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

Costis Skiadas*

Journal of Political Economy, forthcoming

Abstract

Assuming Brownian/Poisson uncertainty, a certainty equivalent (CE) based on the smooth second-order expected utility of Klibanoff, Marinacci, and Mukerji (*Econometrica*, 2005) is shown to be approximately equal to an expected-utility CE. As a consequence, the corresponding continuous-time recursive utility form is the same as for Kreps-Porteus utility. The analogous conclusions are drawn for a smooth divergence CE, based on a formulation of Maccheroni, Marinacci, and Rustichini (*Econometrica*, 2006), but only under Brownian uncertainty. Under Poisson uncertainty, a smooth divergence CE can be approximated with an expected-utility CE if and only if it is of the entropic type. A non-entropic divergence CE results in a new class of continuous-time 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 for helpful discussions with Nabil Al-Najjar, Larry Epstein, Ronald Gindrat, Lars Hansen, Soohun Kim, Peter Klibanoff, Fabio Maccheroni, Mark Machina, Massimo Marinacci, Jianjun Miao, Sujoy Mukerji, Dimitris Papanikolaou and Viktor Todorov. I am especially thankful for the feedback I have received from Christian Hellwig, Monika Piazzesi and Mark Schroder. I'm responsible for any errors. The latest version of this article, along with the Appendix and computer code for the example of Section 2, can be downloaded at http://www.kellogg.northwestern.edu/faculty/skiadas/home.htm

1 Introduction

The distinction between risk and uncertainty or ambiguity (terms we use as synonyms) dates back to Knight (1921) and Keynes (1937) and continues to be the subject of great research interest to this day. The concept is colorfully explained by Silver (2012) in his popular account of statistical ideas: "Say that you'll win a poker hand unless your opponent draws to an inside straight: the chances of that happening are exactly 1 chance in 11. This is risk. It is not pleasant when you take a "bad beat" in poker, but at least you know the odds of it and can account for it ahead of time. ... Uncertainty, on the other hand, is risk that is hard to measure. You might have some vague awareness of the demons lurking out there. You might even be acutely concerned about them. But you have no real idea how many of them there are or when they might strike. Your back-of-the-envelope estimate might be off by a factor of 100 or by a factor of 1,000; there is no good way to know. This is uncertainty. Risk greases the wheels of a free-market economy; uncertainty grinds them to a halt."

The last sentence is a caricature of the old loose idea that aversion to Knightian uncertainty, as famously exhibited in the Ellsberg (1961) experiments, can act as a type of market friction, inhibiting trade or giving rise to higher risk premia than could reasonably be expected if risk estimates were taken at face value. This intuition has motivated a growing literature that employs various formalizations of ambiguity aversion to model economic phenomena that have so far proved a challenge to the neoclassical approach. The most prominent formalization of ambiguity aversion is the nonsmooth maxmin expected utility[1] of Gilboa and Schmeidler (1989). As an example, among many, Caballero and Krishnamurthy (2008) use maxmin expected utility to formulate a model of crises and central bank policy driven by Knightian uncertainty. Another early proposed representation of ambiguity aversion, due to Bewley (1986), is also nonsmooth.

The last decade saw a resurgence of interest in formalizations of ambiguity aversion, with emphasis on smooth utility representations. In this paper, we focus on two smooth representations of ambiguity aversion that have received some attention in the literature: Second-order expected utility, proposed by Klibanoff, Marinacci, and Mukerji (2005) and Nau (2006), and the smooth case of divergence preferences, proposed by Maccheroni, Marinacci, and Rustichini (2006a) as an extension of the entropic variational preferences of Hansen and Sargent (2001). From an applied theorist's perspective, a main question is, do these preferences have quantitative implications that can not be obtained more parsimoniously within the Bayesian, expected-utility framework? This paper's conclusions give a mostly negative answer for a broad class of potential applications.

To take an example application area, Ju and Miao (2012) and Collard, Mukerji, Sheppard, and Tallon (2011) use second-order expected utility to tackle the challenge of explaining the historically high average premium that investors require for bearing aggregate market risk, given that aggregate

[1] Gilboa (2009) offers a readable introduction to maxmin expected utility and its relationship to the Choquet expected utility of Schmeidler (1989). Dynamic extensions of maxmin expected utility were formulated by Epstein and Schneider (2003) in discrete time, and Chen and Epstein (2002) in continuous time.

consumption varies apparently too little to justify such a high premium. This so-called equity premium puzzle has been one of the central themes of asset pricing since it was posed by Mehra and Prescott (1985). It would certainly be interesting if ambiguity aversion in the form of smooth second-order expected utility resolved the matter. This paper's first main conclusion, however, raises the following difficulty. While Ju and Miao (2012) and Collard et. al. (2011) use annual data, in practice agents receive information at a much higher frequency. Given the type of information structures commonly assumed in macroeconomics and finance, the utility penalty due to ambiguity aversion modeled through smooth second-order expected utility decreases as the frequency increases and disappears entirely in the continuous-time limit. Over short time intervals, only small incremental risks are resolved, and for such small risks, second-order expected utility is well-approximated by the corresponding expected utility obtained by eliminating ambiguity aversion.

An essential assumption underlying this paper's arguments is that the information stream is well-approximated by a continuous-time model in which all uncertainty is driven by a fixed finite number of Brownian motions and Poisson processes. This assumption is consistent with almost all continuous-time or high-frequency models in macroeconomics and finance for well-understood probabilistic reasons.[2] Over a short time interval, a Brownian shock represents a small change with high probability, just like a stock price is expected to vary by some small amount on any given day. A Poissonian shock represents a large change with a small probability that comes as a surprise, just like a stock price can occasionally surprise with a "multiple-sigma" drop. The fact that the drift and volatility of various processes can depend on the entire history of Brownian and Poisson shocks makes this a very rich modeling framework.

The distinction between Brownian and Poissonian uncertainty is important in explaining this paper's insights with regard to smooth divergence preferences. It is well-known (see Example 3) that over a single period, entropic variational utility can be equivalently expressed as expected utility. Ambiguity aversion in this case is an argument for greater risk aversion, but does not take us outside the realm of expected utility. (A continuous-time dynamic version of this fact is shown in Skiadas 2003.) This paper's second main conclusion is that in high frequency with Brownian information, smooth divergence preferences can be approximated by entropic variational preferences and hence by preferences admitting an expected utility representation. Once again, the incorporation of smooth ambiguity aversion does not take us beyond the realm of expected utility. Under Poisson information, however, the analogous conclusion is valid if and only if the divergence preferences are of the entropic type. This dichotomy presents the interesting possibility that in a mixed Brownian/Poisson model, ambiguity aversion modeled through smooth divergence preferences can be calibrated to affect the pricing of Brownian and Poissonian uncertainty differently.

While the paper's conclusions have so far been described entirely in terms of single-period preferences, it is important to note that these preferences are taken to a dynamic setting by embedding

[2] Brownian motions and Poisson processes are the building blocks of all processes with stationary and independent increments that are continuous in probability. (See, for example, Chapter I of Protter (2004).) By bounding and discretizing the possible jump sizes, a finite number of Poisson processes are enough in terms of practical modeling.

them in a smooth recursive utility function. The approach is consistent with the way Klibanoff, Marinacci, and Mukerji (2009) and Maccheroni, Marinacci, and Rustichini (2006b) extend their respective static formulations to dynamic settings. Chapter 6 of Skiadas (2009) discusses a related literature and provides a simple axiomatic foundation for recursive utility with an arbitrary certainty equivalent (CE), which can be specialized to the CEs of interest in this paper. In the case of an expected-utility (EU) CE, the resulting recursive utility is Kreps and Porteus (1978) utility, whose continuous-time version, assuming smoothness, is given by Duffie and Epstein (1992). The present paper's CE approximations suggest the modifications to Duffie-Epstein utility that are required if one is to replace the EU CE with a smooth second-order EU CE or a smooth divergence CE. Under Brownian uncertainty, no modifications are required. The same is true under Poissonian uncertainty and a smooth second-order EU CE. Only in the case of Poissonian uncertainty and a non-entropic divergence CE we obtain a genuinely new continuous-time utility functional form.

Citing earlier versions of the present paper, Hansen and Sargent (2011), Gindrat and Lefoll (2010) and Maccheroni, Marinacci, and Ruffino (forthcoming) offered some suggestions for breaking the approximate equivalence of smooth second-order expected utility and expected utility for small risks. The first paper effectively takes the ambiguity aversion coefficient to infinity as the frequency increases to infinity, the second paper blows up the drift of the risk source as the frequency increases to infinity, while the third paper is a one-shot model and does not address at all how a small risk should be calibrated to correspond to a sampled stochastic process at increasingly higher frequencies. None of these papers negate any of the present paper's conclusions.

The remainder of the paper is organized in five sections. Section 2 gives a numerical illustration of the irrelevance of ambiguity aversion under a smooth second-order expected-utility CE in a Brownian setting. Section 3 presents the uncertainty model over each step of the information tree. Section 4 presents the CE approximations. Section 5 discusses implications for continuous-time recursive utility. Section 6 concludes with clarifying remarks and further discussion of some related work. An online Appendix contains proofs and additional results and details.

2 A Numerical Example

A central conclusion of this paper is the irrelevance of ambiguity aversion modeled through smooth second-order expected utility in a high-frequency Brownian-information setting. This section presents a simple numerical illustration of this point, using off-the-shelf components: The discrete tree representing Brownian information is the same as in textbook accounts of binomial option pricing theory, and the single-period preferences are taken to a dynamic setting by modifying the widely used parameterization of Kreps and Porteus (1978) utility due to Epstein and Zin (1991) and Weil (1989). The example also helps frame the ensuing analysis more generally by clarifying the sense in which small-risk CE approximations relate to a dynamic setting.

Time is discrete and indexed by $t = 0, h, 2h, 3h, \ldots$, where h denotes the time length of each period in years. A consumption plan c specifies, for each time t, a consumption amount c_t that is contingent on the resolution of time-t information. With c we associate a utility process U, where U_t represents the time-t utility of the consumption stream $(c_t, c_{t+h}, c_{t+2h}, \dots)$. We select utility units to correspond to equivalent consumption perpetuities, meaning that at time t the agent is indifferent between receiving the constant perpetuity $(U_t, U_t, U_t, \dots)$ and the remaining plan $(c_t, c_{t+h}, c_{t+2h}, \dots)$. In order to truncate the infinite horizon, we freeze all information from time T onward, which implies that $c_T = c_{T+h} = c_{T+2h} = \dots$ and therefore $U_T = c_T$.

For each time $t < T$, a (conditional) CE υ_t is a function that reduces the random variable U_{t+h} to a random variable $\upsilon_t (U_{t+h})$ whose value is resolved at time t. We assume, for now, that υ_t is the CE implied by expected utility with a constant coefficient of relative risk aversion γ; that is, $\upsilon_t = u_\gamma^{-1} \mathbb{E}_t u_\gamma$, where $\mathbb{E}_t$ denotes a conditional expectation operator given time-t information, and $u_\gamma (x) = \left(x^{1-\gamma} - 1\right) / (1 - \gamma)$ (with $u_1 (x) = \log x$). The Epstein-Zin-Weil specification recursively computes the process U in terms of c, starting with the terminal value $U_T = c_T$, and proceeding backward in time according to the recursion:

$$U_t = u_\delta^{-1} \left(\left(1 - e^{-\beta h}\right) u_\delta (c_t) + e^{-\beta h} u_\delta (\upsilon_t (U_{t+h})) \right), \tag{1}$$

where δ is the inverse of the elasticity of intertemporal substitution (EIS) and β is the rate of impatience.[3]

Let us now modify the CE υ_t to correspond to a simple instance of second-order expected utility, resulting in a recursive utility of the type adopted by Ju and Miao (2012). An agent considers priors Q^1 and Q^2 as equally plausible but choosing one or the other entails ambiguity. Prior Q^i defines the EU CE $\upsilon_t^i = u_\gamma^{-1} \mathbb{E}_t^i u_\gamma$, where $\mathbb{E}^i$ denotes expectation under Q^i. The agent's CE is

$$\upsilon_t (U_{t+h}) = u_\alpha^{-1} \left(\frac{1}{2} u_\alpha \left(\upsilon_t^1 (U_{t+h})\right) + \frac{1}{2} u_\alpha \left(\upsilon_t^2 (U_{t+h})\right) \right),$$

where $\alpha \geq \gamma$, reflecting higher aversion to uncertainty associated with the choice of a prior than to risk given a prior. Let $\mathbb{E}$ denote expectation under the compound prior $0.5Q^1 + 0.5Q^2$. The preceding CE reduces to $\upsilon_t = u_\gamma^{-1} \mathbb{E}_t u_\gamma$ if and only if $\alpha = \gamma$. We are interested in the quantitative significance of the ambiguity-aversion parameter α.

Towards a numerical example, suppose that log-consumption follows a binomial-tree approximation of the process $\log c_t = \log 100 + \mu t + 0.03 B_t$, where B is a standard Brownian motion. The drift parameter μ is 0 under prior Q^1 and 0.05 per year under prior Q^2; ambiguity is about the expected consumption growth rate, while consumption volatility is perfectly observable. Let also $\beta = 0.03$ per year ($e^{-\beta} \approx 0.97$), $\gamma = 2$, EIS $= 1/\delta = 1.5$ and $T = 10$ years. The familiar discretization of $\log c$ on a binomial tree is spelled out in Section 2.B of Dixit and Pindyck (1994) and is consistent with the Brownian uncertainty model introduced later in Section 3. Based on

[3] Time-additive expected discounted power or logarithmic utility results if and only if $\gamma = \delta$. Allowing γ and δ to differ partially disentangles preferences for smoothing across time and states of the world, resulting in a richer framework for modeling equilibrium interest rates and risk premia.

the latter, we compute[4] the following time-zero utility values, where the frequencies correspond to $h = 1/4,\ 1/12,\ 1/52$ and $1/365$:

frequency:	quarterly	monthly	weekly	daily
$\alpha = \gamma = 2$	$U_0 = 123.70$	123.71	123.72	123.72
$\alpha = 20,\ \gamma = 2$	$U_0 = 122.19$	123.21	123.60	123.70
$\alpha = 40,\ \gamma = 2$	$U_0 = 120.56$	122.65	123.47	123.68
$\alpha = \gamma = 40$	$U_0 = 107.84$	107.03	106.74	106.67

Note that in the first three rows, the sensitivity of the utility value on the ambiguity aversion parameter α evaporates as we increase the frequency (across columns). In contrast, the first and last rows show that Epstein-Zin-Weil utility (corresponding to $\alpha = \gamma$) remains sensitive to the coefficient of relative risk aversion γ at all frequencies. To get an idea of why this is happening, let's say $h = 10^{-4}$. Over a single period, log-consumption can vary plus-minus one standard deviation, whose size is $0.03 \times \sqrt{h} = 3 \times 10^{-4}$. Recall that ambiguity is entirely about the expected growth rate μ. In this case, μ can be one of two values, whose difference is $0.05 \times h = 5 \times 10^{-6}$, a number that is dwarfed by the standard deviation. In fact, the ratio of the standard deviation to the expected growth range goes to infinity as h goes to zero, for any fixed range of μ values. In the following sections, these ideas are formalized and extended through CE approximations.

