{u(c_T)}{1 - e^{-\beta h}}.$$

The normalized version of the above utility is recursive utility with the KMM conditional CE (35)*, where*

$$\varphi(\cdot) = \phi\left(\frac{u(\cdot)}{1 - e^{-\beta h}} \right),$$

and the same aggregator (33)*, as for the time-additive expected utility of Example 5. Just as in the latter, the corresponding normalized utility process $U_t(c)$ is given by* (32)*, and is easily verified to satisfy the claimed recursion.*

3.4 Divergence Conditional CE

As in the single-period case, we specify a divergence conditional CE in terms of a vNM index u, a divergence index $\phi : (0, \infty) \to \mathbb{R}$, as defined in section 2.1.3, and a positive scalar θ. For each prior Q, we define the conditional CE

$$\upsilon_t^Q(\cdot) = u^{-1}\left(\mathbb{E}_t^Q u(\cdot) + \theta \mathbb{E}_t \phi\left(\frac{\xi_{t+h}^Q}{\xi_t^Q}\right)\right), \quad \text{where} \quad \xi_t^Q = \mathbb{E}_t\left[\frac{dQ}{dP}\right].$$

Finally, we define the *divergence conditional CE* by

$$\upsilon_t(\cdot) = \inf_{Q \in \Pi} \upsilon_t^Q(\cdot).$$

Remark 7 *Given any $F \in \mathcal{F}_t^0$, let $F_1, \ldots, F_n$ be the elements of $\mathcal{F}_{t+h}^0$ whose union is F. Letting $p_i = P(F_i \mid F)$ and $q_i = Q(F_i \mid F)$, we have*

$$\mathbb{E}\left[\phi\left(\frac{\xi_{t+h}^Q}{\xi_t^Q}\right) \mid F\right] = \sum_{i=1}^{n} \phi\left(\frac{q_i}{p_i}\right) p_i.$$

Given a recursive utility with a divergence conditional CE, we now single out a reference consumption plan c and we let $U = U(c)$ denote the corresponding utility process. In order to appreciate the simplifications resulting from the continuous-time formulation to follow, we derive below the discrete-model solution to the minimization problem defining the conditional CE, under a condition that guarantees that the minimum is achieved. For the Proposition's statement, we define the function

$$G(x, y) = \phi'^{-1}\left(-\frac{1}{\theta}(u(x) + y)\right), \quad x \in (\ell, \infty), \quad y < -\theta\phi'(0+) - u(x),$$

and we let U_t^* denote the $\mathcal{F}_t$-measurable random variable that equals the maximum value of U_{t+h} given time-t information; that is,

$$U_t^*(\omega) = \max\{U_{t+h}(\omega) : \omega \in F\}, \quad \omega \in F \in \mathcal{F}_t^0.$$

Proposition 8 *Suppose that*

$$\mathbb{E}_t\left[\phi'^{-1}\left(\phi'(0+) + \frac{1}{\theta}(u(U_t^*) - u(U_{t+h}))\right)\right] < 1, \quad t < T - h. \tag{36}$$

For any time $t < T - h$, there exists a unique $\lambda_t \in L(\mathcal{F}_t)$ such that

$$\mathbb{E}_t[G(U_{t+h}, \lambda_t)] = 1 \quad \text{and} \quad \lambda_t < -\theta\phi'(0+) - u(U_{t+h}). \tag{37}$$

Given such a λ_t, let the positive martingale ξ be defined recursively by

$$\xi_{t+h} = \xi_t G(U_{t+h}, \lambda_t), \quad \xi_0 = 1, \tag{38}$$

and let Q be the unique probability such that $\xi^Q = \xi$ (given by $Q(F) = \mathbb{E}[\xi_T 1_F]$ for every event F). Then $\upsilon_t(U_{t+h}) = \upsilon_t^Q(U_{t+h})$.

Proof. See Appendix. ■

Example 9 (Gini conditional CE) *This example is covered by Theorem 24 of MMR. We use it to illustrate the application of the last proposition, as well as an interesting class of divergence recursive utilities that is not within the Kreps-Porteus class (a fact that, as we will see in the following section, remains true in the continuous-time formulation).*

We assume that

$$\phi(x) = \frac{1}{2}(x-1)^2,$$

and therefore

$$\upsilon_t = \inf_{Q \in \Pi} u^{-1}\left(\mathbb{E}_t^Q u(\cdot) + \frac{\theta}{2} Var_t\left[\frac{\xi_{t+h}^Q}{\xi_t^Q}\right]\right), \tag{39}$$

where Var_t denotes the conditional variance operator under P given time-t information. In this case, $G(x,y) = 1 - \theta^{-1}(u(x) + y)$. We assume the validity of condition (36)*, which in the current context is the same as*

$$u(U_{t+h}) - \mathbb{E}_t u(U_{t+h}) < \theta. \tag{40}$$

Equation (37) *results in $\lambda_t = -\mathbb{E}_t[u(U_{t+h})]$, and therefore*

$$\frac{\xi_{t+h}^Q}{\xi_t^Q} = G(U_{t+h}, \lambda_t) = 1 - \frac{1}{\theta}\left(u(U_{t+h}) - \mathbb{E}_t[u(U_{t+h})]\right).$$

Calculating the minimum using the above expression results in the conditional CE expression

$$\upsilon_t(U_{t+h}) = u^{-1}\left(\mathbb{E}_t u(U_{t+h}) - \frac{1}{2\theta} Var_t[u(U_{t+h})]\right).$$

Condition (40) *guarantees that U_{t+h} is valued within the range where the above expression defines a strictly increasing function. Consistent with our conclusion in Proposition 3(c), the above mean-variance specification is not consistent with expected utility.*

Another important special case is the conditional version of an entropic CE, introduced in Example 1. The formulation is summarized in the following example, where the Donsker-Varadhan variational identity is derived as a corollary of the above proposition.

Example 10 (Entropic conditional CE) *Suppose that $\phi(x) = x\log x - x + 1$ and therefore*

$$\upsilon_t = \inf_{Q \in \Pi} u^{-1}\left(\mathbb{E}_t^Q\left[u(\cdot) + \theta \log\left(\frac{\xi_{t+h}^Q}{\xi_t^Q}\right)\right]\right). \tag{41}$$

Condition (36) *is always satisfied in this case. Applying Proposition 8, we find that the conditional density process of the minimizing probability Q satisfies*

$$\frac{\xi_{t+h}^Q}{\xi_t^Q} = \frac{\exp\left(-\theta^{-1} u(U_{t+h})\right)}{\mathbb{E}_t \exp\left(-\theta^{-1} u(U_{t+h})\right)}.$$

Calculating the minimum using the above expression results in the conditional CE expression

$$\upsilon_t = (\psi \circ u)^{-1} \mathbb{E}_t (\psi \circ u), \quad \textit{where} \quad \psi(u) = \theta \left(1 - \exp\left(-\frac{1}{\theta}u\right)\right).$$

Recursive utility with an entropic conditional CE is therefore within the Kreps-Porteus class. Proposition 3(c) implies that any smooth non-entropic divergence conditional CE takes us outside the Kreps-Porteus class.

Finally, we show how a finite-horizon discrete-time version of a utility that appears in the work of Hansen, Sargent and their coauthors can be embedded in the above setting, after normalization. (See Hansen and Sargent (2001), Hansen, Sargent, Turmuhambetova, and Williams (2006), as well as Skiadas (2003).)

Example 11 (Hansen-Sargent utility form) *We consider the non-normalized utility specification*

$$\tilde{U}_0(c) = \inf_{Q \in \Pi} \mathbb{E}^Q \left[\sum_{n=0}^{N-1} e^{-\beta n h} \left(u(c_{nh}) + \theta \log \xi^Q_{nh} \right) + \frac{e^{-\beta T}}{1 - e^{-\beta h}} \left(u(c_T) + \theta \log \left(\xi^Q_T \right) \right) \right].$$

$\tilde{U}_0(c)$ *is the initial value of the process* $\tilde{U}(c)$ *computed recursively, starting with the terminal value*

$$\tilde{U}_T(c) = \frac{u(c_T)}{1 - e^{-\beta h}},$$

and following the backward recursion

$$\tilde{U}_t(c) + \frac{\theta}{1 - e^{-\beta h}} \log \xi^Q_t = \inf_Q \mathbb{E}^Q_t \left[u(c_t) + \theta \log \xi^Q_t + e^{-\beta h} \left(\tilde{U}_{t+h}(c) + \frac{\theta}{1 - e^{-\beta h}} \log \left(\xi^Q_{t+h} \right) \right) \right],$$

which can be simplified to

$$\tilde{U}_t(c) = u(c_t) + e^{-\beta h} \inf_Q \mathbb{E}^Q_t \left(\tilde{U}_{t+h} + \frac{\theta}{1 - e^{-\beta h}} \log \left(\frac{\xi^Q_{t+h}}{\xi^Q_t} \right) \right).$$

Therefore, the normalized version $U(c)$, *which is defined by* (32), *solves the recursion*

$$U_t(c) = u^{-1} \left[\left(1 - e^{-\beta h}\right) u(c_t) + e^{-\beta h} u \left(\upsilon_t \left(U_{t+h}(c) \right) \right) \right], \quad U_T(c) = c_T,$$

where the the conditional CE υ *is defined in* (41). *Given this fact, the argument of the last example shows that the Hansen-Sargent utility form is within the Kreps-Porteus class. The utility is not generally time-additive, even though the corresponding intertemporal aggregator takes the same form* (33) *as for the expected discounted utility of Example 5.*

4 Continuous-Time Formulation

We are finally in a position to superimpose the single-period small-risk CE approximations of Section 2.1 onto the recursive structure of the last section. Rather than stating the results as approximations, we consider the limiting case in a continuous-time model. Intuitively, we think of an information tree, where at each time-t spot the conditional uncertainty resolved by time $t + dt$ is linearly spanned by the stochastically independent factors $dB_t^1, \ldots, dB_t^d$, representing infinitesimal risks, some of which are Brownian and some Poissonian. The approximations of Section 2.1, which become exact relationships in the limit, can be applied with $h = dt$ and $B = dB_t^i$ for each factor i to simplify the form of the conditional CE over the infinitesimal time interval $[t, t + dt]$. Based on this intuition, we will formulate continuous-time versions of last section's recursive utilities in the form of backward stochastic differential equations (BSDEs), which are well-studied mathematical objects that can serve as a starting point for rigorous analysis and applications directly in continuous time.

