RFS Advance Access published April 17, 2013

Ambiguous Volatility and Asset Pricing in Continuous Time

Larry G. Epstein
Boston University

Shaolin Ji
Shandong University

We formulate a model of utility for a continuous-time framework that captures aversion to ambiguity about both volatility and drift. Corresponding extensions of some basic results in asset pricing theory are presented. First, we derive arbitrage-free pricing rules based on hedging arguments. Because ambiguous volatility implies market incompleteness, hedging arguments determine prices only up to intervals. In order to obtain sharper predictions, we apply the model of utility to a representative agent endowment economy and study equilibrium asset returns. A version of the consumption capital asset pricing model is derived, and the effects of ambiguous volatility are described. (*JEL* D53, D81, G12)

This paper formulates a model of utility for a continuous-time framework that captures the decision-maker's concern with ambiguity or model uncertainty. Its novelty lies in the range of model uncertainty that is accommodated, specifically in the modeling of ambiguity about both drift and volatility, and in corresponding extensions of some basic results in asset pricing theory. First, we derive arbitrage-free pricing rules based on hedging arguments. Ambiguous volatility implies market incompleteness and thus, in general, rules out perfect hedging. Consequently, hedging arguments determine prices only up to intervals. However, sharper predictions can be obtained by assuming preference maximization and equilibrium. Thus we apply the model of utility to a representative agent endowment economy to study

This work was supported by the National Science Foundation (awards SES-0917740, SES-1216339 to L.G.E.), the National Natural Science Foundation of China (award numbers 11171187, 10921101 to S.J.), the Programme of Introducing Talents of Discipline to Universities of China (award number B12023 to S.J.), and the Program for New Century Excellent Talents in University of China. We have benefited also from very helpful comments by Pietro Veronesi and two referees, and from discussions with Shige Peng, Zengjing Chen, Mingshang Hu, Gur Huberman, Julien Hugonnier, Hiro Kaido, Jin Ma, Semyon Malamud, Guihai Zhao, and especially Jianfeng Zhang. Epstein is grateful also for the generous hospitality of EIEF and CIRANO where much of this work was completed during extended visits. The contents of this paper appeared originally under the title "Ambiguous volatility, possibility and utility in continuous time," first posted March 5, 2011. That paper now focuses on a mathematically rigorous treatment of the model of utility and excludes the application to asset pricing. Send correspondence to Larry Epstein, Boston University, Department of Economics, 270 Bay State Road, Boston MA 02215; telephone: (617) 353-4142. E-mail: lepstein@bu.edu.

doi:10.1093/rfs/hht018

equilibrium asset returns in a sequential Radner-style market setup. A version of the consumption capital asset pricing model is derived and the effects of ambiguous volatility are described. A pivotal role for "state prices" is demonstrated in both the hedging and equilibrium analyses, thus extending to the case of comprehensive ambiguity this cornerstone element of asset pricing theory.

The model of utility is a continuous-time version of multiple priors (or maxmin) utility formulated by Gilboa and Schmeidler (1989) for a static setting. Related continuous time models are provided by Chen and Epstein (2002) and Anderson, Hansen, and Sargent (2003).[1] In *all* existing literature on continuous-time utility, ambiguity is modeled so as to retain the property that all priors are equivalent—that is, they agree which events are null. This universal restriction is driven not by an economic rationale but rather by the technical demands of continuous-time modeling, specifically by the need to work within a probability space framework. Notably, in order to describe ambiguity, authors invariably rely on Girsanov's theorem for changing measures. It provides a tractable characterization of alternative hypotheses about the true probability law, but it also limits alternative hypotheses to correspond to measures that are both mutually equivalent and that differ from one another only in what they imply about drift. This paper defines a more general framework within which one can model the utility of an individual who is not completely confident in any single probability law for either drift or volatility. This is done while maintaining a separation between risk aversion and intertemporal substitution, as in Duffie and Epstein (1992b).

At a technical level, the analysis requires a significant departure from existing continuous-time modeling because ambiguous volatility cannot be modeled within a probability space framework, where there exists a probability measure that defines the set of null (or impossible) events. In our companion paper Epstein and Ji (2013), we exploit and extend recent advances in stochastic calculus that do not require a probability space framework. The reader is referred to that paper for a rigorous treatment of the technical details involved in defining a utility function that accommodates aversion to ambiguity about volatility, including for proofs regarding utility, and also for extensive references to the noted mathematics literature. Our treatment below is less formal but is otherwise largely self-contained. Proofs are provided here for all the asset pricing results.

The paper proceeds as follows. The next section elaborates on motivation and then provides an informal outline of our approach to modeling ambiguous volatility. Section 2 describes the new model of utility. Our asset pricing results follow in Section 3. Concluding remarks are offered in Section 4. Proofs are collected in appendices.

[1] The discrete-time counterpart of the former is axiomatized in Epstein and Schneider (2003).