3 Single-Period Uncertainty

In a dynamic setting, this paper's CE approximations apply conditionally at each node of a binomial tree that in a continuous-time limit converges to either a Brownian motion or a Poisson process. Eliminating notation that is irrelevant for the argument, we focus on the single-period uncertainty from the perspective of any node on the binomial tree. This means there are only two states to consider, comprising the underlying state space $\{0, 1\}$.

We introduce two uncertainty models, corresponding to Brownian and Poissonian risks, parameterized by the time-period length $h \in (0, 1)$, which should be thought of as being small but not negligible. Terms of higher order than h are negligible. To express approximation errors that are negligible relative to h, we use the familiar little-oh notation: $o(h)$ represents some function $f : (0, 1) \to \mathbb{R}$ such that $f(h)/h \to 0$ as $h \to 0$. (Every instance of little oh can represent a potentially different function f.)

Each model is specified by a probability P and a random variable B, normalized so that

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

where $\mathbb{E}$ denotes expectation under P. More concretely, we consider the following specifications:

[4] The computer code is available at the author's web page: http://www.kellogg.northwestern.edu/faculty/skiadas/research/research.htm.

- **Brownian Uncertainty:**

$$\begin{cases} B(1) = +\sqrt{h} & \text{with probability} \quad P(1) = 1/2, \\ B(0) = -\sqrt{h} & \text{with probability} \quad P(0) = 1/2. \end{cases} \tag{3}$$

- **Poissonian Uncertainty:** For some function[5] $\epsilon(h) = o(h)$,

$$\begin{cases} B(1) = 1 - h - \epsilon(h) & \text{with probability} \quad P(1) = h + \epsilon(h), \\ B(0) = 0 - h - \epsilon(h) & \text{with probability} \quad P(0) = 1 - h - \epsilon(h). \end{cases} \tag{4}$$

Note that (P, B) is parameterized by h, even though the dependence on h is notationally suppressed.

In a continuous-time limit, h corresponds to a time infinitesimal dt and B corresponds to the infinitesimal increment dB of either a standard Brownian motion or a compensated Poisson process with unit arrival rate. Billingsley (1999) (Theorem 14.1 and Example 12.3) can be consulted for the rigorous statement and proof of the facts that, as h goes to zero, a random walk whose increments are i.i.d. copies of B converges to a standard Brownian motion in the case of specification (3) and a compensated Poisson process with unit arrival rate in the case of specification (4). We proceed under the assumption that (P, B) is specified by either (3) or (4).

Taken as given is a random variable $U : \{0, 1\} \to \mathbb{R}$ that is also parameterized by h and converges to the scalar U_0 as h goes to zero. We think of U_0 as the utility level of a given consumption plan at a given reference node of a binomial tree and U as the uncertain utility value one period ahead. In a recursive utility specification, U_0 is computed as a function of consumption at the reference node and a CE value of U. Our focus will be on approximations of the CE of U.

Since there are only two states, for any given h, we have the following *canonical decomposition*:

$$U = U_0 + \mu h + \Sigma B, \quad \text{where } \mu = \frac{\mathbb{E}U - U_0}{h} \text{ and } \Sigma = \frac{\mathbb{E}[BU]}{\mathbb{E}[B^2]} = \frac{\mathbb{E}[BU]}{h}. \tag{5}$$

(Adding an $o(h)$ term to U does not affect the results.) We assume that μ and Σ do **not** vary with h, and we refer to them, respectively, as the *drift* and *volatility* of U. (The continuous-time counterpart of (5) is the Ito decomposition $dU = \mu dt + \Sigma dB$.)

The ambiguity-averse CEs of interest involve multiple priors. In the Brownian model, a change of prior corresponds to a change of the drift of B, while keeping the volatility of B approximately the same. In the Poissonian model, a change of prior corresponds to a change in the arrival rate associated with B. Let us now look at these ideas more formally, paying special attention to the dependence of priors on h.

For any probability Q other than P, we write $\mathbb{E}^Q$ for the expectation operator under Q, and dQ/dP for the density of Q with respect to P, defined as the random variable that takes the

[5] While any such choice of $\epsilon(h)$ yields the same results, if (2) is to hold exactly, it must be the case that $\epsilon(h) = 0.5 - h - \sqrt{0.25 - h}$. We make this choice of $\epsilon(h)$ for exposition economy. Alternatively, one can carry out the analysis under the weaker restriction $\mathbb{E}[B^2] = h + o(h)$ without any change of substance.

value $Q(\omega)/P(\omega)$ at state $\omega \in \{0,1\}$. A probability Q defines the scalar ρ through the canonical decomposition of dQ/dP, which is computed analogously to (5) :

$$\frac{dQ}{dP} = 1 + \rho B, \quad \text{where} \quad \rho h = \mathbb{E}\left[B\frac{dQ}{dP}\right] = \mathbb{E}^{Q} B. \tag{6}$$

Definition 1 *A* prior *is a probability Q on $\{0,1\}$ that assigns positive mass to each state and is indexed by h so that $\rho \equiv \mathbb{E}^{Q}B/h$ does not vary with h (and therefore ρ is the drift of B under Q).*

The requirement that $\mathbb{E}^{Q}B = \rho h$ implies the concrete expressions:

$$Q(1) = 1 - Q(0) = \begin{cases} (1+\rho\sqrt{h})/2 & \text{in the case of Brownian uncertainty,} \\ (1+\rho)\, h + o(h) & \text{in the case of Poissonian uncertainty.} \end{cases}$$

Under the prior Q, a properly normalized random walk whose increments are i.i.d. copies of $B^{Q} \equiv B - \rho h = B - \mathbb{E}^{Q}B$ converges, as h goes to zero, to a standard Brownian motion in the case of Brownian uncertainty, and to a compensated Poisson process with arrival rate $1+\rho$ in the case of Poissonian uncertainty. A canonical decomposition with respect to B can always be transformed to one with respect to B^{Q}, for example, (5) becomes

$$U = U_0 + (\mu + \rho\Sigma)\, h + \Sigma B^{Q}. \tag{7}$$

The first two moments of B^{Q} under Q can be computed using only (2) and (6) as

$$\mathbb{E}^{Q}B^{Q} = 0 \quad \text{and} \quad \mathbb{E}^{Q}\left[(B^{Q})^2\right] = h + \rho\mathbb{E}\left[B^3\right] - (\rho h)^2. \tag{8}$$

In the Brownian model, $\mathbb{E}\left[B^3\right] = 0$ and therefore the second moment of B^{Q} under Q is $h + o(h)$, reflecting the fact that the drift but not the volatility of B or U are affected by a change of prior. In the Poissonian model, $\mathbb{E}\left[B^3\right] = h + o(h)$ and the second moment of B^{Q} under Q is $(1+\rho)\, h + o(h)$, reflecting the fact that the arrival rate under Q is $1+\rho$.

4 Certainty-Equivalent Approximations

This section defines the smooth ambiguity-averse CEs of interest and presents the main CE-approximation results. Some requisite terminology and notation are first introduced.

Fixed throughout is a constant $\ell \in [-\infty, 1)$ serving as a lower bound on consumption (possibly equal to $-\infty$). A *certainty equivalent (CE)* is an increasing and continuous function of the form $\upsilon : (\ell,\infty)^2 \to (\ell,\infty)$ with the property $\upsilon(x,x) = x$ for all $x \in (\ell,\infty)$. For each CE υ of interest, we will approximate $\upsilon(U)$, where U is specified in (5). We think of $\upsilon(U)$ as a conditional CE of the one-period-ahead continuation utility from the perspective of a reference node on a binomial tree.

We use the term *von Neumann-Morgenstern (vNM) index* to mean any strictly increasing, continuous function of the form $u : (\ell,\infty) \to \infty$. Differentiability assumptions are key for our approximations. For $n = 1, 2$, we use the notation

$$C^{n}_{\text{vNM}} = \text{ set of } n \text{ times continuously differentiable vNM indices.} \tag{9}$$

A pair of a prior Q and a vNM index u defines the *expected-utility (EU) CE* $\upsilon = u^{-1}\mathbb{E}^{Q}u$.

Our main focus will be the extent to which smooth ambiguity averse CEs can be approximated, up to an error term $o(h)$, by EU CEs. To correctly read this section's approximations as h goes to zero, it is important to keep in mind what quantities vary with h, as summarized in the following table, which includes preference parameters introduced below.

quantities that	do *not* vary with h	*can* vary with h
uncertainty model	states $0, 1$	P, B
priors	drift ρ, arrival rate $1+\rho$	corresponding probability Q on $\{0,1\}$
payoff structure	U_0, μ, Σ	U
certainty equivalents	$u,\ \varphi,\ \pi^1, \ldots, \pi^S$	$Q^1, \ldots, Q^S$, minimizing probability Q

4.1 Second-Order Expected-Utility CE

Consider an agent contemplating (so to speak) an EU CE with a given vNM index u, under each one of the priors $Q^1, \ldots, Q^S$. By Definition 1, for each $s \in \{1, \ldots, S\}$, there exists a constant ρ^s that is the drift of B under Q^s. The agent is uncertain about this drift. 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 $\pi^1, \ldots, \pi^S \in (0,1)$, where $\sum_s \pi^s = 1$. The *second-order EU CE* υ is defined in terms of a second vNM index φ by

$$\upsilon(U) = \varphi^{-1}\left(\sum_{s=1}^{S} \varphi\left(u^{-1}\mathbb{E}^{Q^s} u(U)\right)\pi^s\right), \quad U \in (\ell, \infty)^2. \tag{10}$$

The utility function $\varphi \circ \upsilon$ is of the type formulated by Klibanoff, Marinacci, and Mukerji (2005).

If $u = \varphi$, then $\upsilon = u^{-1}\mathbb{E}^{Q}u$, where Q is the compound prior: $Q = \sum_s Q^s \pi^s$. For $u \neq \varphi$, priors cannot be exactly compounded, but we have the following approximation result.

Theorem 2 (EU Approximation of Second-Order EU CE) *Suppose that* (P, B) *is given by either the Brownian specification* (3) *or the Poisson specification* (4)*, and that* υ *is the second-order EU CE* (10)*, for some*[6] $u \in C^2_{vNM}$ *and* $\varphi \in C^1_{vNM}$ *and priors* $Q^1, \ldots, Q^S$ *(see Definition 1). Then*

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

A rigorous proof can be found in Section A.1.2 of the online Appendix, along with more detailed small-risk approximations of an EU CE. The basic ideas can be outlined as follows. For each prior Q^s, the corresponding EU CE appearing as an argument of φ in (10) can be approximated, up to $o(h)$, by the expectation of U under Q^s minus a risk adjustment term that is proportional to h. (This is the familiar Arrow-Pratt approximation in the Brownian case. The approximation is extended to the Poissonian case in the Appendix.) It follows that in approximating υ with error

[6] In the Poisson case, it is enough to assume that $u \in C^1_{\mathrm{vNM}}$.

$o(h)$, it is sufficient to take a linear approximation of φ. This gets to the heart of the argument. In computing each EU CE inside φ in (10), we are facing a payoff U whose variance is of order h. In computing the outer CE in (10) we are facing a range of CE values whose variance is order $o(h)$, allowing the linearization of φ. Given the latter, the terms φ^{-1} and φ in the CE definition cancel out. Since φ becomes irrelevant as h goes to zero, we may as well set it equal to u, which is exactly the case that allows the compounding of priors.

4.2 Divergence CE

The second extension of an EU CE we consider corresponds to smooth divergence preferences, which are within the axiomatic setting of Maccheroni et. al. (2006a). We use the term *divergence index* to mean any strictly convex differentiable function of the form $\varphi : (0, \infty) \to \mathbb{R}$ such that

$$\varphi(1) = \varphi'(1) = 0 \quad \text{and} \quad \lim_{x \to \infty} \varphi'(x) = \infty. \tag{11}$$

A *divergence CE* is defined in terms of a vNM index u and a divergence index φ by

$$\upsilon(U) = \inf_{Q \in \Pi} u^{-1} \left(\mathbb{E}^Q u(U) + \mathbb{E}\varphi \left(\frac{dQ}{dP} \right) \right), \tag{12}$$

where Π denotes the set of all probabilities on $\{0, 1\}$ that place a nonzero mass to each state. (Recall that $\mathbb{E}$ denotes expectation relative to P, which is also an element of Π.)

As shown in the online Appendix, in the Brownian model, the infimum in (12) is achieved as a minimum for all sufficiently small $h > 0$, while in the Poissonian model, a minimum is achieved by some Q in Π if and only if (U_0, Σ) belongs to the set[7]

$$D = \{(U_0, \Sigma) \in (\ell, \infty) \times \mathbb{R} : U_0 + \Sigma > \ell \text{ and } u(U_0 + \Sigma) - u(U_0) < -\varphi'(0+)\}. \tag{13}$$

If either $\Sigma \leq 0$ or $\varphi'(0+) = -\infty$ (as in the entropic case below) then $(U_0, \Sigma) \in D$ for all $U_0 > \ell - \Sigma$.