4.1 Stochastic Setting

This section contains an informal and fairly self-contained explanation of the stochastic setting. The rigorous theory is contained in Jacod and Shiryaev (2003). An excellent introduction to multivariate point processes is given by Brémaud (1981).

We take as given an underlying probability space $(\Omega, \mathcal{F}, P)$, with corresponding expectation operator $\mathbb{E}$. On this space are defined d mutually stochastically independent processes, forming the column vector

$$B = \left(B^1, \ldots, B^k, B^{k+1}, \ldots, B^d\right)',$$

where

- B^i is a standard Brownian motion for $i = 1, \ldots, k$, and
- B^i is a compensated Poisson process with unit arrival rate for $i = k + 1, \ldots, d$.

The last statement means that there exist independent Poisson processes $N^{k+1}, \ldots, N^d$ such that $\mathbb{E}N_t^i = t$ and $B_t^i = N_t^i - t$ for each $i \in \{k + 1, \ldots, d\}$. We note that the symbol B does not stand for Brownian motion, as is often the case, but rather for *basis.*

The underlying filtration $\{\mathcal{F}_t : t \in [0, T]\}$ is defined as the smallest filtration such that B_t is $\mathcal{F}_t$-measurable and $\mathcal{F}_t$ contains all P-null events for every time t. Intuitively, at time t the agent observes the realization of B up to time t, which defines the counterpart of a "spot" in last section's finite model. $\mathbb{E}_t$ denotes the conditional expectation operator given time-t information $\mathcal{F}_t$. We assume that $\mathcal{F} = \mathcal{F}_T$ and therefore $\mathbb{E}_T[x] = x$ for every random variable x. A process X is *adapted* if X_t is $\mathcal{F}_t$-measurable for every time t. We will not enter into the technical definition of a *predictable* process X, but we think of the concept heuristically as the condition that X_t is $\mathcal{F}_{t-}$-measurable. (In the discrete-time model, predictability of X means that X_t is $\mathcal{F}_{t-h}$-measurable, but the notion is more subtle in the continuous-time limit.) For any process X whose paths have left limits, we use the heuristic notation $dX_t = X_{t+dt} - X_{t-}$, where X_{t-} denotes the left limit of X at t.

For our purposes, we define a *volatility process* to be any d-dimensional predictable process σ such that $\int_0^T \sigma_t' \sigma_t dt < \infty$ with probability one. A process M in this context is a locally square-integrable martingale if and only if there exists some volatility process σ such that

$$M_t = M_0 + \int_0^t \sigma_u' dB_u, \quad \text{or equivalently} \quad dM_t = \sigma_t' dB_t. \tag{42}$$

If $M_t = \int_0^t \alpha_u' dB_u$ and $N_t = \int_0^t \beta_u' dB_u$, for volatility processes α and β, then the conditional covariation process of M and N is well-defined and is given by $\int_0^t \alpha_u' \beta_u du$. We summarize this statement with the heuristic notation

$$\mathbb{E}_{t-}\left[(\alpha_t' dB_t)\left(\beta_t' dB_t\right)\right] = \alpha_t' \beta_t dt.$$

In particular, representation (42) implies that $\sigma_t dt = \mathbb{E}_{t-}\left[dM_t dB_t\right]$. We therefore think of $\sigma_t^i dt$ as the time-t conditional factor loading of dM_t on factor dB_t^i.

As before, we use the term *prior* to refer to any probability on $\mathcal{F}$ that is equivalent to P (meaning that it defines the same null events as P). Given any prior Q, we let $\mathbb{E}^Q$ denote the corresponding expectation operator, and we define the martingale ξ^Q, predictable process ρ^Q and adjusted basis process B^Q by[5]

$$\xi_t^Q = \mathbb{E}_t\left[\frac{dQ}{dP}\right], \quad \frac{d\xi_t^Q}{\xi_{t-}^Q} = \rho_t^{Q\prime} dB_t \quad \text{and} \quad B_t^Q = B_t - \int_0^t \rho_t^Q dt. \tag{43}$$

We let Π denote the set of every prior Q such that ρ^Q is a volatility process. Girsanov's theorem implies that for any $Q \in \Pi$,

$$B^Q \text{ is a local martingale under } Q.$$

A heuristic derivation of this fact is

$$\mathbb{E}_{t-}^Q\left[dB_t\right] = \mathbb{E}_{t-}\left[dB_t \frac{\xi_T^Q}{\xi_{t-}^Q}\right] = \mathbb{E}_{t-}\left[dB_t\left(\frac{\xi_{t+dt}^Q}{\xi_{t-}^Q} - 1\right)\right] = \mathbb{E}_{t-}\left[dB_t \frac{d\xi_t^Q}{\xi_{t-}^Q}\right] = \rho_t^Q dt.$$

The first equation is a version of Bayes' rule, and the second equation follows from the law of iterated expectations and the fact that B is a martingale.

The fact that B^Q is a local martingale under Q has specific implications for the d factors. For the Brownian factors, Lévy's characterization of Brownian motion implies that $B^{Q1}, \ldots, B^{Qk}$ are

[5] The conditional density process ξ^Q can be recovered from ρ^Q by the formula

$$\log \xi_t^Q = \sum_{i=1}^k \int_0^t \left(\rho_s^{Q\prime} dB_s^i - \frac{1}{2}\left(\rho_s^{Qi}\right)^2 ds\right) + \sum_{i=k+1}^d \int_0^t \left(\log\left(1+\rho_s^{Qi}\right) dB_s^i + \left(\log\left(1+\rho_s^{Qi}\right) - \rho_s^{Qi}\right) ds\right),$$

as can be verified by an application of Ito's lemma.

independent standard Brownian motions under the probability Q. For the Poissonian factors, we note that

$$B_t^{Qi} = N_t^i - \int_0^t \lambda_s^{Qi} ds, \quad \text{where} \quad \lambda_t^{Qi} = 1 + \rho_t^{Qi}, \quad i = k+1, \ldots, d, \tag{44}$$

and therefore λ^{Qi} is the arrival rate process of the point process N^i under the probability Q.

Finally, Ito's lemma plays the role of Taylor series approximations in the discrete model. We state a simple version that is sufficient for our purposes. Consider any process of the form

$$dX_t = \mu_t dt + \sigma_t' dB_t^Q,$$

where σ is a volatility process, and μ is a *drift* process, meaning that μ is a a predictable process satisfying $P\left[\int_0^t |\mu_u| \, du < \infty\right] = 1$. For any twice continuously differentiable function $f : \mathbb{R} \to \mathbb{R}$,

$$df(X_t) = \alpha_t dt + \beta_t' dB_t^Q,$$

where

$$\begin{aligned} \alpha_t &= f'(X_t) \mu_t + \frac{1}{2} f''(X_t) \sum_{i=1}^k \left(\sigma_t^i\right)^2 + \sum_{i=k+1}^d \left(f\left(X_{t-} + \sigma_t^i\right) - f(X_{t-}) - f'(X_{t-}) \sigma_t^i \right) \lambda_t^{Qi}, \\ \beta_t^i &= f'(X_{t-}) \sigma_t^i, \quad i = 1, \ldots, k; \quad \beta_t^i = f\left(X_{t-} + \sigma_t^i\right) - f(X_{t-}), \quad i = k+1, \ldots, d. \end{aligned}$$

4.2 Continuous-Time Recursive Utility

This section formulates recursive utility with a general conditional CE in last section's continuous-time stochastic setting. The formulation is specialized in the following section to correspond to expected utility, KMM and divergence CEs.

As in the discrete-time case, we assume consumption is valued in the interval (ℓ, ∞), for some fixed $\ell \in [-\infty, 1)$. A *consumption plan* is any (ℓ, ∞)-valued adapted process c, where c_t represents a consumption rate for $t < T$, and c_T represents a lump-sum terminal consumption. We write $\mathbf{1}$ to denote the consumption plan that is identically equal to one (unit consumption rate at all times terminated by unit lump-sum consumption). For each consumption plan c, we will construct a corresponding utility process $U(c)$, which is normalized so that $U(s\mathbf{1}) = s\mathbf{1}$ for any $s \in (\ell, \infty)$.