Example 3 (Entropic CE) *The CE* (12) *is defined to be* entropic *if* $\varphi(y) = \theta(y \log y - y + 1)$. *In this case, a well-known identity states that* $\upsilon = (\psi \circ u)^{-1} \mathbb{E}(\psi \circ u)$, *where*

$$\psi(x) = \theta\left(1 - \exp\left(-\frac{x}{\theta}\right)\right). \tag{14}$$

Remark 9 of the online Appendix further elaborates on this case.

We just saw that an entropic CE is an EU CE. The following result investigates the extent to which a nonentropic smooth divergence CE is an EU CE. Analogously to (9), we write C^n_{div} for the set of n times continuously differentiable divergence indices.

[7] If $\varphi'(0+) > -\infty$ and the jump size Σ is positive and sufficiently high, a minimizing probability assigns zero probability to the positive jump.

Theorem 4 (EU Approximation of Divergence CE) *Suppose that* υ *is the divergence CE* (12), *for some*[8] $u \in C^2_{vNM}$ *and* $\varphi \in C^2_{div}$.

(a) If (P, B) *is given by the Brownian specification* (3), *then*

$$\upsilon(U) = (\psi \circ u)^{-1} \mathbb{E}(\psi \circ u)(U) + o(h), \tag{15}$$

where ψ *is defined by* (14) *with* $\theta = \varphi''(1)$.

(b) Suppose (P, B) *is given by the Poisson specification* (4). *There exist prior* Q *and* $\tilde{u} \in C^1_{vNM}$ *such that* $\upsilon(U) = \tilde{u}^{-1}\mathbb{E}^Q\tilde{u}(U) + o(h)$ *for all* $(U_0, \Sigma) \in D$ *if and only if* $Q = P$ *and the CE* υ *is entropic in the sense of Example 3.*

The theorem is proved in Sections A.1.4 and A.1.6 of the online Appendix. In the Brownian case, the basic idea is that it is sufficient to take a quadratic approximation of φ near unity, which is determined by $\theta = \varphi''(1)$, since $\varphi(1) = \varphi'(1) = 0$. The quadratic approximation of φ is therefore the same for the original φ as for the entropic φ of Example 3, leading to the claim of part (a). Example 3 states that every entropic CE is an EU CE. Part (b) of the preceding theorem provides a strong converse: *In the presence of Poisson uncertainty, if a smooth divergence CE can be approximated by any smooth expected-utility CE, then it must be entropic.* The argument in this case is too technical to grasp intuitively in simple terms but it is spelled out in the online Appendix. Also given in Theorem 11 of the Appendix is an explicit approximation of a non-entropic divergence CE under Poisson uncertainty, which is important in determining the functional form of continuous-time recursive utility with a divergence CE, a topic we turn to next.

5 Continuous-Time Recursive Utility

This paper's CE approximations map to functional forms of continuous-time recursive utility. Ideally, this should be demonstrated as the convergence of recursive utility along sequences of binomial trees, but such an approach is highly technical and currently incomplete even for the case of an EU CE.[9] Applied theory papers using continuous-time recursive utility take as a starting point the continuous-time version of Kreps and Porteus (1978) utility formulated by Duffie and Epstein (1992) using a compelling but heuristic argument. In this section, we take a similar approach to establish continuous-time recursive utility forms corresponding to the smooth ambiguity-averse CEs of interest. Section A.2 of the online Appendix provides additional details and generality.

We put the Brownian and Poisson cases together in one model. We assume that the information tree (defined on some underlying state space) is generated over the time set $[0, T]$ by two stochastic processes,[10] forming the column vector $B = (B^1, B^2)'$. Taken as given is an underlying probability

[8] In the Poisson case of part (b) it is enough to assume that $u \in C^1_{\text{vNM}}$ and $\varphi \in C^1_{\text{div}}$, with the same proof.
[9] The recent working paper by Kraft and Seifried (2011) is a first step in addressing this gap in the literature.
[10] The expanded presentation in the online Appendix allows for more than two sources of risk.

P on the set of time-T events, with corresponding expectation operator $\mathbb{E}$. Relative to P, the processes B^1 and B^2 are stochastically independent, B^1 is a standard Brownian motion and B^2 is a compensated Poisson process with unit arrival rate (which can be taken to be the definition of the unit of time).

Let us fix a reference consumption plan c, with corresponding utility process U. (Naturally, c and U are adapted to the given information tree.) For every $t < T$, c_t represents a time-t consumption rate. As in Section 2, we set $U_T = c_T$ (allowing the measurement of utility in terms of equivalent perpetuities). We assume the recursive-utility form (1) , but with u_δ denoting a general smooth vNM index (not necessarily a constant-EIS one). The continuous-time version of this utility recursion can be heuristically stated as[11]

$$u_\delta\left(U_{t-}\right) = \left(1 - e^{-\beta dt}\right) u_\delta\left(c_t\right) + e^{-\beta dt} u_\delta\left(\upsilon_t\left(U_{t+dt}\right)\right), \tag{16}$$

where U_{t-} denotes the time-t utility value just prior to any time-t jump, and dt is a time infinitesimal, analogous to the quantity h of the discrete-time analysis.

Analogously to the canonical decomposition (5) , suppose the utility dynamics are

$$dU_t = \mu_t dt + \Sigma_t dB_t. \tag{17}$$

Here $\Sigma_t = \left(\Sigma_t^1, \Sigma_t^2\right)$, where Σ_t^1 is the volatility of Brownian component and Σ_t^2 is the time-t jump size, conditionally on there being a time-t jump. Terms that are order $o\left(dt\right)$ are treated as zero (for example, $e^{-\beta dt} = 1 - \beta dt$). The CE approximations of interest can all be stated as

$$\upsilon_t\left(U_{t+dt}\right) = \mathbb{E}_{t-}\left[U_{t+dt}\right] - \mathcal{A}\left(U_t, \Sigma_t\right) dt = U_{t-} + \left(\mu_t - \mathcal{A}\left(U_t, \Sigma_t\right)\right) dt, \tag{18}$$

where the $\mathcal{A}\left(U_t, \Sigma_t\right) dt$ term represents a risk/ambiguity-aversion adjustment to the risk-neutral CE under prior P. Now substitute (18) into (16) , take a first-order Taylor expansion of u_δ around U_{t-}, solve for μ_t and insert the resulting expression back into the utility dynamics (17) to find

$$dU_t = -\left(f\left(c_t, U_t\right) - \mathcal{A}(U_t, \Sigma_t)\right) dt + \Sigma_t dB_t, \quad U_T = c_T, \tag{19}$$

where, as in equation (27) of Duffie and Epstein (1992), $f\left(c, \upsilon\right) = \beta\left(u_\delta\left(c\right) - u_\delta\left(\upsilon\right)\right)/u_\delta'\left(\upsilon\right).$

Equation (19) is a so-called backward stochastic differential equation to be solved jointly in $\left(U, \Sigma\right)$. This is a fixed-point problem, whose solution requires regularity conditions[12] on $\left(f, \mathcal{A}\right)$ and the consumption plan c. Since f is entirely unrelated to the CE specification, equations (18) and (19) make transparent how a small-risk CE approximation maps to a corresponding continuous-time

[11] The additive structure in (16) is not important—a more general intertemporal aggregator is assumed in the Appendix.

[12] Corresponding existence and uniqueness results were obtained by Pardoux and Peng (1990), Duffie and Epstein (1992), and many others since, albeit under assumptions that are violated in common homothetic applications, including the continuous-time version of Epstein-Zin-Weil utility (unless log-consumption is assumed to be bounded). Existence and uniqueness results on the latter were developed by Duffie and Lions (1996) and Schroder and Skiadas (1999), results that still require generalization, for example, to include jumps.

recursive utility. We conclude by specifying the functional form of $\mathcal{A}$ corresponding to the smooth CEs of interest, referring to the online Appendix for detailed derivations.

Suppose first that υ is the smooth second-order EU CE of Theorem 2, with the underlying probability selected to be the compound prior: $P = \sum_s \pi^s Q^s$. Approximation (18) in this case is the same as for the EU CE $\upsilon_t = u^{-1}\mathbb{E}_t u$. The corresponding risk-adjustment term is

$$\mathcal{A}(U_t, \Sigma_t) = \frac{a^u(U_t)}{2}\left(\Sigma_t^1\right)^2 + \Sigma_t^2 - \frac{u\left(U_0 + \Sigma_t^2\right) - u(U_0)}{u'(U_0)}, \tag{20}$$

where $a^u = -u''/u'$ denotes the coefficient of absolute risk aversion of u. The corresponding continuous-time utility (19) is Duffie-Epstein utility, extended to include Poisson jumps.

Finally, suppose that υ is the smooth divergence CE of Theorem 4. The corresponding risk-adjustment term is

$$\mathcal{A}(U_t, \Sigma_t) = \frac{a^{\psi\circ u}(U_t)}{2}\left(\Sigma_t^1\right)^2 + \Sigma_t^2 - \frac{\varphi^*\left(u\left(U_0 + \Sigma_t^2\right) - u(U_0)\right)}{u'(U_0)}, \tag{21}$$

where $\varphi^*(x) = \min_{y\in(0,\infty)}\{\varphi(y) + xy\}$ and ψ is the exponential form (14) with $\theta = \varphi''(1)$.

Note that φ is entropic if and only if $\varphi^* = \psi$, in which case expression (21) can be obtained from (20) after replacing u with $\psi \circ u$, as it should, according to Example 3. A nonentropic φ enters the first (Brownian) term of (21) only through the local variable $\theta = \varphi''(1)$, while the second (Poissonian) term depends on the global structure of φ, thus providing some flexibility to adjust aversion toward Brownian and Poisson risks separately. Skiadas (forthcoming) gives an example of how this can be accomplished with a concrete parameterization of φ.

6 Concluding Remarks

This paper's arguments hinge on the assumption that the parameter α of Section 2 and more generally the function φ representing ambiguity aversion remain constant as the frequency is taken to infinity, similarly to the way the risk aversion parameter for Epstein-Zin-Weil utility and more generally the vNM index u of Kreps-Porteus utility remain constant as one transitions from discrete time to the Duffie-Epstein continuous-time limit. (Discrete-time recursive utility with a second-order EU CE reduces to Kreps-Porteus utility exactly when $\varphi = u$.) The general idea is that φ should capture ambiguity aversion as a fixed aspect of preferences that applies in every uncertainty environment, just as u captures risk aversion in the Kreps-Porteus specification. Recall that we have normalized utility to correspond to the payment rate of an equivalent perpetuity. This device allows us to think of the continuation utility one period ahead as a state-contingent perpetuity that can be embedded in the same static setting, no matter what the frequency, thus anchoring risk aversion and ambiguity aversion.

Responding to an earlier version of this paper, Hansen and Sargent (2011) verified the irrelevance of a fixed φ in a specific continuous-time context and proposed an alternative frequency-dependent parameterization of φ, finely tuned to preserve the effect of ambiguity aversion in the Brownian limit of recursive utility with a second-order EU CE. Their parameterization implies that ambiguity aversion goes to infinity as the frequency goes to infinity, thus resulting in a different continuous-time limit, which, as the authors point out, is not smooth in the level of the continuation utility. In other words, the limiting conditional CE over infinitesimal Brownian risks implied by the Hansen-Sargent formulation is a very different CE type than the smooth second-order EU CE that has been analyzed in this paper. Moreover, since the Hansen-Sargent CE is effectively defined only over infinitesimal risks (defined in the limit), it is not clear how it can be related to preferences over larger risks.

Another modification of second-order expected utility motivated by an early version of the present paper was proposed by Gindrat and Lefoll (2010). Their proposal can be thought of as amplifying the order of magnitude of the instantaneous drift to account for ambiguity.

In their current form, these alternative formulations seem driven by the desire to preserve a specific formalization of smooth ambiguity aversion in the continuous-time limit, rather than any compelling decision-theoretic foundations. The fact remains that the most direct continuous-time interpretation of a smooth CE derived from a static model of second-order expected utility or divergence preferences is not distinguishable from an expected-utility CE under Brownian uncertainty (as well as under Poissonian uncertainty in the case of second-order expected utility).

References

Bewley, T. (1986): "Knightian Decision Theory, Part I," Cowles Foundation Discussion Paper no. 807, Yale University.

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

Caballero, R., and A. Krishnamurthy (2008): "Collective Risk Management in a Flight to Quality Episode," *Journal of Finance*, 63, 2195–2236.

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

Collard, F., S. Mukerji, K. Sheppard, and J.-M. Tallon (2011): "Ambiguity and the Historical Equity Premium," working paper, dept. of Economics, University of Oxford, U.K.

Dixit, A. K., and R. S. Pindyck (1994): *Investment under Uncertainty*. Princeton Univ. Press, Princeton, New Jersey.

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

Duffie, D., and P.-L. Lions (1996): "PDE Solutions of Stochastic Differential Utility," *Journal of Mathematical Economics*, 21, 577–606.

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

Epstein, L., and M. Schneider (2003): "Recursive Multiple Priors," *Journal of Economic Theory*, 113, 1–31.

Epstein, L. G., and S. E. Zin (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *The Journal of Political Economy*, 99, 263–286.

Gilboa, I. (2009): *Theory of Decision Making under Uncertainty*. Cambridge University Press, New York.

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

Gindrat, R., and J. Lefoll (2010): "Smooth Ambiguity Aversion and the Continuous-Time Limit," working paper, University of Geneva and SFI, Switzerland.

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

——— (2011): "Robustness and Ambiguity in Continuous Time," *Journal of Economic Theory*, 146, 1195–1223.

Ju, N., and J. Miao (2012): "Ambiguity, Learning, and Asset Returns," *Econometrica*, 80, 559–591.

Keynes, J. M. (1937): "The General Theory of Employment," *The Quarterly Journal of Economics*, 51, 209–223.

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

——— (2009): "Recursive Smooth Ambiguity Preferences," *Journal of Economic Theory*, 144, 930–976.

Knight, F. H. (1921): *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston and New York.

Kraft, H., and F. T. Seifried (2011): "Stochastic Differential Utility as the Continuous-Time Limit of Recursive Utility," working paper, Goethe University, Frankfurt and University of Kaiserslautern.

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

Maccheroni, F., M. Marinacci, and D. Ruffino (forthcoming): "Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis," *Econometrica*.

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.

Mehra, R., and E. C. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161.

Nau, R. F. (2006): "Uncertainty Aversion with Second-Order Probabilities and Utilities," *Management Science*, 52, 136–145.

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

Protter, P. E. (2004): *Stochastic Integration and Differential Equations*. Springer Verlag, New York, second edn.

Schmeidler, D. (1989): "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57, 571–587.

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

Silver, N. (2012): *The Signal and the Noise: Why So Many Predictions Fail, but Some Don't*. Penguin Press HC.

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

——— (2009): *Asset Pricing Theory*. Princeton Univ. Press, Princeton, NJ.

——— (forthcoming): "Scale-invariant asset pricing and consumption/portfolio choice with general attitudes toward risk and uncertainty," *Mathematics and Financial Economics*.

Weil, P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.

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

Costis Skiadas

April 2013

A Online Appendix

This Appendix proves and extends results presented in the main paper. There are two main sections. The first section covers the theory of single-period CE approximations. The second section presents a corresponding theory of continuous-time recursive utility. References that are not part of the main paper can be found at the back of this Appendix.

A.1 Single-Period CE Approximations

Throughout this section, we adopt the single-period stochastic setting of Section 3. In particular, recall Definition 1 of a *prior* Q, whose parameterization by h implies a constant drift ρ of B under Q. Recall also that, independently of the choice of a prior, the risk source B takes the values $+\sqrt{h}$ or $-\sqrt{h}$ in the Brownian case, and $1 - h - \epsilon(h)$ or $0 - h - \epsilon(h)$ in the Poissonian case (where $\epsilon(h) = o(h)$). If ρ is the drift of B under prior Q, then a properly normalized random walk whose increments are given by B converges, as $h \downarrow 0$, to either Brownian motion with drift ρ or a compensated Poisson process with arrival rate $1 + \rho$.

We derive smooth CE approximations of the form

$$v(U) = U_0 + (\mu + \rho\Sigma - \mathcal{A}(U_0, \Sigma, \rho))\, h + o(h) = \mathbb{E}^Q U - \mathcal{A}(U_0, \Sigma, \rho)\, h + o(h), \tag{22}$$

where U_0, μ and Σ refer to the canonical representation (5). (The second equation in (22) follows by applying $\mathbb{E}^Q$ to equation (7).) If Q is given the interpretation of beliefs, then $\mathcal{A}(U_0, \Sigma, \rho)\, h$ represents an approximate adjustment to the risk-neutral CE due to risk/ambiguity aversion. This section's approximation results specify the function $\mathcal{A}$ for the CEs of interest, starting with the benchmark case of a smooth EU CE in subsection A.1.1. The latter is used in subsection A.1.2 to prove the irrelevance of ambiguity aversion as h goes to zero for a smooth second-order EU CE. The remaining subsections develop approximations for the case of a smooth divergence CE, which are again compared to the EU case.

The essential tool in all approximations that follow is Taylor's theorem. For easy reference, we quickly review[13] here the approximation error estimates we will be using. The reader may wish to skip the following paragraph and refer back to it as necessary.

[13] This material is of course standard and can be found, for example, in Apostol, Calculus, 2nd ed., Wiley 1967.

Consider any positive integer n (we'll only need $n = 1$ or 2), any interval $[-\varepsilon, \varepsilon] \subseteq \mathbb{R}$, where $\varepsilon > 0$, and any n times continuously differentiable real-valued function f on an open interval that includes $[-\varepsilon, \varepsilon]$. Let $f^{(i)}$ denote the i^{th} derivative of f, with $f^{(0)} = f$. Let also f_n be the n-degree polynomial such that $f^{(i)}(0) = f_n^{(i)}(0)$ for $i = 0, \ldots, n$. (We will only need $f_1(x) = f(0) + f'(0)x$ and $f_2(x) = f_1(x) + f''(0)x^2/2$.) Then

$$f = f_n + R_n, \quad \text{where } R_n(x) = \int_0^x \left(f^{(n)}(t) - f^{(n)}(0) \right) \frac{(x-t)^{n-1}}{(n-1)!} dt. \tag{23}$$

(This is easily shown by induction in n, using the identity $R_{n-1}(x) = f^{(n)}(0)x^n/n! + R_n(x)$, obtained by applying integration by parts.) It follows that there exists a continuous function $r_n : [-\varepsilon, \varepsilon] \to [0, \infty)$ such that[14]

$$r_n(0) = 0 \quad \text{and} \quad |R_n(x)| \leq r_n(x)|x|^n \text{ for all } x \in [-\varepsilon, \varepsilon]. \tag{24}$$

A.1.1 EU Approximations

We begin with approximation (22) for the case of a smooth EU CE $v = u^{-1}\mathbb{E}^Q u$, both as a benchmark to which other CE approximations are compared and as an intermediate case used in later proofs.

For Brownian small risks, the approximation is a variant to the classic results of Arrow (1965, 1970) and Pratt (1964) applied to the payoff (5), with attention to how the approximation error depends on a change of prior. Brownian approximations are expressed in terms of the Arrow-Pratt coefficient of absolute risk aversion associated with the vNM index u, which we denote

$$a^u = -\frac{u''}{u'}. \tag{25}$$

An Arrow-Pratt approximation is not valid for small Poissonian risks, however, because third and higher moments of such risks are not negligible. We establish an analogous approximation for Poissonian small risks, in which the role of a^u is assumed by the function

$$A^u(U_0, \Sigma) = \Sigma - \frac{u(U_0 + \Sigma) - u(U_0)}{u'(U_0)}. \tag{26}$$

Whereas $a^u(U_0)$ 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,

[14] Note that under the stronger assumption that f is $n+1$ times differentiable, integration by parts allows us to restate the expression for $R_n(x)$ in (23) in the more familiar form

$$R_n(x) = \int_0^x f^{(n+1)}(t) \frac{(x-t)^n}{n!} dt.$$

If $f^{(n+1)}$ is bounded on $[-\varepsilon, \varepsilon]$ (for example, if it is continuous), then there exists a constant K such that $|R_n(x)| \leq Kx^{n+1}$ for all $x \in [-\varepsilon, \varepsilon]$. In other words, we can set $r_n(x) = Kx$ in (24).

and the value U_0 otherwise.[15] Like a^u, the function A^u is a measure of the curvature of u. The inequality $A^u \geq 0$ is equivalent to the gradient inequality for u and therefore the concavity of u. It is also worth noting that if $u \in C^2_{\text{vNM}}$, then $A^u(U_0, \Sigma) = a^u(U_0)\Sigma^2/2 + o(\Sigma^2)$, as can be seen by taking a second-order Taylor series expansion of u in (26). This gives a sense of consistency of the approximations that follow, but here Σ is not required to be small and u need not have a second derivative when only Poissonian risk is involved.

The general form of the smooth EU approximations of interest is given in the following result.

Theorem 5 (EU Approximations) *Suppose Q is a prior and ρ is the drift of B under Q.*

(a) (Brownian risk) If (P, B) is defined by (3), then for all $u \in C^2_{vNM}$,

$$u^{-1}\mathbb{E}^Q u(U) = U_0 + \left(\mu + \rho\Sigma - \frac{1}{2}a^u(U_0)\Sigma^2\right)h + o(h). \tag{27}$$

(b) (Poissonian risk) If (P, B) is defined by (4), then for all $u \in C^1_{vNM}$,

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

The preceding approximations are sufficient for the proof of Theorem 2 in the following subsection. For the analysis of a smooth divergence CE, however, we need more refined estimates of the approximation error. These are provided by two lemmas below, covering the Brownian case and the Poissonian case, respectively. Theorem 5 is a corollary of these two lemmas.

From here on, we impose a normalization that entails no loss of generality and simplifies the exposition. Suppose that u and $\tilde{u}$ are two vNM indices related by a positive affine transformation, meaning that $\tilde{u} = \alpha u + \beta$ for some $\alpha \in (0, \infty)$ and $\beta \in \mathbb{R}$. Then $u^{-1}\mathbb{E}^Q u = \tilde{u}^{-1}\mathbb{E}^Q \tilde{u}$, $a^u = a^{\tilde{u}}$ and $A^u = A^{\tilde{u}}$. It is then clear that to prove Theorem 5, it is sufficient to prove the result after replacing $u(\cdot)$ with $(u(\cdot) - u(U_0))/u'(U_0)$. We therefore proceed under the simplifying normalization:

$$u(U_0) = 0 \quad \text{and} \quad u'(U_0) = 1. \tag{28}$$

Let us also express the canonical decomposition of U as $U = U_0 + \Delta(h)$, where

$$\Delta(h) = \mu h + \Sigma B = (\mu + \rho\Sigma)h + \Sigma B^Q, \tag{29}$$

for any prior Q and corresponding drift ρ of B. Fixing an $\varepsilon \in (0,1)$ such that $U_0 - \varepsilon > \ell$, throughout the rest of the single-period analysis, we assume that $h \in (0, \bar{h})$, where $\bar{h} > 0$ is small enough so that $h \in (0, \bar{h})$ implies $\Delta(h) \in (-\varepsilon, +\varepsilon)$. Note that $\Delta(h)$, ε and $\bar{h}$ do not depend on Q or ρ.

A more detailed error bound for a Brownian EU CE approximation follows. Note that the parameters K_0, K_1 and b apply uniformly across priors.

[15] The reader interested in a graphical representation of A^u can proceed as follows. On the plane, draw the graph of the concave function u and consider the points $p_0 = (U_0, u(U_0))$ and $p_\Sigma = (U_0 + \Sigma, u(U_0 + \Sigma))$. Let p be the point where the horizontal line through p_Σ meets the line that is tangent to the graph of u at p_0. Then $A^u(U_0, \Sigma)$ is the distance between p_Σ and p.

Lemma 6 (Brownian EU Approximation) *Suppose* (P, B) *is defined by* (3) *and* $u \in C^2_{vNM}$ *satisfies the normalization* (28). *Then there exist constants* $K_0, K_1 \in \mathbb{R}$ *and a continuous function* $b : [0, \bar{h}] \to \mathbb{R}$ *that vanishes at zero such that for every prior* Q, *if* ρ *is the drift of* B *under* Q, *then we can write*

$$\mathbb{E}^Q u(U) = \left(\mu + \rho\Sigma - \frac{1}{2} a^u(U_0)\Sigma^2\right) h + \delta(\rho, h) h, \tag{30}$$

with

$$|\delta(\rho, h)| \leq |K_0 + K_1 \rho| h + b(h), \quad h \in (0, \bar{h}). \tag{31}$$

Proof. Applying the Taylor approximation (23) with $f(x) = u(U_0 + x)$ and $n = 2$, we have

$$u(U) = \Delta(h) - \frac{1}{2} a^u(U_0)\Delta(h)^2 + R_2(\Delta(h)). \tag{32}$$

Recall that $\mathbb{E}^Q B^Q = 0$ and $\mathbb{E}^Q\left[(B^Q)^2\right] = h - (\rho h)^2$ by (8). Using these facts and the second expression for $\Delta(h)$ in (29), we compute

$$\mathbb{E}^Q \Delta(h) = (\mu + \rho\Sigma) h \quad \text{and} \quad \mathbb{E}^Q[\Delta(h)^2] = \Sigma^2 h + \left(\mu^2 + 2\mu\Sigma\rho\right) h^2. \tag{33}$$

Therefore, applying $\mathbb{E}^Q$ to both sides of (32), we obtain expression (30), where

$$\delta(\rho, h) = -\frac{1}{2} a^u(U_0)\left(\mu^2 + 2\mu\Sigma\rho\right) h + \frac{1}{h}\mathbb{E}^Q R_2(\Delta(h)). \tag{34}$$

We apply the bound (24) to the last term. Since the possible values of $\Delta(h)$ are $\mu h \pm \Sigma\sqrt{h}$,

$$r_2(\Delta(h)) \leq r(h) \equiv \max\left\{r_2\left(\mu h + \Sigma\sqrt{h}\right), r_2\left(\mu h - \Sigma\sqrt{h}\right)\right\},$$

resulting in the bound $\mathbb{E}^Q |R_2(\Delta(h))| \leq r(h)\,\mathbb{E}^Q[\Delta(h)^2]$. Using this bound and (33) in (34), we obtain

$$|\delta(\rho, h)| \leq \left|\mu^2 + (2\mu\Sigma)\rho\right| \left(\frac{1}{2}|a^u(U_0)| + r(h)\right) h + \Sigma^2 r(h).$$

Note that r is continuous (and therefore bounded), vanishes at zero and does not depend on ρ. ■

The analogous result for the Poissonian model (4) is given below. While in Section 3 the function $\epsilon(h)$ is restricted as discussed in footnote 5, the only assumption on $\epsilon(h)$ used here is that for $\bar{h} > 0$ small enough, $\epsilon(h)/h$ is a continuous function on $[0, \bar{h}]$ that vanishes at zero. Note also that δ_0 and δ_1 do not depend on the choice of ρ.

Lemma 7 (Poissonian EU Approximation) *Suppose* (P, B) *is defined by* (4) *and* $u \in C^1_{vNM}$ *satisfies the normalization* (28). *For* $\bar{h} > 0$ *small enough, there exist continuous functions* $\delta_0, \delta_1 :$ $[0, \bar{h}] \to \mathbb{R}$ *that vanish at zero such that for every prior* Q, *if* ρ *is the drift of* B *under* Q, *then*

$$\mathbb{E}^Q u(U) = (\mu + \rho\Sigma - A^u(U_0, \Sigma)(1 + \rho)) h + (\delta_0(h) + \delta_1(h)(1 + \rho)) h, \quad h \in (0, \bar{h}). \tag{35}$$

Proof. Given the normalization (28), $A^u(U_0,\Sigma)=\Sigma-u(U_0+\Sigma)$. Let

$$\alpha(h)=(\mu-\Sigma)h-\Sigma\epsilon(h). \tag{36}$$

The distribution of U under Q is

$$U=\begin{cases} U_0+\Sigma+\alpha(h), & \text{with } Q\text{-probability } (1+\rho)h+\epsilon(h), \\ U_0+\alpha(h), & \text{with } Q\text{-probability } 1-(1+\rho)h-\epsilon(h), \end{cases} \tag{37}$$

We compute the expectation $\mathbb{E}^Q u(U)$, using the first-order Taylor expansions

$$\begin{aligned} u(U_0+\Sigma+\alpha(h)) &= u(U_0+\Sigma)+u'(U_0+\Sigma)\alpha(h)+R^1(\alpha(h)), \\ u(U_0+\alpha(h)) &= \alpha(h)+R^0(\alpha(h)), \end{aligned}$$

where $|R^\omega(\alpha)|\leq r^\omega(\alpha)|\alpha|$, $\omega\in\{0,1\}$, for continuous functions r^ω that vanish at zero (as in (24) with $n=1$). Note that neither $\alpha(h)$ nor $r^\omega(\alpha)$ depend on the choice of ρ. The rest of the proof is just a matter of explicitly computing $\mathbb{E}^Q u(U)$ using the distribution (37), applying the above first-order approximation to each term and reorganizing the resulting expression to bring it the form of equation (35). ■

Proof of Theorem 5. As already noted, it is sufficient to prove the result under the normalization (28), which implies that $u^{-1}(0)=U_0$ and $u^{-1\prime}(0)=1$. Part (a) follows from Lemma 6 by applying u^{-1} on both sides of expression (30), followed by a first-order Taylor series expansion of u^{-1} around zero. The analogous argument proves part (b) using Lemma 7.