We henceforth fix a reference consumption plan c, and we simplify the notation for the corresponding utility process by writing U instead of $U(c)$. We wish to establish a continuous-time version of the utility backward recursion (28), which we heuristically express as

$$U_{t-} = \Phi\left(dt, c_t, \upsilon_t\left(U_{t+dt}\right)\right), \quad U_T = c_T. \tag{45}$$

The aggregator Φ is assumed to have continuous partial derivatives, denoted Φ_{dt}, Φ_c and Φ_υ. We postulate strictly increasing preferences throughout, and therefore Φ_c and Φ_υ are strictly positive.

Moreover, Φ must satisfy the consistency condition $U_{t-} = \Phi(0, c_t, U_{t-})$. The function Φ can be spot-dependent, that is, a predictable function of the state ω and time t. Any such dependence of Φ on (ω, t) is notationally suppressed, however, since it plays no role in what follows.

4.2.1 Conditional CE structure

Let μ and Σ be, respectively, the drift and volatility processes of U :

$$dU_t = \mu_t dt + \Sigma_t' dB_t. \tag{46}$$

(Note that $dU_t = U_{t+dt} - U_{t-}$ includes a possible time-t jump.) The coefficients in (46) can be interpreted as

$$\mu_t dt = \mathbb{E}_{t-}[dU_t] \quad \text{and} \quad \Sigma_t dt = \mathbb{E}_{t-}[dU_t dB_t]. \tag{47}$$

Therefore, $U_{t-} + \mu_t dt$ represents the risk-neutral certainty equivalent of U_{t+dt} under the prior P.

For the time being, the conditional CE υ is only assumed to satisfy

$$\upsilon_t(U_{t+dt}) = U_{t-} + \mu_t dt + \mathcal{D}_t(U_{t-}, \Sigma_t) dt. \tag{48}$$

The term $\mathcal{D}_t(U_{t-}, \Sigma_t) dt$ represents an adjustment to the risk-neutral conditional CE under P. This adjustment can reflect a prior that is different than P, as well as risk or ambiguity aversion. In the following section, we will specify the functional form of $\mathcal{D}_t$ for expected utility, smooth KMM, and smooth divergence conditional CEs. Further specifications that are consistent with (48), involving source-dependent first or second-order risk aversion, can be found in Skiadas (2008).

4.2.2 Recursive utility as a BSDE

In order to transform the utility backward recursion (45) to a mathematically rigorous version, we expand its right-hand side in a first-order Taylor expansion with respect to the arguments dt and $\upsilon_t(U_{t+dt})$, and we use expression (48) to obtain

$$U_{t-} = U_{t-} + \Phi_{dt}(0, c_t, U_{t-}) dt + \Phi_\upsilon(0, c_t, U_{t-}) (\mu_t + \mathcal{D}_t(U_{t-}, \Sigma_t)) dt.$$

As a technical aside, the utility process U is assumed to have paths that are left-continuous with right limits, and therefore one can substitute U_t for U_{t-} in the above equation without affecting its validity. We will continue to use the redundant notation U_{t-}, however, as an aid to interpretation. Continuing with our calculation, we note that the drift of U is given as

$$\mu_t = -[f(c_t, U_{t-}) + \mathcal{D}_t(U_{t-}, \Sigma_t)], \quad \text{where} \quad f(c_t, U_{t-}) = \frac{\Phi_{dt}(0, c_t, U_{t-})}{\Phi_\upsilon(0, c_t, U_{t-})}.$$

If Φ is time-varying or stochastic (but predictable), then f inherits the respective property. The heuristic recursive specification (46) of the utility process U can therefore be expressed as

$$dU_t = -[f(c_t, U_{t-}) + \mathcal{D}_t(U_{t-}, \Sigma_t)] dt + \Sigma_t' dB_t, \quad U_T = c_T. \tag{49}$$

Equation (49) is an instance of a backward stochastic differential equation (BSDE). Given the terminal value U_T, a solution to the BSDE consists of an adapted pair (U, Σ) such that (49) holds. We also refer to the process U as a solution to the BSDE if (U, Σ) is a BSDE solution for some Σ. We think of BSDE (49) as a continuous-time backward recursion for computing U. To see that, we use expressions (47) to heuristically write the relationship between the drift and volatility terms in (49) as

$$-\frac{\mathbb{E}_t\left[U_{t+dt}\right]-U_{t-}}{dt}=f\left(c_t, U_{t-}\right)+\mathcal{D}_t\left(U_{t-}, \frac{\operatorname{Cov}_{t-}\left[U_{t+dt}, dB_t\right]}{dt}\right),$$

which represents an implicit rule for computing U_{t-} in terms of U_{t+dt}.

The fixed-point nature of BSDE (49) means that special restrictions must be placed on the functions f and $\mathcal{D}$ to guarantee the existence and uniqueness of a solution. BSDE existence and uniqueness results, based on the type of Lipschitz-growth assumptions on the driver familiar from SDE theory, were first obtained by Pardoux and Peng (1990) and Duffie and Epstein (1992) (see also El Karoui, Peng, and Quenez (1997) and El Karoui and Mazliak (1997)). These conditions are violated in common homothetic applications, which includes the continuous-time version of the widely used parametrization of Epstein and Zin (1989). Existence, uniqueness and basic properties for continuous-time Epstein-Zin utility in a Brownian filtration were shown in Appendix A of Schroder and Skiadas (1999) (see also Kobylanski (2000)). Extensions of BSDE theory to include Poissonian risk include Barles, Buckdahn, and Pardoux (1997), Pardoux (1997), Pardoux, Pradeilles, and Rao (1997), Becherer (2006), Royer (2006), and others, albeit, under Lipschitz restrictions on the BSDE driver that rule out interesting representations of risk aversion. The relationship of discrete models and BSDEs is formally studied in an a rapidly increasing literature on the numerical solution of BSDEs (see, for example, Zhang (2004), Bouchard and Elie (2008), and references therein).

4.2.3 Separation of beliefs and risk/ambiguity aversion

The three conditional CEs we study in this paper allow for a meaningful notion of beliefs represented by a prior Q. We isolate the associated common structure here in an abstract way that will be fleshed out in the following section for specific functional forms.

Given any prior Q, recall that $\lambda^{Qi} = 1 + \rho^{Qi}$ is the arrival rate process of the point process N^i under Q, for any $i \in \{k+1, \ldots, d\}$, and let λ^Q denote the $(d-k)$-dimensional vector that lists these arrival rate processes. From now on, we specialize the conditional CE representation (48), by further assuming that there exists a function[6] $\mathcal{A}$ such that for every prior Q,

$$\upsilon_t\left(U_{t+dt}\right)=\mathbb{E}_{t-}^{Q}\left[U_{t+dt}\right]-\mathcal{A}\left(U_{t-}, \Sigma_t, \lambda_t^{Q}\right) dt. \tag{50}$$

The first term on the right-hand side of (50) represents the risk-neutral conditional CE under the prior Q, while the second term represents a risk-aversion adjustment. Key in this interpretation is

[6] In the simplest case, the domain of the function $\mathcal{A}$ is a subset of $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d-k}$, although $\mathcal{A}$ could also be allowed to be stochastic, as long as it is predictable, corresponding to spot-dependent risk aversion.

the fact that the function $\mathcal{A}$ is the same for any choice of the prior Q. Equation (50), therefore, represents a whole class of conditional CEs, parameterized by the prior Q, each with the same risk-aversion function $\mathcal{A}$.

Let us relate (50) to last section's CE representation (48) and hence to the utility BSDE (49). For any prior Q, we use the definition of B^Q in (43) to write the utility dynamics (46) as

$$dU_t = \mu_t^Q dt + \Sigma_t' dB_t^Q, \quad \text{where} \quad \mu_t^Q = \mu_t + \Sigma_t' \rho_t. \tag{51}$$

The CE representation (50) can therefore be restated as

$$\upsilon_t \left(U_{t+dt}\right) = U_{t-} + \left(\mu_t + \Sigma_t' \rho_t^Q - \mathcal{A}\left(U_{t-}, \Sigma_t, \lambda_t^Q\right)\right) dt,$$

which is the same as (48) with

$$\mathcal{D}_t \left(U_{t-}, \Sigma_t\right) = \Sigma_t' \rho_t^Q - \mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^Q). \tag{52}$$

The above is a decomposition of $\mathcal{D}_t \left(U_{t-}, \Sigma_t\right)$ into a belief-adjustment term and a risk-aversion term. The corresponding utility BSDE is

$$dU_t = -\left(f\left(c_t, U_t\right) + \Sigma_t' \rho_t^Q - \mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^Q)\right) dt + \Sigma_t' dB_t, \quad U_T = c_T, \tag{53}$$

which can be equivalently stated as

$$dU_t = -\left(f\left(c_t, U_t\right) - \mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^Q)\right) dt + \Sigma_t' dB_t^Q, \quad U_T = c_T.$$

There remains to specify $\mathcal{A}$ for smooth expected utility, KMM and divergence CEs.

4.3 Continuous-Time Recursive Utility Functional Forms

This final section establishes the functional form of the risk-aversion function $\mathcal{A}$ appearing in the utility BSDE (53) for an expected utility CE, corresponding to Kreps-Porteus utility, a smooth KMM CE, and a smooth divergence CE. The arguments presented are essentially a combination of the single-period approximations of section 2.2.