A.1.2 Proof of Theorem 2

Theorem 2 on the EU approximation of a second-order EU CE is essentially a corollary of Theorem 5. Equation (27) applied to each Q^s implies that we can approximate the second-order EU CE, up to $o(h)$, by using a linear approximation of φ, leading to the cancellation of the terms φ^{-1} and φ in the CE definition. For more detail, let us first consider the Brownian case. Equation (27) and a first-order Taylor series expansion of φ around U_0 imply that

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

Therefore,

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

Note that $\rho=\mathbb{E}^Q B/h$, where $Q=\sum_s Q^s\pi^s$. Taking a first-order Taylor approximation of φ^{-1} around $\varphi(U_0)$ and using Theorem 5 once again, we conclude:

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

The proof for the Poissonian case is essentially the same, replacing the risk-adjustment term $a^u(U_0)\Sigma^2/2$ with $A^u(U_0,\Sigma)(1+\rho^s)$, for each prior Q^s.

A.1.3 Divergence CE: A Preliminary Lemma and Examples

Recall that a divergence CE is defined in terms of a vNM index u, a reference prior Q and a divergence index φ by (12), where Π denotes the set of all probabilities that assign a positive mass to each state. We focus on the case in which the infimum is achieved by some $Q \in \Pi$, a case that is fully characterized in Lemma 8 below. We will see later that the infimum is always achieved as a minimum for sufficiently small Brownian risks, and it is always achieved for Poissonian risks with negative jumps. We let $\Omega = \{0, 1\}$ denote the state space needed for the current application. It's worth noting, however, that the following lemma remains true for any finite Ω and underlying probability $P \in \Pi$, with a straightforward extension of the given proof (where terms involving state one become summations over all states other than state zero).

Lemma 8 *Let $U_{\max} = \max_{\omega \in \Omega} U_\omega$. The infimum in* (12) *is achieved as a minimum by some $Q \in \Pi$ if and only if*

$$\mathbb{E}\varphi'^{-1}\left(\varphi'(0+) + u(U_{\max}) - u(U)\right) < 1, \tag{38}$$

in which case the minimizing Q is given by $dQ/dP = \varphi'^{-1}(\alpha - u(U))$, where the scalar α uniquely solves

$$\mathbb{E}\varphi'^{-1}(\alpha - u(U)) = 1 \quad \textit{and} \quad \alpha > \varphi'(0+) + u(U_{\max}). \tag{39}$$

Proof. We relabel the states if necessary so that $U(0) = U_{\max}$. A probability $Q \in \Pi$ is identified with a $q \in (0, 1)$, where $Q(0) = 1 - q$ and $Q(1) = q$. In particular, P is identified with $p \in (0, 1)$. Then $V(q) \equiv \mathbb{E}^Q u(U) + \mathbb{E}\varphi(dQ/dP)$ can be written as

$$V(q) = u(U_{\max}) + q\{u(U(1)) - u(U_{\max})\} + p\varphi\left(\frac{q}{p}\right) + (1-p)\varphi\left(\frac{1-q}{1-p}\right).$$

The strictly convex function V is minimized at $q \in (0, 1)$ if and only if its derivative a q vanishes, a condition that is easily shown to be equivalent to

$$q = p\varphi'^{-1}(\alpha - u(U(1))), \quad \text{where} \quad \alpha = u(U_{\max}) + \varphi'\left(\frac{1-q}{1-p}\right). \tag{40}$$

Since $\varphi'((1-q)/(1-p)) > \varphi'(0+)$, condition (40) is easily seen to imply (39) and therefore (38), which proves the "only if" part.

Conversely, suppose that condition (38) is satisfied and define the function

$$f(\alpha) = \mathbb{E}\varphi'^{-1}(\alpha - u(U)), \quad \alpha > \varphi'(0+) + u(U_{\max}).$$

Clearly, f is strictly increasing and continuous. Because of condition (38), $f(\alpha)$ takes values below one as α approaches $\varphi'(0+) + u(U_{\max})$. On the other hand, by the definition of a divergence index, $\phi'(\infty) = \infty$, and therefore $f(\alpha)$ takes values greater than one for sufficiently large α. This proves the existence of a unique value of $\alpha > \varphi'(0+) + u(U_{\max})$ such that $f(\alpha) = 1$, which is exactly condition (39). By construction, $\alpha - u(U) > \varphi'(0+)$ and therefore the vector q is well-defined by the first equation of condition (40). This definition of q and the identities $f(\alpha) = 1$ and $U_{\max} = U(0)$ can be easily shown to imply the last equation of condition (40). This proves condition (40), which, as we have seen, is sufficient for the optimality of the prior defined by q. ■

Remark 9 (Entropic CE) *The entropic CE was defined in Example 3 as a divergence CE with divergence index* $\varphi(y) = \theta(y \log y - y + 1)$ *(where, necessarily,* $\theta = \varphi''(1)$*). This divergence index is related to the function* ψ *defined in* (14) *by the convex duality:* $\varphi(y) = \max_x\{\psi(x) - xy\}$ *and* $\psi(x) = \min_y\{\varphi(y) + xy\}$*. In this case,* $\varphi'(0+) = -\infty$ *and therefore condition* (38) *is satisfied for any* U*. The corresponding minimizing probability* Q *is given by Lemma 8 through the density* $dQ/dP = \exp(-u(U)/\theta) / \mathbb{E}\exp(-u(U)/\theta)$*. Computing the corresponding minimum proves that an entropic divergence CE coincides with the EU CE* $\upsilon = (\psi \circ u)^{-1} \mathbb{E}(\psi \circ u)$*, which is the claim of Example 3. This is an well-known identity, appearing in Donsker and Varadhan (1975).*

Example 10 (Quadratic Divergence) *Suppose that* $\varphi(y) = (\theta/2)(y-1)^2$ *for some* $\theta \in (0, \infty)$*. This case is covered by Theorem 24 of Maccheroni, Marinacci, and Rustichini (2006a). We briefly review the main conclusion as a corollary of Lemma 8. Assume the validity of Condition* (38)*, which in this context reduces to* $u(U) - \mathbb{E}u(U) < \theta$*. Equation* (39) *results in* $\alpha = \mathbb{E}[u(U)]$*, and therefore the minimizing prior* Q *is given by* $dQ/dP = 1 - (u(U) - \mathbb{E}u(U))/\theta$*. Calculating the minimum using the preceding expression results in*

$$\upsilon(U) = u^{-1}\left(\mathbb{E}u(U) - \frac{1}{2\theta} Var[u(U)]\right).$$

The same condition $u(U) - \mathbb{E}u(U) < \theta$ *guarantees that* U *is valued within the range where the right-hand side defines a strictly increasing function.*

A.1.4 Proof of Theorem 4(a)

In this subsection, we prove part (a) of Theorem 4, providing an EU approximation of a smooth divergence CE for Brownian small risks. For a preliminary informal indication as to why the result holds, we use restrictions (11) on a divergence index φ in a second-order Taylor expansion to compute

$$\mathbb{E}\varphi\left(\frac{dQ}{dP}\right) = \mathbb{E}\varphi(1 + \rho B) = \frac{\theta}{2}\rho^2 h + o(h). \tag{41}$$

To first order, therefore, it does not matter whether we use the smooth divergence index φ or a corresponding entropic divergence index, as given in Example 3, with $\theta = \varphi''(1)$. Example 3 expresses the entropic divergence CE as an EU CE, implying that the original CE with divergence index φ can be approximated by an EU CE. The rigorous proof that follows is quite a bit more elaborate than this heuristic argument suggests, because we have to justify the interchange of the operations of minimization and approximation.

In order to avoid redundant notation, we adopt the vNM index normalization (28) throughout. If u and $\tilde{u}$ are two vNM indices such that that $\tilde{u} = \alpha u + \beta$ for some $\alpha \in (0, \infty)$ and $\beta \in \mathbb{R}$, and φ

and $\tilde{\varphi}$ are divergence indices related by $\tilde{\varphi} = \alpha\varphi$, then the pair (u, φ) defines the same divergence CE as the pair $(\tilde{u}, \tilde{\varphi})$. It is then easy to see that to prove Theorem 4(a), it is sufficient to prove it after replacing $u(\cdot)$ with $(u(\cdot) - u(U_0))/u'(U_0)$ and $\varphi(\cdot)$ with $\varphi(\cdot)/u'(U_0)$ This shows that our adoption of normalization (28) entails no loss in generality.

As in subsection A.1.1, we adopt the notation (29), we fix an $\varepsilon \in (0,1)$ such that $U_0 - \varepsilon > \ell$, and we choose $\bar{h} > 0$ small enough so that $h \in \left(0, \bar{h}\right)$ implies $\Delta(h) \in (-\varepsilon, +\varepsilon)$. It is straightforward to check that in the Brownian context, condition (38) is satisfied for all sufficiently small h. We further decrease $\bar{h}$ if necessary so that, by Lemma 8, a minimizing prior exists for every $h \in \left(0, \bar{h}\right)$.

Let us fix any prior Q and corresponding drift $\rho = h^{-1}\mathbb{E}^Q B$. In the Brownian model assumed here, $Q(1) = 1 - Q(0) = \left(1 + \rho\sqrt{h}\right)/2$. Therefore,

$$\begin{aligned} V_h(\rho) &\equiv \mathbb{E}^Q u(U) + \mathbb{E}\varphi(dQ/dP) \\ &= u\left(U_0 + \mu h - \Sigma\sqrt{h}\right) + \frac{1}{2}\left[\left(1 + \rho\sqrt{h}\right) H_h + \varphi\left(1 + \rho\sqrt{h}\right) + \varphi\left(1 - \rho\sqrt{h}\right)\right], \end{aligned} \tag{42}$$

where

$$H_h = u\left(U_0 + \mu h + \Sigma\sqrt{h}\right) - u\left(U_0 + \mu h - \Sigma\sqrt{h}\right).$$

Let ρ_h be the value of ρ that minimizes $V_h(\rho)$. Setting the derivative of V_h at ρ_h to zero, we find

$$H_h + \varphi'\left(1 + \rho_h\sqrt{h}\right) - \varphi'\left(1 - \rho_h\sqrt{h}\right) = 0. \tag{43}$$

Since $\lim_{h\downarrow 0} H_h = 0$, it follows that

$$\lim_{h\downarrow 0} \rho_h\sqrt{h} = 0. \tag{44}$$

We now feed this limit back into the optimality condition to prove that in fact ρ_h converges. Equation (43) can be rearranged to read

$$\rho_h = -\Sigma\frac{\left(u\left(U_0 + \mu h + \Sigma\sqrt{h}\right) - u\left(U_0 + \mu h - \Sigma\sqrt{h}\right)\right)/2\Sigma\sqrt{h}}{\left(\varphi'\left(1 + \rho_h\sqrt{h}\right) - \varphi'\left(1 - \rho_h\sqrt{h}\right)\right)/2\rho_h\sqrt{h}}.$$

As $h \downarrow 0$, the numerator converges to $u'(U_0)$, which we normalized to one, and the denominator converges to $\varphi''(1) = \theta$. Therefore, as $h \downarrow 0$, ρ_h converges to $\rho_0 \equiv -\Sigma/\theta$, which is the value of ρ that minimizes the quadratic

$$\mathcal{G}(\rho) \equiv \mu + \rho\Sigma - \frac{1}{2}a^u(U_0)\Sigma^2 + \frac{\theta}{2}\rho^2. \tag{45}$$

Summarizing,

$$V_h(\rho_h) = \min_\rho V_h(\rho), \quad \mathcal{G}(\rho_0) = \min_\rho \mathcal{G}(\rho), \quad \lim_{h\downarrow 0} \rho_h = \rho_0. \tag{46}$$

In the next stage of the proof, we approximate V_h. A suitable approximation for the first term, $E^Q u(U)$, of $V_h(\rho)$ is given by Lemma 6. We now derive an analogous approximation for the second term, $\mathbb{E}\varphi(1 + \rho B)$. (This is essentially a special case of the argument of Lemma 6, but it is simpler to argue from first principles.) Given any $\varepsilon \in (0,1)$, further reduce the value of $\bar{h}$ if necessary so that $\rho\sqrt{h}, \rho_h\sqrt{h} \in (-\varepsilon, \varepsilon)$ for all $h \in \left(0, \bar{h}\right)$ and $\rho \in (\rho_0 - 1, \rho_0 + 1)$. In the remainder of this proof, we assume that every instance of the pair (ρ, h) lies in $(\rho_0 - 1, \rho_0 + 1) \times \left(0, \bar{h}\right)$, and therefore ρB is valued in $(-\varepsilon, \varepsilon)$. We approximate the term $\mathbb{E}\varphi(1 + \rho B)$ in the definition of $V_h(\rho)$ by applying the second-order Taylor approximation (23) with $f(x) = \varphi(1 + x)$ and $n = 2$. Since B takes the values $\pm\sqrt{h}$, $\varphi(1) = \varphi'(1) = 0$ and $\varphi''(1) = \theta$, the error bound (24) applied to this context gives

$$\left|\mathbb{E}\varphi(1 + \rho B) - \frac{\theta}{2}\rho^2 h\right| \leq \rho^2 hc\left(\rho\sqrt{h}\right),$$

for a function $c : [-\varepsilon, \varepsilon] \to \mathbb{R}_+$ that is continuous and vanishes at zero. Combining this with approximation (30) and bound (31) of Lemma 6, we obtain

$$V_h(\rho) = \mathcal{G}(\rho)\, h + \chi(\rho, h)\, h, \tag{47}$$

where the continuous function χ satisfies the bound

$$|\chi(\rho, h)| \leq |K_0 + K_1\rho|\, h + b(h) + \rho^2 c\left(\rho\sqrt{h}\right),$$

where b and c are continuous and vanish at zero, and the parameters K_0, K_1, b and c do not vary with ρ. Since ρ_h converges, we have the key facts

$$\lim_{h\downarrow 0} \chi(\rho_0, h) = 0 \quad \text{and} \quad \lim_{h\downarrow 0} \chi(\rho_h, h) = 0. \tag{48}$$

On the other hand, (46) and (47) imply the string of inequalities

$$\mathcal{G}(\rho_0)\, h + \chi(\rho_0, h)\, h = V_h(\rho_0) \geq V_h(\rho_h) = \mathcal{G}(\rho_h)\, h + \chi(\rho_h, h)\, h \geq \mathcal{G}(\rho_0)\, h + \chi(\rho_h, h)\, h.$$

The last two displays prove that

$$u(\upsilon(U)) = V_h(\rho_h) = \mathcal{G}(\rho_0)\, h + o(h).$$

Given the normalization (28), a first-order Taylor expansion of u^{-1} around zero yields the conclusion

$$\upsilon(U) = U_0 + \mathcal{G}(\rho_0)\, h + o(h), \quad \rho_0 = -\frac{\Sigma}{\theta}. \tag{49}$$

Computing the value $\mathcal{G}(\rho_0)$ and using the identity $a^{\psi\circ u} = a^u + \left(a^\psi \circ u\right)(u')$ with the normalization (28), we also have

$$\mathcal{G}(\rho_0) = \mu - \frac{1}{2}\left(a^u(U_0) + \frac{1}{\theta}\right)\Sigma^2 = \mu - \frac{1}{2}a^{\psi\circ u}(U_0)\,\Sigma^2.$$

Therefore, by part Theorem 5(a),

$$(\psi \circ u)^{-1}\, \mathbb{E}(\psi \circ u)(U) = U_0 + \mathcal{G}(\rho_0)\, h + o(h). \tag{50}$$

The combination of equations (49) and (50) completes the proof.

A.1.5 Divergence CE: Poissonian Approximation

The purpose of this subsection is to derive the exact form of CE approximation (22) for a smooth divergence CE and Poisson uncertainty. The approximation is used in the following subsection to prove Theorem 4(b).

We proceed under the assumption that (P, B) is the Poissonian model (4) and υ is the divergence CE (12) for some $u \in C^1_{\text{vNM}}$ and $\varphi \in C^1_{\text{div}}$. For simplicity, we focus on the case in which the infimum defining the divergence CE is achieved as a minimum within Π for all sufficiently small h. Lemma 8 (or a direct calculation) implies that the latter condition is equivalent to the membership of (U_0, Σ) to the set D defined in (13). If either $\Sigma \leq 0$ or $\varphi'(0+) = -\infty$ (as in the entropic case), then $(U_0, \Sigma) \in D$ for all $U_0 > \ell - \Sigma$. If $\varphi'(0+) > -\infty$ and Σ is positive and sufficiently high, a minimizing probability does assign zero probability to the positive jump, which is the case being excluded here.

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

$$A^u_\zeta(U_0, \Sigma) = \Sigma - \frac{\zeta(u(U_0 + \Sigma) - u(U_0))}{u'(U_0)}, \tag{51}$$

which extends the notation in (26), since $A^u = A^u_{\text{identity}}$. In fact, the last identity can be viewed as a limiting case (for $\theta = \infty$) of the more interesting identity

$$A^{\psi \circ u} = A^u_\psi, \quad \text{where} \quad \psi(u) = \theta\left(1 - \exp\left(-\frac{u}{\theta}\right)\right). \tag{52}$$

The following result fleshes out the CE approximation (22) in the current context. Note that in the entropic case, the function ζ defined in (53) is given as $\zeta = \psi$, which, in light of Example 3 and identity (52), makes approximation (53) consistent with the EU approximation of Theorem 5(b).

Theorem 11 (Divergence CE Poissonian Approximation) *Suppose (P, B) is given by the Poisson specification (4), and υ is the divergence CE (12) for some $u \in C^1_{vNM}$ and $\varphi \in C^1_{div}$. Then for all $(U_0, \Sigma) \in D$,*

$$\upsilon(U) = U_0 + \left(\mu - A^u_\zeta(U_0, \Sigma)\right)h + o(h), \quad \textit{where} \quad \zeta(x) = \min_{y \in (0,\infty)} \{\varphi(y) + xy\}, \tag{53}$$

and A^u_ζ is defined in (51).

Proof. As in the proof of Theorem 4(a), we assume the vNM index normalization (28). Again this entails no loss of generality: If u and $\tilde{u}$ are vNM indices such that $\tilde{u} = \alpha u + \beta$, where $\alpha \in (0, \infty)$ and $\beta \in \mathbb{R}$, and φ and $\tilde{\varphi}$ are divergence indices such that $\tilde{\varphi} = \alpha\varphi$, then (u, φ) defines the same divergence CE as $(\tilde{u}, \tilde{\varphi})$, and $A^u_\zeta = A^{\tilde{u}}_{\tilde{\zeta}}$, where $\tilde{\zeta}(x) = \min_{y>0}\{\tilde{\varphi}(y) + xy\}$. It is therefore clear that to prove the theorem, it is sufficient to do so after replacing $u(\cdot)$ with $(u(\cdot) - u(U_0))/u'(U_0)$ and $\varphi(\cdot)$ with $\varphi(\cdot)/u'(U_0)$.

In addition to normalization (28), we adopt the notation (29) and we assume $h \in (0, \bar{h})$, where $\bar{h}$ is defined just as in the discussion immediately following equation (29).

Given the Poisson specification (4), Lemma 8 implies the existence of a minimizing probability in Π, for all $(U_0, \Sigma) \in D$. Given any prior Q, let $\rho = \mathbb{E}^Q B/h$ and recall that $dQ/dP = 1 + \rho B$. Using the notation in (36) and (37), we define

$$\begin{aligned} V_h(\rho) &\equiv \mathbb{E}^Q u(U) + \mathbb{E}\varphi(dQ/dP) \\ &= u(U_0 + \alpha(h)) + ((1+\rho)h + \epsilon(h)) H_h \\ &\quad + (h + \epsilon(h))\varphi(1 + \rho(1 - h - \epsilon(h))) + (1 - h - \epsilon(h))\varphi(1 + \rho(-h - \epsilon(h))), \end{aligned}$$

where $H_h = u(U_0 + \Sigma + \alpha(h)) - u(U_0 + \alpha(h))$. Let ρ_h be the value of ρ that minimizes $V_h(\rho)$.

Setting the derivative of V_h at ρ_h to zero, we find

$$H_h + \left(1 + \frac{\epsilon(h)}{h}\right)(1 - h - \epsilon(h)) G_h = 0, \tag{54}$$

where

$$G_h = \varphi'(1 + \rho_h(1 - h - \epsilon(h))) - \varphi'(1 - \rho_h(h + \epsilon(h))). \tag{55}$$

As h goes to zero, H_h converges to $u(U_0 + \Sigma)$ and $\epsilon(h)/h$ converges to zero. From equation (54), it follows that

$$\lim_{h \downarrow 0} G_h = -u(U_0 + \Sigma). \tag{56}$$

We show that ρ_h must then also converge. Since $1+\rho B$ is strictly positive, $1+\rho_h(1 - h - \epsilon(h)) > 0$, which implies that $\rho_h > -2$ for all $h \in (0, \bar{h})$, provided $\bar{h}$ is sufficiently small. Since φ' is increasing, the quantity $\varphi'(1 - \rho_h(h + \epsilon(h)))$ in (55) is bounded above by a constant for all $h \in (0, \bar{h})$. Because of (56), it must then be the case that the first term in (55) is also bounded above by a constant for sufficiently small h, and therefore ρ_h must also be bounded above for all sufficiently small h. Since ρ_h is bounded both below and above as h goes to zero,

$$\lim_{h \downarrow 0} \rho_h(h + \epsilon(h)) = 0. \tag{57}$$

Finally, equation (55) can be rearranged to

$$\rho_h = -1 + \rho_h(h + \epsilon(h)) + \varphi'^{-1}(G_h + \varphi'(1 - \rho_h(h + \epsilon(h)))),$$

which combined with (56) and (57) implies that

$$\lim_{h \downarrow 0} \rho_h = \rho_0 \equiv -1 + \varphi'^{-1}(-u(U_0 + \Sigma)).$$

Note that ρ_0 is the value of ρ that minimizes

$$\mathcal{G}(\rho) \equiv \mu + \rho\Sigma - A^u(U_0, \Sigma)(1 + \rho) + \varphi(1 + \rho).$$

Summarizing, condition (46) of the proof of Theorem 4(a) holds in the current context, too.

In the next stage of the proof, we approximate V_h. A suitable approximation for the first term, $E^Q u(U)$, of $V_h(\rho)$ is given by Lemma 7. The analogous computation for $\mathbb{E}\varphi(1+\rho B)$ yields an approximation of the form

$$|\mathbb{E}\varphi(1+\rho B) - \varphi(1+\rho)h| \leq \delta(\rho, h)h,$$

where $\delta(\rho_0, h)$ and $\delta(\rho_h, h)$ converge to zero as h goes to zero. Adding up the two approximations results in expression (47) for a function χ such that the limits (48) hold. In the proof of Theorem 4(a), these conditions together with (46) were used to prove equation (49). The exact same argument in the current context yields

$$\upsilon(U) = U_0 + \mathcal{G}(\rho_0)h + o(h), \quad 1 + \rho_0 = \varphi'^{-1}(-u(U_0 + \Sigma)).$$

Equation (53) follows from the identity

$$\mathcal{G}(\rho_0) = \mu - A^u_\zeta(U_0, \Sigma), \quad \text{where} \quad \zeta(x) = \min_{y \in (0,\infty)} \{\varphi(y) + xy\},$$

which can be shown easily using the definitions and the observation that for $x = u(U_0 + \Sigma)$, the minimum value $\zeta(x)$ is achieved by $y = \varphi'^{-1}(-x) = 1 + \rho_0$. ■

A.1.6 Proof of Theorem 4(b)

By Example 3 (which was proved in Remark 9), if φ is entropic, then the CE υ can be expressed (exactly) as an EU CE. This proves the theorem's "if" part. To show the converse, suppose that there exist a prior Q and some $\tilde{u} \in C^1_{\text{vNM}}$ such that $\upsilon(U) = \tilde{u}^{-1}\mathbb{E}^Q\tilde{u}(U) + o(h)$ for all $(U_0, \Sigma) \in D$, and let $\tilde{\rho}$ be the drift of B under Q. Combining Theorems 5(b) and 11, it follows that

$$-A^u_\zeta(U_0, \Sigma) = \tilde{\rho}\Sigma - A^{\tilde{u}}(U_0, \Sigma)(1 + \tilde{\rho}), \quad (U_0, \Sigma) \in D.$$

The following lemma proves, in particular, that this assumption implies that $\tilde{\rho} = 0$ (and therefore $Q = P$) and φ is entropic (and therefore $\zeta = \psi$). This completes the theorem's proof.

We conclude with the key lemma, which is stated here slightly more generally than needed above, since we are going to use it again in the continuous-time analysis to follow.

Lemma 12 *Given any $\varphi \in C^1_{div}$, define the set D by (13), the notation A^u_ζ by (51), and let*

$$\theta = \varphi''(1), \quad \psi(u) = \theta\left(1 - \exp\left(-\frac{u}{\theta}\right)\right) \quad \textit{and} \quad \zeta(x) = \min_{y \in (0,\infty)} \{\varphi(y) + xy\}. \tag{58}$$

Then the following two statements are equivalent, for any $\tilde{u} \in C^1_{vNM}$, $\Sigma \in \mathbb{R}$ and $\rho, \tilde{\rho} \in (-1, \infty)$.

1. $\rho\Sigma - A^u_\zeta(U_0, \Sigma)(1 + \rho) = \tilde{\rho}\Sigma - A^{\tilde{u}}(U_0, \Sigma)(1 + \tilde{\rho})$ *for all* $(U_0, \Sigma) \in D$.

2. $\rho = \tilde{\rho}$ *and* $\varphi(y) = \theta(y \log y - y + 1)$ *for all* $y > 0$.

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

$$\tilde{u}'(1) = u'(1) \quad \text{and} \quad \tilde{u}(1) = u(1) = 0,$$

which entails no loss of generality since $A^{\tilde{u}}$ is invariant to a positive affine transformation of $\tilde{u}$, and A^u_ζ is invariant to adding a constant to u.