We fix a reference vNM index $u \in C^3_{\text{vNM}}$, defined as in section 2.1.1, and we recall the form of the key risk-aversion coefficients associated with u :

$$a^u = -\frac{u''}{u'}, \quad A_\zeta^u \left(U_{t-}, \Sigma_t^i\right) = \Sigma_t^i - \frac{\zeta\left(u\left(U_{t-} + \Sigma_t^i\right) - u\left(U_{t-}\right)\right)}{u'\left(U_{t-}\right)}, \quad A^u = A^u_{\text{identity}}.$$

We use the same notation in specifying the three conditional CE forms as for the discrete case, except that for a divergence CE the reference prior is $Q \in \Pi$ rather than P, resulting in a reference vector of arrival processes λ^Q. The other piece of key notation is ϕ^*, which stands for the convex conjugate of ϕ, defined in (25).

The paper's main conclusions are summarized in the single expression

$$\mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^Q) = \frac{a^{\psi \circ u}(U_{t-})}{2} \sum_{i=1}^{k} \left(\Sigma_t^i\right)^2 + \sum_{i=k+1}^{d} A_\zeta^u\left(U_{t-}, \Sigma_t^i\right) \lambda_t^{Qi}, \tag{54}$$

where the coefficients are specified for each smooth conditional CE case as follows:

- **Kreps-Porteus utility (expected utility CE).** ψ and ζ are identity functions, and Q is the agent's prior.

- **Klibanoff-Marinacci-Mukerji CE.** Same as for Kreps-Porteus utility, with Q being the compound prior – the function φ of the discrete formulation plays no role.

- **Divergence CE,** with reference prior $Q \in \Pi$, under the interior-solution condition

$$u\left(U_{t-} + \Sigma_t^i\right) - u\left(U_{t-}\right) < -\theta\phi'(0+), \quad i = k+1, \ldots, d. \tag{55}$$

  Equation (54) holds with Q being the reference prior,

$$\psi(u) = \theta\left(1 - \exp\left(-\frac{u}{\theta}\right)\right) \quad \text{and} \quad \zeta(\delta) = -\theta\phi^*\left(-\frac{\delta}{\theta}\right). \tag{56}$$

  Note that $a^{\psi \circ u} = a^u + \theta^{-1}u'$, while the function ϕ appears only in the risk-adjustment for Poissonian risk, and therefore recursive utility with a smooth divergence CE is Kreps-Porteus utility if there is only Brownian risk $(k = d)$. Finally, if there is Poissonian risk $(k < d)$, recursive utility with a smooth divergence CE is within the Kreps-Porteus class if and only if $\phi(x) = x \log x - x + 1$, in which case $\zeta = \psi$ and $A_\zeta^u = A^{\psi \circ u}$.

Before proceeding with a more detailed explanation of the above claims, we consider the special case of a quadratic divergence CE, or Gini CE in the language of MMR, and its relationship to a continuous-time recursive utility specification introduced in Schroder and Skiadas (2008). The last bullet point implies that this type of recursive utility is not within the Kreps-Porteus class.

Example 12 (Gini CE) *We specialize the divergence conditional CE specification by assuming that*

$$\phi(x) = \frac{1}{2}(x-1)^2, \quad \textit{and therefore} \quad \phi^*(x^*) = x^* + \frac{1}{2}(x^*)^2 \quad \textit{and} \quad \zeta(\delta) = \delta - \frac{1}{2\theta}\delta^2.$$

We also assume the validity of the interior-solution condition (55)*, which in this context reduces to*

$$u\left(U_{t-} + \Sigma_t^i\right) < u\left(U_{t-}\right) + \theta, \quad i = k+1, \ldots, d.$$

In this case, expression (54) *is specialized by setting*

$$A_\zeta^u\left(U_{t-}, \Sigma_t^i\right) = A^u\left(U_{t-}, \Sigma_t^i\right) + \frac{u'(U_{t-})}{2\theta}\left[\Sigma_t^i - A^u\left(U_{t-}, \Sigma_t^i\right)\right]^2, \quad i = k+1, \ldots, d.$$

If u is assumed to be the identity function, the Brownian and Poissonian risk adjustment terms become identical and quadratic in form, resulting in the specification

$$\mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^Q) = \frac{1}{2\theta} \left(\sum_{i=1}^{k} \left(\Sigma_t^i\right)^2 + \sum_{i=k+1}^{d} \left(\Sigma_t^i\right)^2 \lambda^{Qi} \right).$$

Assuming a differential intertemporal aggregator of the form $f_t(c_t, U_t) = g_t(c_t - U_t)$, the resulting recursive utility is within the specification of Section 6 of Schroder and Skiadas (2008). As explained there, the utility is quasilinear with respect to $\mathbf{1}$, a property that is shared with the normalized version of expected discounted exponential utility. The latter, however, results in higher than quadratic order terms in $\mathcal{A}$, given Poissonian risk, while the quadratic specification of $\mathcal{A}$ has notable tractability advantages, as demonstrated in the analysis of Schroder and Skiadas (2008).

The remainder of this section clarifies and elaborates on the conclusions summarized above.

4.3.1 Expected-utility conditional CE

We first establish the continuous-time version of the recursive utility of Kreps and Porteus (1978), in what amounts to an extension of the Duffie and Epstein (1992) analysis that clarifies the role of Poissonian risk. We specialize the setting of section 4.2.3 by assuming that

$$\upsilon_t(U_{t+dt}) = u^{-1}\mathbb{E}_{t-}^Q u(U_{t+dt}), \tag{57}$$

for some $Q \in \Pi$ and $u \in C^3_{\mathrm{vNM}}$ (or just C^2_{vNM} if there is no Brownian risk), and we verify that equation (50) holds with

$$\mathcal{A}\left(U_{t-}, \Sigma_t, \lambda_t^Q\right) = \frac{a^u(U_{t-})}{2} \sum_{i=1}^{k} \left(\Sigma_t^i\right)^2 + \sum_{i=k+1}^{d} A^u\left(U_{t-}, \Sigma_t^i\right) \lambda_t^{Qi}. \tag{58}$$

Ito's lemma together with the utility dynamics (51) implies that

$$u(U_{t+dt}) = u(U_{t-}) + u'(U_{t-})\left(\mu_t^Q - \mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^Q)\right)dt + \beta' dB_t^Q,$$

where the function $\mathcal{A}$ is defined in (58), and β is a volatility process whose precise form we do not need here. Therefore,

$$u^{-1}\mathbb{E}_{t-}^Q u(U_{t+dt}) = u^{-1}\left(u(U_{t-}) + u'(U_{t-})\left(\mu_t^Q - \mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^Q)\right)dt\right).$$

Since $\left(u^{-1}\right)'(u(U_{t-})) = 1/u'(U_{t-})$, it follows that

$$u^{-1}\mathbb{E}_{t-}^Q u(U_{t+dt}) = U_{t-} + \mu_t^Q dt - \mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^Q)dt = \mathbb{E}_{t-}^Q[U_{t+dt}] - \mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^Q)dt.$$

4.3.2 Klibanoff-Marinacci-Mukerji conditional CE

For the single-period KMM CE formulation of section 2.1.2, we saw that for either a Brownian or Poissonian small risk, the KMM CE is approximately not dependent on φ, and it is approximated by an expected utility CE with the compound prior. We now verify the same claims in the current continuous-time setting, thus concluding that continuous-time recursive utility with a smooth KMM is the same as Kreps-Porteus utility.

The irrelevance of the function φ can be made in greater generality than the KMM formulation, as follows. Taken as primitive are the conditional CEs

$$\upsilon_t^s\left(U_{t+dt}\right)=\mathbb{E}_{t-}\left[U_{t+dt}\right]+\mathcal{D}_t^s\left(U_{t-},\Sigma_t\right)dt,\quad s=1,\ldots,S, \tag{59}$$

and corresponding strictly positive predictable process $\pi^1,\ldots,\pi^S$ such that $\sum_{s=1}^S\pi_t^s=1$ for all t. The conditional CE υ is defined, for some $\varphi\in C^2_{\mathrm{vNM}}$, by

$$\upsilon_t\left(U_{t+dt}\right)=\varphi^{-1}\left(\sum_{s=1}^S\varphi\left(\upsilon_t^s\left(U_{t+dt}\right)\right)\pi_t^s\right). \tag{60}$$

The claim is that

$$\upsilon_t\left(U_{t+dt}\right)=\sum_s\upsilon_t^s\left(U_{t+dt}\right)\pi_t^s=\mathbb{E}_{t-}\left[U_{t+dt}\right]+\mathcal{D}_t\left(U_{t-},\Sigma_t\right)dt,\quad \text{where } \mathcal{D}_t=\sum_s\mathcal{D}_t^s\pi_t^s, \tag{61}$$

and therefore the corresponding recursive utility solves BSDE (49), which does not depend on φ.

To verify this claim, we use the identity $\mathbb{E}_{t-}\left[U_{t+dt}\right]=U_{t-}+\mu_t dt$ and equation (59) to obtain the first-order expansion

$$\sum_s\varphi\left(\upsilon_t^s\left(U_{t+dt}\right)\right)\pi_t^s=\varphi\left(U_{t-}\right)+\varphi'\left(U_-\right)\left(\mu_t dt+\mathcal{D}_t\left(U_{t-},\Sigma_t\right)dt\right).$$

Applying φ^{-1} on both sides, and using another first-order expansion results in (61).

The KMM formulation specializes the above setting by postulating priors $Q^1,\ldots,Q^S$ such that

$$\upsilon_t^s\left(U_{t+dt}\right)=u^{-1}\mathbb{E}_{t-}^{Q^s}u\left(U_{t+dt}\right),\quad s=1,\ldots,S,$$

and by requiring that the processes π^s are updated by Bayes rule:

$$\pi_t^s=\frac{\xi_{t-}^s\pi_0^s}{\sum_{r=1}^S\xi_{t-}^r\pi_0^r},\quad \text{where}\quad \xi_t^s=\mathbb{E}_t\left[\frac{dQ^s}{dP}\right].$$

Equations (59) can now be further refined by letting

$$\mathcal{D}_t^s\left(U_{t-},\Sigma_t\right)=\Sigma_t'\rho_t^{Q^s}-\mathcal{A}\left(U_{t-},\Sigma_t,\lambda_t^{Q^s}\right),\quad s=1,\ldots,S, \tag{62}$$

where the risk-aversion adjustment function $\mathcal{A}$ is given by equation (58). As before, the compound prior is defined by

$$Q=\sum_s Q^s\pi_0^s\quad\text{and therefore}\quad \xi^Q=\sum_s\xi^s\pi_0^s. \tag{63}$$

In addition to the irrelevance of φ, we can now claim that the conditional CE υ is an expected utility CE with vNM index u and prior Q, and therefore the corresponding recursive utility is Kreps-Porteus utility. To verify this claim, we first note that the second equation in (63) together with the Bayes formula defining π_t^s results in

$$\frac{d\xi_t^Q}{\xi_{t-}^Q} = \sum\nolimits_s \frac{d\xi_t^s}{\xi_{t-}^s}\pi_t^s, \quad \text{and therefore} \quad \rho_t^Q = \sum\nolimits_s \rho_t^{Q^s}\pi_t^s \quad \text{and} \quad \lambda_t^Q = \sum\nolimits_s \lambda_t^{Q^s}\pi_t^s.$$

Since, in equation (62), the dependence of $\mathcal{D}_t^s(U_{t-},\Sigma_t)$ on both ρ^{Q^s} and λ^{Q^s} is linear, we have

$$\mathcal{D}_t(U_{t-},\Sigma_t) = \sum\nolimits_s \mathcal{D}_t^s(U_{t-},\Sigma_t)\,\pi_t^s = \Sigma_t'\rho_t^Q - \mathcal{A}\left(U_{t-},\Sigma_t,\lambda_t^Q\right),$$

and therefore (61) becomes

$$\upsilon_t(U_{t+dt}) = \mathbb{E}_{t-}^Q[U_{t+dt}] - \mathcal{A}\left(U_t,\Sigma_t,\lambda_t^Q\right)dt = u^{-1}\mathbb{E}_{t-}^Q u(U_{t+dt}).$$

4.3.3 Divergence conditional CE

In the discrete-time model we defined a divergence CE with the underlying probability P taken as the reference prior. In our current setting this assumption is limiting, since we have selected P so as the basic Poisson processes $N^{k+1},\ldots,N^d$ are normalized to have unit arrival rate. With a single Poissonian risk source, as in section 2.2, the arrival rate can be adjusted by changing the unit of time, and therefore there is no loss of generality in assuming that P is the reference prior. In this section we allow for multiple Poissonian risk sources, and we define a divergence CE under a new prior $Q\in\Pi$ implying the arrival rates forming the vector $\lambda^Q = (\lambda^{Qi})_{i=k+1,\ldots,d}$. For any other prior $R\in\Pi$, we define the positive Q-martingale $\xi^{R/Q}$ and the predictable process $\rho^{R/Q}$ by

$$\xi_t^{R/Q} = \mathbb{E}_t^Q\left[\frac{dR}{dQ}\right] \quad \text{and} \quad \frac{d\xi_t^{R/Q}}{\xi_{t-}^{R/Q}} = \rho_t^{R/Q\prime}dB_t^Q.$$

Remark 13 *The usual change-of-measure formula for conditional expectations and the integration-by-parts formula (for semimartingales) imply the relationships*

$$\xi^{R/Q} = \frac{\xi^R}{\xi^Q}; \quad \rho^{R/Qi} = \rho^{Ri} - \rho^{Qi}, \quad i=1,\ldots,k; \quad 1+\rho_t^{R/Qi} = \frac{1+\rho^{Ri}}{1+\rho^{Qi}}, \quad i=k+1,\ldots,d.$$

The remaining primitives needed to define a divergence conditional CE are the same as in the discrete model: a positive scalar θ and a divergence index $\phi\in C^3_{\mathrm{vNM}}$, as defined in section 2.1.3. (Without Brownian risk, it is enough to assume that $\phi\in C^2_{\mathrm{vNM}}$.) We define the divergence conditional CE υ by

$$\upsilon_t(U_{t+dt}) = \inf_{R\in\Pi}\upsilon_t^R(U_{t+dt}), \tag{64}$$

where

$$\upsilon_t^R(U_{t+dt}) = u^{-1}\left(\mathbb{E}_{t-}^R u(U_{t+dt}) + \theta\mathbb{E}_{t-}^Q\phi\left(\frac{\xi_{t+dt}^{R/Q}}{\xi_{t-}^{R/Q}}\right)\right).$$

We wish to translate this heuristic expression to a formula for $\mathcal{A}$, which specifies a corresponding utility BSDE. Condition (55) is assumed if there is Poissonian risk ($k < d$). We will see shortly that condition (55) is equivalent to the existence of a minimum in (64). Since every element of Π is equivalent to P, condition (55) is essentially the requirement that the agent does not entirely switch off one of the Poissonian risk sources in the CE calculation.

We begin by applying Ito's lemma to compute the term

$$\mathbb{E}_{t-}^{R} u\left(U_{t+dt}\right) = \mathbb{E}_{t-}^{R} u\left(U_{t-} + \mu_t^R dt + \Sigma_t' dB_t^R\right) = u\left(U_{t-}\right) + u'\left(U_{t-}\right)\left(\mu_t^R - \mathcal{A}(U_{t-}, \Sigma_t, \lambda_t^R)\right) dt,$$

with $\mathcal{A}$ defined in (58), just as we did for an EU conditional CE but with R in place of P. An analogous computation yields

$$\mathbb{E}_{t-}^{Q} \phi\left(1 + \rho_t^{R/Q\prime} dB_t^Q\right) = \sum_{i=1}^{k} \frac{1}{2}\left(\rho_t^{R/Qi}\right)^2 dt + \sum_{i=k+1}^{d} \phi\left(1 + \rho_t^{R/Qi}\right)\lambda_t^{Qi} dt.$$

Combining the above terms, we find

$$\begin{aligned} \upsilon_t^R\left(U_{t+dt}\right) &= U_{t-} + \mu_t^R dt + \sum_{i=1}^{k} \mathcal{C}_t\left(\rho_t^{R/Qi}\right) dt + \sum_{i=k+1}^{d} \mathcal{J}_t\left(\rho_t^{R/Qi}\right)\lambda_t^{Qi} dt, \qquad (65) \\ \text{where} \quad \mathcal{C}_t\left(\rho_t^i\right) &= \rho_t^i \Sigma_t^i + \frac{\theta}{u'\left(U_{t-}\right)} \frac{\left(\rho_t^i\right)^2}{2} - \frac{a^u\left(U_{t-}\right)}{2}\left(\Sigma_t^i\right)^2, \\ \text{and} \quad \mathcal{J}_t\left(\rho_t^i\right) &= \rho_t^i \Sigma_t^i + \frac{\theta}{u'\left(U_{t-}\right)} \phi\left(1 + \rho_t^i\right) - A^u\left(U_{t-}, \Sigma_t^i\right)\left(1 + \rho_t^i\right). \end{aligned}$$

The above terms are minimized separately, nothing that $\mathcal{C}_t$ is quadratic and $\mathcal{J}_t$ is strictly convex. The assumed inequality (55) is equivalent to the condition $\mathcal{J}_t'\left(-1 + \varepsilon\right) < 0$ for some sufficiently small $\varepsilon > 0$, which is necessary and sufficient for $\mathcal{J}_t$ to be minimized by some ρ^i such that $1 + \rho^i$ is strictly positive. It follows that the right-hand-side of (65) is minimized by the value $\rho^{R/Q}$, where the Brownian terms are given by

$$\rho_t^{R/Qi} = -\frac{1}{\theta} u'\left(U_{t-}\right)\Sigma_t^i, \quad i = 1, \ldots, k,$$

and the Poissonian terms are given by

$$1 + \rho_t^{R/Qi} = \phi'^{-1}\left(-\frac{1}{\theta}\left(u\left(U_{t-} + \Sigma_t^i\right) - u\left(U_{t-}\right)\right)\right), \quad i = k+1, \ldots, d.$$

Substituting the minimizing value of $\rho^{R/Q}$ in (65) results in

$$\upsilon_t\left(U_{t+dt}\right) = U_{t-} + \left(\mu_t^Q - \frac{a^{\psi \circ u}\left(U_{t-}\right)}{2} \sum_{i=1}^{k}\left(\Sigma_t^i\right)^2 - \sum_{i=k+1}^{d} A_\zeta^u\left(U_{t-}, \Sigma_t^i\right)\lambda_t^{Qi}\right) dt, \qquad (66)$$

where ψ and ζ are defined in (56). The above expression is the same as (50) with the risk aversion function $\mathcal{A}$ given by equation (54).