Suppose the Lemma's condition 1 is true and define

$$f(x) = \frac{1+\rho}{1+\tilde{\rho}}\,\zeta(x).$$

The assumed condition is equivalent to

$$\frac{f(u(U_0+\Sigma) - u(U_0))}{u'(U_0)} = \frac{\tilde{u}(U_0+\Sigma) - \tilde{u}(U_0)}{\tilde{u}'(U_0)}, \tag{59}$$

for all $(U_0, \Sigma) \in D$. Letting $U_0 = 1$ and $z = 1 + \Sigma$ it follows that if $u(z) < -\varphi'(0+)$,

$$f(u(z)) = \tilde{u}(z) \quad \text{and therefore } f'(u(z))\,u'(z) = \tilde{u}'(z).$$

Assuming

$$x = u(U_0) < -\varphi'(0+) \quad \text{and} \quad y = u(U_0+\Sigma) - u(U_0) < -\varphi'(0+),$$

condition (59) becomes

$$f'(x)\,f(y) = f(x+y) - f(x).$$

Differentiating with respect to y and taking logarithms results in

$$\log f'(x) + \log f'(y) = \log f'(x+y).$$

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

$$\log f'(x) = ax. \tag{60}$$

Since $\zeta(x) = \min_{y>0}\{\varphi(y) + xy\}$, if $x = -\varphi'(y)$, then $\zeta(x) = \varphi(y) + xy$ and $\zeta'(x) = y$ (by the envelope theorem). Using this fact in identity (60) with $x = -\varphi'(y)$, we obtain

$$\log\left(\frac{1+\rho}{1+\tilde{\rho}}\right) + \log y = -a\varphi'(y), \quad y > 0.$$

Since $\varphi'(1) = 0$, it follows that $\tilde{\rho} = \rho$. Since $\varphi''(1) = \theta$, it follows that $a\theta = -1$. Therefore, φ solves the ODE

$$\varphi(1) = 0, \quad \varphi'(y) = \theta\log(y), \quad y > 0,$$

whose unique solution is $\varphi(y) = \theta(y\log y - y + 1)$. ■

A.2 Continuous-Time Recursive Utility

This section shows and elaborates on the claims of Section 5. As outlined there, the main objective is to establish the appropriate modifications to Duffie and Epstein (1992) utility, if any, for the smooth ambiguity averse CEs of interest, under Brownian and Poisson uncertainty. This section's analysis follows more closely the approach in Skiadas (2008), which is extended here to include Poisson jumps. The setting also extends that of the main paper in that we allow for multiple Brownian/Poissonian sources of risk and a more general intertemporal aggregator. The arguments can be followed with only an intuitive grasp of elements of the stochastic calculus, which are outlined in the first subsection. The advanced theory can be found in Jacod and Shiryaev (2003).

A.2.1 Stochastic Setting

Given is a probability space $(\Omega, \mathcal{F}, P)$ on which are defined d mutually stochastically independent processes forming the column vector $B = (B^1, \ldots, B^k, B^{k+1}, \ldots, B^d)'$, 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 every time t and process $i \in \{k+1, \ldots, d\}$.

The underlying filtration $\{\mathcal{F}_t : t \in [0, T]\}$, where $T > 0$ is a given finite time horizon, is defined as the smallest filtration such that B_t is $\mathcal{F}_t$-measurable for every time t (and each $\mathcal{F}_t$ contains the P-null events, a technicality that can be ignored). We assume that $\mathcal{F} = \mathcal{F}_T$, without loss of generality. Intuitively, uncertainty is represented by the possible paths of B. Time-t information results from observing the realized path of B up to time t. Given this information, conditional uncertainty resolved by time $t+dt$ is spanned by the stochastically and linearly independent factors $dB_t^1, \ldots, dB_t^d$, representing infinitesimal risks, some of which are Brownian and some Poissonian. For $d = 1$, one can think of last section's CE approximations, which become exact relationships in the continuous-time limit, as applying over each interval $[t, t+dt]$, conditionally on time-t information, with dt in place of h and dB in place of B.

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. 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.

Conditional expectation given time-t information $\mathcal{F}_t$ (resp. $\mathcal{F}_{t-}$) is denoted $\mathbb{E}_t$ (resp. $\mathbb{E}_{t-}$). A *volatility process* is any d-dimensional predictable process σ such that $\int_0^T \sigma_t' \sigma_t dt < \infty$ with probability one. A local martingale M can be heuristically thought of as an adapted process whose instantaneous increments have conditionally zero mean: $\mathbb{E}_{t-} dM_t = 0$. Given an integrability condition, such an increment dM_t can be expressed as a linear combination of the instantaneous linear factors $dB_t^1, \ldots, dB_t^d$. This intuition is formalized by a martingale representation theorem, which in the current context states that a process M is a locally square-integrable martingale if and only if there exists some volatility process σ such that

$$M_t = M_0 + \int_0^t \sigma_u' dB_u, \quad \text{or equivalently} \quad dM_t = \sigma_t' dB_t. \tag{61}$$

A *prior* is any probability on $\mathcal{F}$ that is equivalent to P (meaning that it defines the same null events as P). Associated with a prior Q are a martingale ξ^Q, a d-dimensional predictable process ρ^Q and a d-dimensional adapted process B^Q, defined by

$$\xi_t^Q = \mathbb{E}_t \left[\frac{dQ}{dP} \right], \quad \frac{d\xi_t^Q}{\xi_{t-}^Q} = \rho_t^{Q\prime} dB_t \quad \text{and} \quad B_t^Q = B_t - \int_0^t \rho_t^Q dt. \tag{62}$$

(Note that the ratio $\xi_{t+dt}^Q / \xi_{t-}^Q = 1 + \rho_t^{Q\prime} dB_t$ is the analog of the ratio $dQ/dP = 1 + \rho B$ in the single period model, where Q and P represent transition probabilities over a single binomial step. Here Q and P represent probabilities over entire paths.) We let Π denote the set of every prior Q such that ρ^Q is a volatility process.[16] Girsanov's theorem implies that for any $Q \in \Pi$,

$$B^Q \text{ is a local martingale under } Q. \tag{63}$$

Expectation relative to a prior Q is denoted $\mathbb{E}^Q$. Heuristically, we have $\mathbb{E}_{t-}^Q dB_t^Q = 0$.

Priors in this setting correspond to beliefs about the drift process of each Brownian motion and the arrival rate process of each Poisson process, all of which can be path dependent. For the Brownian factors, Lévy's characterization of Brownian motion implies that $B^{Q1}, \ldots, B^{Qk}$ are independent standard Brownian motions under the probability Q, and therefore ρ^{Qi} is the drift of Brownian motion B^{Qi} under Q. For the Poissonian factors, we note that

$$B_t^{Qi} = N_t^i - \int_0^t \left(1 + \rho_s^{Qi}\right) ds, \quad i = k+1, \ldots, d, \tag{64}$$

which in combination with (63) implies that $1 + \rho^{Qi}$ is the arrival rate process of the point process N^i under the probability Q.

We will derive utility dynamics through a formal application of Ito's rule, which we now review. Consider any process of the form

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

where σ is a volatility process and μ^Q is a *drift* process, meaning that it is predictable and $\int_0^t \left|\mu_u^Q\right| du < \infty$ with probability one. For any twice continuously differentiable function $f : \mathbb{R} \to \mathbb{R}$,

[16] 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.

Ito's rule states that

$$df\left(X_t\right) = \alpha_t^Q dt + \beta_t' dB_t^Q. \tag{65}$$

where

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

Since B^Q is a local martingale under Q, we have the heuristic expression

$$\mathbb{E}_{t-}^Q\left[f\left(X_{t+dt}\right)\right] = f\left(X_{t-}\right) + \alpha_t^Q dt. \tag{66}$$

A.2.2 Recursive Utility as a BSDE

As in the single-period analysis, we assume consumption is valued in the interval (ℓ,∞), for some $\ell \in [-\infty, 1)$. Ignoring technical integrability conditions, a *consumption plan* is any (ℓ,∞)-valued adapted (progressively measurable) process c, where c_t represents a consumption rate for $t < T$, and c_T represents terminal consumption (which can be thought of as a constant perpetuity paying out at the rate c_T at all times after T). The consumption plan that is identically equal to one (including unit terminal consumption) is denoted $\mathbf{1}$. For each consumption plan c, we will define a corresponding utility process $U(c)$, which is normalized so that $U(s\mathbf{1}) = s\mathbf{1}$ for any $s \in (\ell,\infty)$. The interpretation is that at time t, conditionally on time-t information, the agent is indifferent between the plans c and $U_t(c)\,\mathbf{1}$.

We henceforth fix a reference consumption plan c and we simplify the notation for the corresponding utility process by writing U instead of $U(c)$. Heuristically, we assume that the process U is computed by a backward-in-time recursion of the form[17]

$$U_{t-} = \Phi\left(dt, c_t, \upsilon_t\left(U_{t+dt}\right)\right), \quad U_T = c_T, \tag{67}$$

where $\upsilon_t\left(U_{t+dt}\right)$ is a time-t conditional CE of U_{t+dt}, to be specified later on, and Φ is an intertemporal aggregator. We assume throughout that Φ has continuous partial derivatives, denoted Φ_{dt}, Φ_c and Φ_υ, with Φ_c and Φ_υ taking values in $(0,\infty)$, reflecting preference monotonicity. Moreover, we assume that Φ satisfies the consistency condition $\Phi(0,c,\upsilon) = \upsilon$. Most common in applications is the specification

$$\Phi\left(dt, c, \upsilon\right) = u_\delta^{-1}\left(\left(1-e^{-\beta dt}\right)u_\delta(c) + e^{-\beta dt}u_\delta(\upsilon)\right), \tag{68}$$

for some vNM index u_δ and impatience rate β. This includes Epstein-Zin-Weil utility and the example of Section 2, where u_δ is parameterized to imply a constant EIS.

[17] Simple behavioral axioms that lead to recursive utility with an arbitrary conditional CE can be found in Chapter 6 of Skiadas (2009), whose end-notes give further background on the foundations of recursive utility.

To make mathematical sense of the heuristic recursion (67), suppose the processes μ and Σ represent, respectively, the drift and volatility of U :

$$dU_t = \mu_t dt + \Sigma_t' dB_t = \left(\mu_t + \rho_t^Q \cdot \Sigma_t\right) dt + \Sigma_t' dB_t^Q, \tag{69}$$

for any prior Q, with B^Q defined in (62). Note that $dU_t = U_{t+dt} - U_{t-}$ includes any time-t jump. Since $\mathbb{E}_{t-}^Q dB_t^Q = 0$, the second equation in (69) implies that the risk-neutral conditional CE of U_{t+dt} under the prior Q is

$$\mathbb{E}_{t-}^Q U_{t+dt} = U_{t-} + \left(\mu_t + \rho_t^Q \cdot \Sigma_t\right) dt. \tag{70}$$

Reflecting the single-period CE approximation form (22), the continuous-time conditional CEs of interest in this paper can be expressed as

$$\upsilon_t\left(U_{t+dt}\right) = U_{t-} + \left(\mu_t + \rho_t^Q \cdot \Sigma_t - \mathcal{A}\left(U_{t-}, \Sigma_t, \rho_t^Q\right)\right) dt. \tag{71}$$

This claim is verified in the following subsection, where explicit expressions for $\mathcal{A}$ are given.

Let us now expand the right-hand side of the heuristic recursion (67) in a first-order Taylor approximation with respect to the arguments dt and $\upsilon_t\left(U_{t+dt}\right)$, using expression (71), to obtain

$$U_{t-} = \Phi\left(0, c_t, U_{t-}\right) + \Phi_{dt}\left(0, c_t, U_{t-}\right) dt + \Phi_{\upsilon}\left(0, c_t, U_{t-}\right)\left(\mu_t + \rho_t^Q \cdot \Sigma_t - \mathcal{A}\left(U_{t-}, \Sigma_t, \rho_t^Q\right)\right) dt.$$

Given the normalization $U_{t-} = \Phi\left(0, c_t, U_{t-}\right)$, the preceding equation can be rearranged to

$$-\mu_t = f\left(c_t, U_{t-}\right) + \rho_t^Q \cdot \Sigma_t - \mathcal{A}(U_{t-}, \Sigma_t, \rho_t^Q), \quad \text{where} \quad f\left(c, \upsilon\right) = \frac{\Phi_{dt}\left(0, c, \upsilon\right)}{\Phi_{\upsilon}\left(0, c, \upsilon\right)}. \tag{72}$$

For example, if Φ is given by (68), then $f\left(c, \upsilon\right) = \beta\left(u_\delta\left(c\right) - u_\delta\left(\upsilon\right)\right)/u_\delta'\left(\upsilon\right)$.

Combining (69) and (72), we have transformed the heuristic recursive specification (67) of the utility process U to the backward stochastic differential equation (BSDE):

$$dU_t = -\left(f\left(c_t, U_t\right) + \rho_t^Q \cdot \Sigma_t - \mathcal{A}(U_{t-}, \Sigma_t, \rho_t^Q)\right) dt + \Sigma_t' dB_t, \quad U_T = c_T. \tag{73}$$

Given the terminal value U_T, a solution to the BSDE consists of an adapted pair (U, Σ) such that (73) holds. We also refer to the process U as a solution to the BSDE if (U, Σ) is a BSDE solution for some Σ. The fixed-point nature of BSDE (73) means that special restrictions must be imposed on the primitives (including integrability restrictions on c and ρ^Q) to guarantee the existence and uniqueness of a solution (see footnote 12 and Skiadas (2008) for relevant references).

A.2.3 Smooth Ambiguity Aversion in Continuous Time

A general form of continuous-time recursive utility has been expressed as BSDE (73), with the function $\mathcal{A}$ defined heuristically in the CE expression (71). This section specifies the function $\mathcal{A}$ for smooth conditional CEs corresponding to second-order expected utility and divergence preferences, and characterizes the instances in which the resulting specification is equivalent to one with an EU CE. The claimed expressions, which are analogous to last section's single-period CE approximations, are based on formal applications of Ito's lemma, although the irrelevance of the function φ for a smooth second-order EU CE will be seen to be even simpler than that.

Expected Utility CE We first establish the continuous-time version of the recursive utility of Kreps and Porteus (1978) in what amounts to a variant of the argument of Duffie and Epstein (1992), extended to include Poisson jumps. We assume that

$$\upsilon_t\left(U_{t+dt}\right) = u^{-1}\mathbb{E}^Q_{t-}u\left(U_{t+dt}\right), \tag{74}$$

for some $Q \in \Pi$ and $u \in C^2_{\mathrm{vNM}}$ (or just C^1_{vNM} if there is no Brownian risk). Recall that the risk-aversion functions a^u and A^u are defined in (25) and (26), respectively. Theorem 5 suggests that the CE expression (71), and therefore BSDE (73), is satisfied with

$$\mathcal{A}\left(U_{t-},\Sigma_t,\rho_t\right) = \frac{a^u\left(U_{t-}\right)}{2}\sum_{i=1}^{k}\left(\Sigma^i_t\right)^2 + \sum_{i=k+1}^{d}A^u\left(U_{t-},\Sigma^i_t\right)\left(1+\rho^i_t\right). \tag{75}$$

This claim follows by a formal application of Ito's lemma similarly to (66):

$$\mathbb{E}^Q_{t-}u\left(U_{t+dt}\right) = u\left(U_{t-}\right) + u'\left(U_{t-}\right)\left(\mu_t + \rho^Q_t\cdot\Sigma_t - \mathcal{A}(U_{t-},\Sigma_t,\rho^Q_t)\right)dt, \tag{76}$$

where $\mathcal{A}$ is defined by (75). Further applying u^{-1} on both sides of equation (76) and using Ito's lemma (in a trivial sense) confirms (71) the current context.