If there is no Poissonian risk $(k = d)$, then expression (66) reduces to an expected utility conditional CE with prior Q and vNM index $\psi \circ u$, and the corresponding recursive utility is within the class of continuous-time Kreps-Porteus utilities studied by Duffie and Epstein (1992).

4.3.4 Characterization of recursive entropic utility

Finally, we show that in the presence of Poissonian risk, the continuous-time recursive divergence utility of the last subsection is in the Kreps-Porteus class if and only if its conditional CE is entropic.

We assume throughout that $0 \leq k < d$. The conditional CE υ is the divergence conditional CE defined in the last subsection, and is therefore given by (66), where ψ and ζ are defined by (56). We show that *there exists some prior* W *and* $w \in C^2_{\text{vNM}}$ *such that*

$$\upsilon_t (U_{t+dt}) = U_{t-} + \left(\mu^W_t - \frac{a^w (U_{t-})}{2} \sum_{i=1}^{k} \left(\Sigma^i_t\right)^2 - \sum_{i=k+1}^{d} A^w \left(U_{t-}, \Sigma^i_t\right) \lambda^{Wi}_t \right) dt \qquad (67)$$

for all values of (U_{t-}, Σ_t) *satisfying* (55) *if and only if*

$$W = Q \quad \text{and} \quad \phi(x) = x \log x + x - 1. \qquad (68)$$

The "if" part is immediate, since (68) implies that $\zeta = \psi$ and $A^u_\zeta = A^{\psi \circ u}$. Conversely, suppose that for some prior W and $w \in C^2_{\text{vNM}}$, equation (67) is true for all values of (U_{t-}, Σ_t) satisfying (55). Let D be the set of all $(\nu, \sigma) \in (\ell, \infty) \times \mathbb{R}$ such that $\nu + \sigma > \ell$ and $u(\nu + \sigma) - u(\nu) < -\theta\phi'(0+)$. The interior-solution condition (55) states that $\left(U_{t-}, \Sigma^i_t\right) \in D$ for every Poissonian factor i. Isolating any such factor $i \in \{k+1, \ldots, d\}$, the equality of the conditional CEs (66) and (67) implies that

$$\Sigma^i_t \rho^{Qi}_t - A^u_\zeta \left(U_{t-}, \Sigma^i_t\right) \left(1 + \rho^{Qi}_t\right) = \Sigma^i_t \rho^{Wi}_t - A^w \left(U_{t-}, \Sigma^i_t\right) \left(1 + \rho^{Wi}_t\right),$$

for all values $\left(U_{t-}, \Sigma^i_t\right)$ in D. An application of the following lemma shows the validity of condition (68). (The same lemma was used as the essential part in the proof of Proposition 3.)

Lemma 14 *Suppose* ψ *and* ζ *are defined by* (56). *Then the following statements are equivalent, for any* $w \in C^2_{vNM}$, $\sigma \in \mathbb{R}$ *and* $\rho^q, \rho^w \in (-1, \infty)$.

1. $\sigma\rho^q - A^u_\zeta(\nu, \sigma)(1 + \rho^q) = \sigma\rho^w - A^w(\nu, \sigma)(1 + \rho^w)$ *for all* $(\nu, \sigma) \in D$.
2. $\rho^q = \rho^w$ *and* $\phi(x) = x \log x - x + 1$.

Proof. See Appendix, A.3. ■

A Appendix: Proofs

This Appendix contains proofs omitted from the main text.

A.1 Proof of Proposition 2

Throughout this proof we assume that h is sufficiently small so that U_h is valued in $[U_0 - \varepsilon, U_0 + \varepsilon]$ for a some fixed $\varepsilon > 0$ (so that $U_0 - \varepsilon > \ell$). We simplify the notation by writing $\rho = \rho^Q$, and we note the moments conditions

$$\mathbb{E}^{Q_h}\left[(B^Q)^2\right] = h - (\rho h)^2 = h + o(h) \quad \text{and} \quad \mathbb{E}^{Q_h}\left[\left(B^Q\right)^3\right] = -2\rho\left(1 - \rho^2 h\right)h^2 = o(h). \tag{69}$$

(a) A second-order Taylor series expansion of u around U_0 gives

$$u(U_h) = u(U_0) + u'(U_0)\left(\mu^Q h + \Sigma B^Q\right) + \frac{1}{2}u''(U_0)\left(\mu^Q h + \Sigma B^Q\right)^2 + C_h\left(\mu^Q h + \Sigma B^Q\right)^3$$

where C_h is a random variable such that

$$|C_h| \le \frac{1}{6}\max\left\{u'''(U_0 + \delta) : \delta \in [-\varepsilon, +\varepsilon]\right\}.$$

Expanding, taking expectations under Q_h, and using the moment conditions (69), we find that

$$\mathbb{E}^{Q_h}u(U_h) = u(U_0) + u'(U_0)\left\{\left(\mu + \rho\Sigma - \frac{1}{2}a^u(U_0)\Sigma^2\right)h + r_1(\rho)h^2 + r_2(\rho)h^3\right\}. \tag{70}$$

where r_1 and r_2 are continuous functions of ρ, whose exact form is not important for our argument. The above error estimate will be used in the proof of part (c). The proof of part (a) is completed with a first-order Taylor expansion of u^{-1} around $u(U_0)$.

(b) Equation (15) and a first-order Taylor series expansion of φ around U_0 imply that

$$\varphi(\upsilon^s(U_h)) = \varphi(U_0) + \varphi'(U_0)\left(\mu + \rho^s\Sigma - \frac{1}{2}a^u(U_0)\Sigma^2\right)h + o(h), \quad s \in \{1, \ldots, S\}.$$

Therefore,

$$\upsilon(U_h) = \varphi^{-1}\left(\varphi(U_0) + \varphi'(U_0)\left(\mu + \rho\Sigma - \frac{1}{2}a^u(U_0)\Sigma^2\right)h + o(h)\right), \quad \rho = \sum\nolimits_s \rho^s\pi^s.$$

Noting that $\rho = \rho^Q$, where $Q = \sum_s Q^s\pi^s$, taking a first-order Taylor approximation of φ^{-1} around $\varphi(U_0)$, and using part (a), we conclude:

$$\upsilon(U_h) = U_0 + \left(\mu^Q - \frac{1}{2}a^u(U_0)\Sigma^2\right)h + o(h) = u^{-1}\mathbb{E}^Q u(U_h) + o(h).$$

(c) The quantity in (5) being minimized is convex in $q = Q(1) = 1 - Q(0)$. The minimum is therefore achieved for some $Q \in \Pi$ (implying $0 < Q(1) < 1$) if and only if the derivative with respect to q of the same quantity takes a negative value as q approaches zero. The last condition can easily be seen to be equivalent to

$$u\left(U_h\left(1\right)\right)-u\left(U_h\left(0\right)\right)<-\theta\phi'\left(0+\right), \tag{71}$$

which, in the Brownian risk model, is clearly satisfied for all sufficiently small values of $h\in(0,1)$. We can therefore proceed assuming the existence of a minimum. Given equations (4) on ϕ, a second-order Taylor series approximation of ϕ around one shows approximation (17), for any $Q\in\Pi$. Using approximations (70) and (17) in the CE definition (5), as well as Remark 15 at the end of this proof, we compute

$$u\left(\upsilon\left(U_h\right)\right)=u\left(U_0\right)+u'\left(U_0\right)\min_{\rho}\left\{\mu-\frac{1}{2}a^u\left(U_0\right)\Sigma^2+\Sigma\rho+\frac{\theta}{2u'\left(U_0\right)}\rho^2\right\}h+o\left(h\right),$$

The minimum is achieved for

$$\rho=-\frac{1}{\theta}u'\left(U_0\right)\Sigma.$$

Computing the corresponding minimum and utilizing a first-order Taylor expansion of u^{-1} around $u\left(U_0\right)$, we obtain

$$\upsilon\left(U_h\right)=U_0+\left(\mu-\frac{1}{2}a^{\psi\circ u}\left(U_0\right)\Sigma^2\right)h+o\left(h\right),$$

where $a^{\psi\circ u}=a^u+\left(a^{\psi}\circ u\right)\left(u'\right)$ is the (local) coefficient of absolute risk aversion of $\psi\circ u$. Referring to the expected utility CE approximation (15) with $Q=P$, equation (16) follows.

Remark 15 *The interchange of the approximation and minimization operations is justified by showing that the error terms in the relevant Taylor approximations are uniform in ρ. Assuming sufficiently small h, we have seen that the minimum in the CE definition is achieved for some Q assigning strictly positive probability to each state. This observation allows us to assume that ρ takes values in a compact interval, over which the continuous function ϕ''' is bounded. The error term in* (17) *can therefore be bounded by $Kh^{3/2}$, where K is a constant that does not depend on ρ within the given compact interval. An analogous argument, based on the error estimate in part (a), gives a uniform bound for the error term in* (15).