Second-Order Expected-Utility CE In Theorem 2, we saw that a smooth second-order EU CE over small Brownian or Poissonian risks is approximately equal to an EU CE with the compound prior. This relationship should be exact in the continuous time limit, meaning that the utility BSDE for a smooth second-order EU CE should be specified as in the last subsection, with Q being the compound prior. There are two parts to this claim: The function φ does not appear in the utility BSDE, and the prior Q is the compound prior. Let us now clarify these two parts.

The irrelevance of the function φ can be made in greater generality than the second-order EU formulation, as follows. Suppose that for some $\varphi \in C^1_{\mathrm{vNM}}$,

$$\upsilon_t\left(U_{t+dt}\right) = \varphi^{-1}\left(\sum\nolimits_{s=1}^{S}\varphi\left(\upsilon^s_t\left(U_{t+dt}\right)\right)\pi^s_t\right), \tag{77}$$

where $\pi^1,\ldots,\pi^S$ are $(0,1)$-valued predictable processes such that $\sum_s \pi^s = \mathbf{1}$, and $\upsilon^1,\ldots,\upsilon^S$ are conditional CEs that can be expressed as

$$\upsilon^s_t\left(U_{t+dt}\right) = U_{t-} + \left(\mu_t + \mathcal{D}^s_t\left(U_{t-},\Sigma_t\right)\right)dt, \quad s = 1,\ldots,S. \tag{78}$$

(As always, μ and Σ refer to the dynamics (69) of U.) A first-order Taylor expansion gives

$$\varphi\left(\upsilon^s_t\left(U_{t+dt}\right)\right) = \varphi\left(U_{t-}\right) + \varphi'\left(U_{-}\right)\left(\mu_t + \mathcal{D}^s_t\left(U_{t-},\Sigma_t\right)\right)dt.$$

Multiplying the last equation by π^s_t, adding up over s, applying φ^{-1} on both sides and taking another first-order expansion results in

$$\upsilon_t\left(U_{t+dt}\right) = U_{t-} + \left(\mu_t + \mathcal{D}_t\left(U_{t-},\Sigma_t\right)\right)dt, \quad \text{where} \quad \mathcal{D}_t = \sum\nolimits_{s=1}^{S}\mathcal{D}^s_t\pi^s_t. \tag{79}$$

A corresponding BSDE can be established as in Section A.2.2. *The function* φ *is not part of equation* (79) *or the corresponding BSDE and is therefore irrelevant.*

The second-order EU CE definition specializes the preceding formulation by postulating priors $Q^1, \ldots, Q^S$ and a $u \in C^2_{\text{vNM}}$ 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 weights π_t^s are updated using Bayes' rule:

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

By the EU CE analysis of the last subsection, equations (78) are satisfied with

$$\mathcal{D}_t^s \left(U_{t-}, \Sigma_t\right) = \rho_t^{Q^s} \cdot \Sigma_t - \mathcal{A}(U_{t-}, \Sigma_t, \rho_t^{Q^s}), \quad s = 1, \ldots, S, \tag{80}$$

where the function $\mathcal{A}$ is defined in equation (75). The compound prior is defined by

$$Q = \sum\nolimits_{s=1}^S Q^s \pi_0^s \quad \text{and therefore} \quad \xi^Q = \sum\nolimits_{s=1}^S \xi^s \pi_0^s. \tag{81}$$

In addition to the irrelevance of φ, we can now claim that υ *is an EU CE with prior* Q *and vNM index* u. To verify this claim, we first note that the second equation in (81) together with the Bayes formula defining π_t^s results in

$$\frac{d\xi_t^Q}{\xi_{t-}^Q} = \sum\nolimits_{s=1}^S \frac{d\xi_t^s}{\xi_{t-}^s} \pi_t^s \quad \text{and therefore} \quad \rho_t^Q = \sum\nolimits_{s=1}^S \rho_t^{Q^s} \pi_t^s.$$

Since in equation (80) the dependence of $\mathcal{A}(U_{t-}, \Sigma_t, \rho_t^{Q^s})$ on ρ^{Q^s} is linear (as stated in (75)), we conclude that

$$\mathcal{D}_t \left(U_{t-}, \Sigma_t\right) = \sum\nolimits_{s=1}^S \mathcal{D}_t^s \left(U_{t-}, \Sigma_t\right) \pi_t^s = \rho_t^Q \cdot \Sigma_t - \mathcal{A}(U_{t-}, \Sigma_t, \rho_t^Q).$$

Using this in CE expression (79) and comparing to the corresponding EU expression in Section A.2.3, we conclude that

$$\upsilon_t \left(U_{t+dt}\right) = u^{-1} \mathbb{E}_{t-}^Q u \left(U_{t+dt}\right).$$

Divergence CE Finally, we formulate the continuous-time version of the divergence CE of Section 4.2, extended to include any finite number of Brownian and Poissonian risk sources.

We postulate a reference prior R. For each Brownian risk source $i \in \{1, \ldots, k\}$, the corresponding reference drift term is ρ^{Ri}, and for each Poissonian risk source $i \in \{k+1, \ldots, d\}$, the corresponding reference arrival rate is $1 + \rho^{Ri}$. For any other prior Q, let the positive R-martingale $\xi^{Q/R}$ and the predictable process $\rho^{Q/R}$ be defined by[18]

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

The other primitives needed to define the conditional CE are $u \in C_{\text{vNM}}^2$ and $\varphi \in C_{\text{div}}^2$ (see section 4.2). If there is no Brownian risk, it is sufficient to assume that $u \in C_{\text{vNM}}^1$ and $\varphi \in C_{\text{div}}^1$.

The continuous-time divergence CE in this context is heuristically defined by

$$\upsilon_t \left(U_{t+dt} \right) = \inf_{Q \in \Pi} u^{-1} \left(\mathbb{E}_{t-}^Q u \left(U_{t+dt} \right) + \mathbb{E}_{t-}^R \varphi \left(\frac{\xi_{t+dt}^{Q/R}}{\xi_{t-}^{Q/R}} \right) \right) . \tag{82}$$

The interior-solution condition

$$u \left(U_{t-} + \Sigma_t^i \right) - u \left(U_{t-} \right) < -\varphi' \left(0+ \right) , \quad i = k+1, \ldots, d, \tag{83}$$

is assumed throughout. Of course the condition is automatically satisfied if there are no Poisson terms $(k = d)$.

As in the single-period analysis, we define θ, ψ and ζ by (58) and A_ζ^u by (51). Theorems 4(a) and 11 suggest that the conditional CE expression (71) and corresponding utility BSDE (73) hold with

$$\mathcal{A}(U_{t-}, \Sigma_t, \rho_t) = \frac{a^{\psi \circ u} \left(U_{t-} \right)}{2} \sum_{i=1}^{k} \left(\Sigma_t^i \right)^2 + \sum_{i=k+1}^{d} A_\zeta^u \left(U_{t-}, \Sigma_t^i \right) \left(1 + \rho_t^i \right) . \tag{84}$$

This is confirmed at the end of this section as a formal application of Ito's lemma, along with the following claims:

- If there is no Poisson risk $(k = d)$, then the above divergence utility specification is equivalent to one with an expected-utility CE, and is therefore within the Duffie-Epstein class of continuous-time Kreps-Porteus utility.

- If there is Poisson risk $(k < d)$, then the above divergence utility specification is equivalent to one with an expected-utility CE if and only if the divergence is of the entropic type: $\varphi(y) = \theta \left(y \log y - y + 1 \right)$, in which case $\zeta = \psi$ and $A_\zeta^u = A^{\psi \circ u}$.

The first claim follows by comparing (84) to (75) and the second claim follows from Lemma 12.

[18] The change-of-measure formula for conditional expectations and the integration-by-parts formula for semimartingales can be used to show the relationships:

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

Example 13 (Quadratic Divergence) *Suppose u is the identity function and, as in Example 10, $\varphi(y) = \theta(y-1)^2/2$ and therefore $\zeta(x) = x - x^2/(2\theta)$. Assuming the validity of condition (83), which in this case states that $\Sigma_t^i < \theta$ for $i > k$, expression (84) reduces to*

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

Schroder and Skiadas (2008) show the tractability advantages of this specification in the presence of Poisson jumps.

We conclude with the formal derivation of the preceding claims. Analogously to the derivation of equation (76), Ito's lemma implies that

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

Combining the preceding expression with (76) results in

$$\begin{aligned} u^{-1}\left(\mathbb{E}_{t-}^Q u\left(U_{t+dt}\right) + \mathbb{E}_{t-}^R\varphi\left(\frac{\xi_{t+dt}^{Q/R}}{\xi_{t-}^{Q/R}}\right)\right) &= U_{t-} + \mu_t dt \\ &\quad + \sum_{i=1}^{k}\mathcal{C}_t\left(\rho_t^{Qi}\right)dt + \sum_{i=k+1}^{d}\mathcal{J}_t\left(\rho_t^{Qi}\right)dt, \end{aligned} \tag{85}$$

where

$$\begin{aligned} \mathcal{C}_t\left(\rho_t^i\right) &= \rho_t^i\Sigma_t^i - \frac{a^u\left(U_{t-}\right)}{2}\left(\Sigma_t^i\right)^2 + \frac{\theta}{2u'\left(U_{t-}\right)}\left(\rho_t^i - \rho_t^{Ri}\right)^2, \\ \mathcal{J}_t\left(\rho_t^i\right) &= \rho_t^i\Sigma_t^i - A^u\left(U_{t-}, \Sigma_t^i\right)\left(1+\rho_t^i\right) + \frac{1}{u'\left(U_{t-}\right)}\varphi\left(\frac{1+\rho_t^i}{1+\rho_t^{Ri}}\right)\left(1+\rho_t^{Ri}\right). \end{aligned}$$

The last two terms are minimized separately, noting that $\mathcal{C}_t$ is quadratic and $\mathcal{J}_t$ is strictly convex. The assumed inequality (83) is equivalent to the condition $\mathcal{J}_t'(-1+\varepsilon) < 0$ 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 (85) is minimized by the value ρ^Q, where the Brownian terms are given by

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

and the Poissonian terms are given by

$$1+\rho_t^{Qi} = \left(1+\rho_t^{Ri}\right)\varphi'^{-1}\left(u\left(U_{t-}\right) - u\left(U_{t-}+\Sigma_t^i\right)\right), \quad i = k+1, \ldots, d.$$

Substituting the minimizing value of ρ^Q in (85) results in

$$\upsilon_t\left(U_{t+dt}\right) = U_{t-} + \left(\mu_t + \rho_t^R \cdot \Sigma_t - \frac{a^{\psi\circ u}\left(U_{t-}\right)}{2}\sum_{i=1}^{k}\left(\Sigma_t^i\right)^2 - \sum_{i=k+1}^{d} A_\zeta^u\left(U_{t-},\Sigma_t^i\right)\left(1+\rho_t^{Ri}\right)\right)dt, \quad (86)$$

where ψ and ζ are defined in (58). Expression (86) is the same as (71) with $Q = R$ and the function $\mathcal{A}$ given by equation (84), as claimed.

If there is no Poissonian risk $(k = d)$, then expression (86) reduces to an EU CE with prior R and vNM index $\psi \circ u$, and therefore the corresponding recursive utility is within the class of continuous-time Kreps-Porteus utilities.

Finally, suppose there is Poissonian risk $(0 \leq k < d)$. The question is, can the divergence CE (86) be expressed as an EU CE relative to some prior and smooth vNM index? We argue that there exists some $Q \in \Pi$ and $\tilde{u} \in C^2_{\text{vNM}}$ such that the CE (86) takes the expected utility form

$$\upsilon_t\left(U_{t+dt}\right) = U_{t-} + \left(\mu_t + \rho_t^Q \cdot \Sigma_t - \frac{a^{\tilde{u}}\left(U_{t-}\right)}{2}\sum_{i=1}^{k}\left(\Sigma_t^i\right)^2 - \sum_{i=k+1}^{d} A^{\tilde{u}}\left(U_{t-},\Sigma_t^i\right)\left(1+\rho_t^{Qi}\right)\right)dt \quad (87)$$

for all values of (U_{t-}, Σ_t) satisfying condition (83) if and only if

$$Q = R \quad \text{and} \quad \varphi\left(x\right) = \theta\left(x\log x + x - 1\right) \quad (\text{where } \theta = \varphi''\left(1\right)). \quad (88)$$

The "if" part is immediate, since (88) implies that $\zeta = \psi$ and $A_\zeta^u = A^{\psi\circ u}$. Conversely, suppose that for some prior Q and $\tilde{u} \in C^2_{\text{vNM}}$, equation (87) is true for all values of (U_{t-}, Σ_t) satisfying (83), that is, for all values of (U_{t-}, Σ_t) such that $\left(U_{t-}, \Sigma_t^i\right) \in D$ for every Poissonian factor i, where D is defined in (13). Isolating any such factor $i \in \{k+1, \ldots, d\}$, the equality of the conditional CEs (86) and (87) implies that

$$\rho_t^{Ri}\Sigma_t^i - A_\zeta^u\left(U_{t-},\Sigma_t^i\right)\left(1+\rho_t^{Ri}\right) = \rho_t^{Qi}\Sigma_t^i - A^{\tilde{u}}\left(U_{t-},\Sigma_t^i\right)(1+\rho_t^{Qi}),$$

for all values of $\left(U_{t-}, \Sigma_t^i\right)$ in D. An application of Lemma 12 shows that then (88) must hold.

Additional References

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

K. J. Arrow (1970), *Essays in the Theory of Risk Bearing,* North Holland, London.

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

J. Jacod and A. N. Shiryaev (2003), *Limit Theorems for Stochastic Processes,* 2nd edition, Springer-Verlag, Berlin Heidelberg.

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

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

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