A.2 Proof of Proposition 3

(a) Expression (11) and first-order Taylor series expansions of u around U_0 and $U_0+\Sigma$ imply

$$\begin{aligned}u\left(U_h\left(0\right)\right)\left(1-\lambda^Q h\right) &= u\left(U_0\right)\left(1-\lambda^Q h\right)+u'\left(U_0\right)\left(\mu^Q-\lambda^Q\Sigma\right)h+o\left(h\right),\\ u\left(U_h\left(1\right)\right)\lambda^Q h &= u\left(U_0+\Sigma\right)\lambda^Q h+o\left(h\right).\end{aligned}$$

Adding the two equations and rearranging results in

$$\mathbb{E}^Q u\left(U_h\right)=u\left(U_0\right)+u'\left(U_0\right)\left(\mu^Q-A^u\left(U_0,\Sigma\right)\lambda^Q\right)h+o\left(h\right). \tag{72}$$

A first-order Taylor expansion of u^{-1} around $u\left(U_0\right)$ completes the proof.

(b) Since μ^Q and λ^Q are both linear functions of ρ^Q, we have

$$\mu^Q = \sum_{s=1}^{S} \mu^{Q(s)} \pi^s \quad \text{and} \quad \lambda^Q = \sum_{s=1}^{S} \lambda^{Q(s)} \pi^s.$$

The just derived approximation for an expected utility CE implies that for each $s \in \{1, \ldots, S\}$,

$$\varphi\left(\upsilon^{Q(s)}\left(U_h\right)\right) = \varphi\left(U_0\right) + \varphi'\left(U_0\right)\left(\mu^{Q(s)} - A^u\left(U_0, \Sigma\right)\lambda^{Q(s)}\right)h + o\left(h\right),$$

and therefore

$$\upsilon\left(U_h\right) = \varphi^{-1}\left(\varphi\left(U_0\right) + \varphi'\left(U_0\right)\left(\mu^Q - A^u\left(U_0, \Sigma\right)\lambda^Q\right)h + o\left(h\right)\right).$$

Finally, a first-order Taylor approximation of φ^{-1} around $\varphi\left(U_0\right)$ results in the CE approximation

$$\upsilon\left(U_h\right) = U_0 + \left(\mu^Q - A^u\left(U_0, \Sigma\right)\lambda^Q\right)h + o\left(h\right) = u^{-1}\mathbb{E}^Q u\left(U_h\right) + o\left(h\right),$$

where the last equation follows from part (a).

(c) Given any prior Q, approximation (72), the definition of A^u, and the fact that $\mu^Q = \mu + \Sigma\rho^Q$ and $\lambda^Q = 1 + \rho^Q$ result in

$$\mathbb{E}^Q u\left(U_h\right) = u\left(U_0\right) + u'\left(U_0\right)\left(\mu - \Sigma\right)h + \left[u\left(U_0 + \Sigma\right) - u\left(U_0\right)\right]\left(1 + \rho^Q\right)h + o\left(h\right).$$

This equation can be reapplied with new variables to conclude, using restrictions (4) on ϕ, that

$$\mathbb{E}\phi\left(\frac{dQ}{dP}\right) = \mathbb{E}\phi\left(1 + \rho^Q B\right) = \phi\left(1 + \rho^Q\right)h + o\left(h\right).$$

Combining the last two approximations, we have

$$\begin{aligned}\min_Q\left\{\mathbb{E}^Q u\left(U_h\right) + \theta\mathbb{E}\phi\left(\frac{dQ}{dP}\right)\right\} &= u\left(U_0\right) + u'\left(U_0\right)\left(\mu - \Sigma\right)h \\ &\quad + \min_\rho\left\{\left[u\left(U_0 + \Sigma\right) - u\left(U_0\right)\right]\left(1 + \rho\right) + \theta\phi\left(1 + \rho\right)\right\}h + o\left(h\right).\end{aligned}$$

The interchange of the minimization and approximation operations is justified along the lines of Remark 15. One can easily confirm that there is a minimizing value of ρ such that $1 + \rho > 0$ if and only if inequality (71) holds, or equivalently $\left(U_0, \Sigma\right) \in D$, in which case the minimum is achieved for

$$1 + \rho = \phi'^{-1}\left(-\frac{1}{\theta}\left(u\left(U_0 + \Sigma\right) - u\left(U_0\right)\right)\right).$$

Moreover, it is easy to confirm that

$$\zeta\left(\delta\right) = \theta\phi\left(\phi'^{-1}\left(-\frac{\delta}{\theta}\right)\right) + \delta\phi'^{-1}\left(-\frac{\delta}{\theta}\right), \quad \text{for any } \delta \in \left(-\infty, -\theta\phi'\left(0+\right)\right).$$

Substituting the optimal value of $1+\rho$ and using the above expression for ζ, we obtain

$$u\left(\upsilon\left(U_h\right)\right) = u\left(U_0\right) + u'\left(U_0\right)\left\{\mu - A_\zeta^u\left(U_0, \Sigma\right)\lambda\right\} h + o\left(h\right).$$

Finally, a first-order approximation of u^{-1} around $u\left(U_0\right)$ gives the claimed expression for $\upsilon\left(U_h\right)$.

In order to prove the final claim of part (c), we note that, by part (a), the CE approximation (27) is equivalent to

$$-A_\zeta^u\left(U_0, \Sigma\right) = \Sigma\rho^W - A^w\left(U_0, \Sigma\right)\left(1+\rho^W\right), \quad \left(U_0, \Sigma\right) \in D.$$

The proof is completed by Lemma 14, which is stated in section 4.3.4 and is proved below.

A.3 Proof of Lemma 14

We show the implication ($1 \implies 2$), since the converse is a matter of simple computation. We assume that

$$w'\left(1\right) = u'\left(1\right) \quad \text{and} \quad w\left(1\right) = u\left(1\right) = 0,$$

which entails no loss of generality since A^w is invariant to a positive affine transformation of w, and A_ζ^u is invariant to adding a constant to u.

Suppose that

$$\sigma\rho^q - A_\zeta^u\left(\nu, \sigma\right)\left(1+\rho^q\right) = \sigma\rho^w - A^w\left(\nu, \sigma\right)\left(1+\rho^w\right), \quad \text{for all } \left(\nu, \sigma\right) \in D.$$

Defining the function

$$f\left(\delta\right) = \frac{1+\rho^q}{1+\rho^w}\zeta\left(\delta\right),$$

the assumed condition is equivalent to

$$\frac{f\left(u\left(\nu+\sigma\right) - u\left(\nu\right)\right)}{u'\left(\nu\right)} = \frac{w\left(\nu+\sigma\right) - w\left(\nu\right)}{w'\left(\nu\right)}. \tag{73}$$

Letting $\nu = 1$ and $z = 1+\sigma$ it follows that for any z such that $u\left(z\right) < -\theta\phi'\left(0+\right)$,

$$f\left(u\left(z\right)\right) = w\left(z\right) \quad \text{and therefore } f'\left(u\left(z\right)\right)u'\left(z\right) = w'\left(z\right).$$

Assuming

$$x = u\left(\nu\right) < -\theta\phi'\left(0+\right) \quad \text{and} \quad y = u\left(\nu+\sigma\right) - u\left(\nu\right) < -\theta\phi'\left(0+\right),$$

condition (73) becomes

$$f'\left(x\right)f\left(y\right) = f\left(x+y\right) - f\left(x\right), \quad x, y \in \left(-\infty, -\theta\phi'\left(0+\right)\right).$$

Differentiating with respect to y and taking logs results in

$$\log f'\left(x\right) + \log f'\left(y\right) = \log f'\left(x+y\right), \quad x, y \in \left(-\infty, -\theta\phi'\left(0+\right)\right).$$

Since f' is continuous, it follows that there exists a scalar a such that

$$\log f'(\delta) = a\delta, \quad \delta \in \left(-\infty, -\theta\phi'(0+)\right). \tag{74}$$

For any $x > 0$, the definition of ϕ^* implies the identity $\phi^*\left(\phi'(x)\right) = \phi'(x)\,x - \phi(x)$. Differentiating and simplifying (using the fact that $\phi'' > 0$) results in

$$\phi^{*\prime}\left(\phi'(x)\right) = x, \quad x \in (0, \infty). \tag{75}$$

Let us now assume that x and δ are related by

$$\phi'(x) = -\frac{\delta}{\theta}, \quad \text{and therefore} \quad x \in (0,\infty) \iff \delta \in \left(-\infty, -\theta\phi'(0+)\right).$$

Differentiating $\zeta(\delta) = -\theta\phi^*(-\delta/\theta)$, substituting $\delta = -\theta\phi'(x)$ and using (75), we find

$$\zeta'\left(-\theta\phi'(x)\right) = x, \quad x \in (0,\infty).$$

Identity (74) with $\delta = -\theta\phi'(x)$ becomes

$$\log\left(\frac{1+\rho^q}{1+\rho^w}\right) + \log x = -a\theta\phi'(x), \quad x \in (0,1).$$

Since $\phi'(1) = 0$, it follows that $\rho^w = \rho^q$. Since $\phi''(1) = 1$, it follows that $a\theta = -1$. Therefore, ϕ solves the ODE

$$\phi(1) = 0, \quad \phi'(x) = \log(x) \quad \text{for all } x > 0,$$

whose unique solution is $\phi(x) = x\log x - x + 1$.

A.4 Proof of Proposition 8

Consider the Lagrangian

$$\mathcal{L}_t(x_{t+h}, y_t) = \mathbb{E}_t\left[x_{t+h}u(U_{t+h}) + \theta\phi(x_{t+h}) + y_t(x_{t+h} - 1)\right], \tag{76}$$

where $x_{t+h} \in L_{++}(\mathcal{F}_{t+h})$ is the variable corresponding to ξ^Q_{t+h}/ξ^Q_t, and $y_t \in L(\mathcal{F}_t)$ is the Lagrange multiplier corresponding to the constraint $\mathbb{E}_t[x_{t+h}] = 1$. For any given y_t, the value $x_{t+h} = G(U_{t+h}, y_t)$ minimizes the quantity inside the expectation in (76) state by state, and therefore also minimizes $\mathcal{L}_t(x_{t+h}, y_t)$. Assuming, for now, that λ_t satisfies (37), recursion (38) defines a positive martingale ξ^Q, which is the conditional density process of the probability $Q \in \Pi$ defined by $Q(F) = \mathbb{E}\left[\xi^Q_T 1_F\right]$ for every event F. For any other probability $R \in \Pi$, we have

$$\begin{aligned}
\mathbb{E}^R_t u(U_{t+h}) + \theta\mathbb{E}_t\phi\left(\xi^R_{t+h}/\xi^R_t\right) &= \mathcal{L}_t\left(\xi^R_{t+h}/\xi^R_t, \lambda_t\right) \\
&\geq \mathcal{L}_t\left(\xi^Q_{t+h}/\xi^Q_t, \lambda_t\right) = \mathbb{E}^Q_t u(U_{t+h}) + \theta\mathbb{E}_t\phi\left(\xi^Q_{t+h}/\xi^Q_t\right),
\end{aligned}$$

which proves that Q is the minimizing probability.

There remains to show the existence of a unique $\lambda_t \in L(\mathcal{F}_t)$ satisfying (37). Fixing any event $F \in \mathcal{F}_t^0$, consider the function

$$g(y) = \mathbb{E}\left[G\left(U_{t+h}, y\right) \mid F\right], \quad y < -\theta\phi'(0+) - u\left(U_t^*\right).$$

Clearly, g is strictly decreasing and continuous. Because of condition (36), as y approaches its upper limit, $g(y)$ takes values smaller than one. As y approaches $-\infty$, then the last part of condition (4) implies that $g(y)$ takes values greater than one. There is, therefore, a unique value of y such that $g(y) = 1$. Letting $\lambda_t(\omega)$ be equal to that value of y for every $\omega \in F$ completes the proof.

References

Arrow, K. J. (1965): *Aspects of the Theory of Risk Bearing.* Yrjo Jahnssonin Saatio, Helsinki.

——— (1970): *Essays in the Theory of Risk Bearing.* North Holland, London.

Barles, G., R. Buckdahn, and E. Pardoux (1997): "Backward Stochastic Differential Equations and Integro-Partial Differential Equations," *Stochastics and Stochastics Reports*, 60, 57–83.

Barro, R. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 121, 823–866.

Becherer, D. (2006): "Bounded Solutions to Backward SDE's with Jumps for Utility Optimization and Indifference Hedging," *Annals of Applied Probability*, 16, 2027–2054.

Billingsley, P. (1999): *Convergence of Probability Measures.* John Wiley & Sons, New York, second edn.

Bouchard, B., and R. Elie (2008): "Discrete Time Approximation of Decoupled Forward-Backward SDE with Jumps," *Stochastic Processes and their Applications*, 118, 53–75.

Brémaud, P. (1981): *Point Processes and Queues: Martingale Dynamics.* Springer-Verlag, New York.

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

Chew, S. H., and L. G. Epstein (1989): "The Structure of Preferences and Attitudes Toward the Timing of the Resolution of Uncertainty," *International Economic Review*, 30, 103–117.

Cont, R., and P. Tankov (2004): *Financial Modelling with Jump Processes.* Chapman&Hall/CRC, London, UK.

Donsker, M. D., and S. R. S. Varadhan (1975): "Asymptotic Evaluation of Certain Markov Process Expectations for Large Time I," *Communications on Pure and Applied Mathematics*, 28, 1–47.

Duffie, D., and L. G. Epstein (1992): "Stochastic Differential Utility," *Econometrica*, 60, 353–394.

El Karoui, N., and L. Mazliak (eds.) (1997): *Backward Stochastic Differential Equations*. Addison Wesley Longman, Essex, England.

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

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669.

Epstein, L., and J. Miao (2003): "A Two-Person Dynamic Equilibrium under Ambiguity," *Journal of Economic Dynamics and Control*, 27, 1253–1288.

Epstein, L. G., and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Gabaix, X. (2008): "Variable Rare Disasters: A Tractable Theory of Ten Puzzles in Macro-Finance," *American Economic Review: Papers & Proceedings*, 98.

Gilboa, I., and D. Schmeidler (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

Hansen, L., T. Sargent, G. Turmuhambetova, and N. Williams (2006): "Robust Control and Model Misspecification," *Journal of Economic Theory*, 128, 45–90.

Hansen, L. P., and T. J. Sargent (2001): "Robust Control and Model Uncertainty," *American Economic Review*, 91, 60–66.

——— (2007): *Robustness*. Pinceton University Press, Princeton, NJ.

Hayashi, T. (2005): "Intertemporal Substitution, Risk Aversion and Ambiguity Aversion," *Economic Theory*, 25, 933–956.

Jacod, J., and A. N. Shiryaev (2003): *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin Heidelberg, second edn.

Johnsen, T. H., and J. B. Donaldson (1985): "The Structure of Intertemporal Preferences under Uncertainty and Time Consistent Plans," *Econometrica*, 53, 1451–1458.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005): "A Smooth Model of Decision Making Under Ambiguity," *Econometrica*, 73, 1849–1892.

——— (2007): "Recursive Smooth Ambiguity Preferences," working paper, MEDS, Kellogg School of Management.

Klibanoff, P., and E. Ozdenoren (2007): "Subjective Recursive Expected Utility," *Economic Theory*, 30, 49–87.

Kobylanski, M. (2000): "Backward Stochastic Differential Equations and Partial Differential Equations with Quadratic Growth," *The Annals of Probability*, 28, 558–602.

Kreps, D., and E. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006a): "Ambiguity Aversion, Robustness, and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498.

——— (2006b): "Dynamic Variational Preferences," *Journal of Economic Theory*, 128, 4–44.

Machina, M. J. (1982): "'Expected Utility' Analysis Without the Independence Axiom," *Econometrica*, 50, 277–324.

Pardoux, E. (1997): "Generalized Discontinuous Backward Stochastic Differential Equations," in *Backward Stochastic Differential Equations*, ed. by N. El Karoui, and L. Mazliak, pp. 207–219. Addison Wesley Longman, Essex, England.

Pardoux, E., and S. Peng (1990): "Adapted Solution of a Backward Stochastic Differential Equation," *Systems and Control Letters*, 14, 55–61.

Pardoux, E., F. Pradeilles, and Z. Rao (1997): "Probabilistic Interpretation of a System of Semi-Linear Parabolic Partial Differential Equations," *Annales de l'Institut Henri Poincaré*, 33, 467–490.

Pratt, J. W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122–136.

Rietz, T. A. (1988): "The Equity Risk Premium: A Solution," *Journal of Monetary Economics*, 22, 117–131.

Rockafellar, R. T. (1970): *Convex Analysis.* Princeton University Press, Princeton, New Jersey.

Royer, M. (2006): "Backward Stochastic Differential Equations with Jumps and Related Non-Linear Expectations," *Stochastic Processes and their Applications*, 116, 1357–1376.

Schroder, M., and C. Skiadas (1999): "Optimal Consumption and Portfolio Selection with Stochastic Differential Utility," *Journal of Economic Theory*, 89, 68–126.

——— (2003): "Optimal Lifetime Consumption-Portfolio Strategies under Trading Constraints and Generalized Recursive Preferences," *Stochastic Processes and Their Applications*, 108, 155–202.

——— (2005): "Lifetime Consumption-Portfolio Choice under Trading Constraints and Nontradeable Income," *Stochastic Processes and their Applications*, 115, 1–30.

——— (2008): "Optimality and State Pricing in Constrained Financial Markets with Recursive Utility under Continuous and Discontinuous Information," *Mathematical Finance*, 18, 199–238.

Selden, L. (1978): "A New Representation of Preferences over 'Certain x Uncertain' Consumption Pairs: The 'Ordinal Certainty Equivalent' Hypothesis," *Econometrica*, 46, 1045–1060.

Skiadas, C. (1998): "Recursive Utility and Preferences for Information," *Economic Theory*, 12, 293–312.

——— (2003): "Robust Control and Recursive Utility," *Finance and Stochastics*, 7, 475–489.

——— (2008): "Dynamic Portfolio Choice and Risk Aversion," in *Handbooks in OR & MS,* Vol. 15, ed. by J. R. Birge, and V. Linetsky, chap. 19, pp. 789–843. Elsevier.

——— (forthcoming): *Asset Pricing Theory.* Princeton Univ. Press, Princeton, NJ.

Strzalecki, T. (2008): "Axiomatic Foundations of Multiplier Preferences," working paper, Dept. of Economics, Northwestern University.

Wang, T. (2003): "Conditional Preferences and Updating," *Journal of Economic Theory*, 108, 286–321.

Zhang, J. (2004): "A Numerical Scheme for BSDEs," *Annals of Applied Probability*, 14, 459–488.