1. Motivation and Informal Outline

1.1 Why ambiguous volatility?

A large literature has argued that stochastic time-varying volatility is important for understanding empirical features of asset markets; for recent examples, see Eraker and Shaliastovich (2008), Drechsler (forthcoming), Bollerslev, Sizova, and Tauchen (2012), Bansal, Kiku, and Yaron (2012), Beeler and Campbell (2012), Bansal et al. (2011), and Campbell et al. (2012), where the first three employ continuous-time models.[2] In macroeconomic contexts, Bloom (2009) and Fernandez-Villaverde et al. (2010) are recent studies that find evidence of stochastic time varying volatility and its effects on real variables. In all of these papers, evidence suggests that relevant volatilities follow complicated dynamics. The common modeling response is to postulate correspondingly complicated parametric laws of motion, including specification of the dynamics of the volatility of volatility. However, one might question whether agents in these models can learn these laws of motion precisely, and more generally, whether it is plausible to assume that agents become completely confident in any particular law of motion. In their review of the literature on volatility derivatives, Carr and Lee (2009) raise this criticism of assuming a particular parametric process for the volatility of the underlying asset. The drawback they note is "the dependence of model value on the particular process used to model the short-term volatility." They write that "the problem is particularly acute for volatility models because the quantity being modeled is not directly observable. Although an estimate for the initially unobserved state variable can be inferred from market prices of derivative securities, noise in the data generates noise in the estimate, *raising doubts that a modeler can correctly select any parametric stochastic process from the menu of consistent alternatives*."

Thus, we are led to develop a model of preference that accommodates ambiguity about volatility. In the model, the individual takes a stand only on bounds rather than on any particular parametric model of volatility dynamics. Thus, maximization of preference leads to decisions that are robust to misspecifications of the dynamics of volatility (as well as drift). Accordingly, we think of this aspect of our model as providing a way to *robustify stochastic volatility modeling*.

To illustrate the latter perspective, consider a stochastic environment with a one-dimensional driving process. By a stochastic volatility model we mean the hypothesis that the driving process has zero drift and that its volatility is stochastic and is described by a single process (σ_t). The specification of a single process for volatility indicates the investor's complete confidence in the implied dynamics. Suppose, however, that (σ_t^1) and (σ_t^2) describe two alternative stochastic volatility models that are put forth by expert econometricians;

[2] Bollerslev, Sizova, and Tauchen (2012, 32) note that a continuous-time model permits "a direct assessment of its ability to match the qualitative features of the data across different sampling frequencies, including intraday cross-correlation patterns as well as longer run dynamic dependencies."

for instance, they might conform to the Hull and White (1987) and Heston (1993) parametric forms, respectively. The models have comparable empirical credentials and are not easily distinguished empirically, but their implications for optimal choice (or for the pricing of derivative securities, which is the context for the earlier quote from Carr and Lee) differ significantly. Faced with these two models, the investor might place probability $\frac{1}{2}$ on each being the true model. But why should she be certain that either one is true? Both $\left(\sigma_t^1\right)$ and $\left(\sigma_t^2\right)$ may fit data well to some approximation, but other approximating models may do as well. An intermediate model such as $\left(\frac{1}{2}\sigma_t^1+\frac{1}{2}\sigma_t^2\right)$ is one alternative, but there are many others that "lie between" $\left(\sigma_t^1\right)$ and $\left(\sigma_t^2\right)$ and that plausibly should be taken into account. Accordingly, we are led to hypothesize that the investor views as possible all volatility processes with values lying in the interval $[\underline{\sigma}_t(\omega), \overline{\sigma}_t(\omega)]$ for every t and ω, where

$$\underline{\sigma}_t(\omega)=\min\left\{\sigma_t^1(\omega),\sigma_t^2(\omega)\right\} \quad \text{and} \quad \overline{\sigma}_t(\omega)=\max\left\{\sigma_t^1(\omega),\sigma_t^2(\omega)\right\}. \tag{1}$$

Given also the conservative nature of multiple priors utility, the individual will be led thereby to take decisions that are robust to (many) misspecifications of the dynamics of volatility. This special case of our model is described further in Section 2.2.

A possible objection to modeling ambiguity about volatility might take the form: "One can approximate the realized quadratic variation of a stock price (for example) arbitrarily well from frequent observations over any short time interval, and thus estimate the law of motion for its volatility extremely well. Consequently, ambiguity about volatility is implausible for a sophisticated agent." However, even if one accepts the hypothesis that, contrary to the view of Carr and Lee, accurate estimation is possible, such an objection relies also on the assumption of a tight connection between the past and future that we relax. We are interested in situations where realized past volatility may not be a reliable predictor of volatility in the future. The rationale is that the stochastic environment is often too complex for a sophisticated individual to believe that her theory, whether of volatility or of other variables, captures all aspects. Being sophisticated, she is aware of the incompleteness of her theory. Accordingly, when planning ahead she believes there may be time-varying factors excluded by her theory that she understands poorly and that are difficult to identify statistically. Thus she perceives ambiguity when looking into the future. The amount of ambiguity may depend on past observations, and may be small for some histories, but it cannot be excluded a priori.

A similar rationale for ambiguity is emphasized by Epstein and Schneider (2008, 2010). Nonstationarity is emphasized by Ilut and Schneider (2011) in their model of business cycles driven by ambiguity. In finance, Lo and Mueller (2010) argue that the (perceived) failures of the dominant paradigm, for example, in the context of the recent crisis, are due to inadequate attention paid to the kind of uncertainty faced by agents and modelers. Accordingly, they suggest a new taxonomy of uncertainty that extends the dichotomy between risk and ambiguity (or "Knightian uncertainty"). In particular, they refer to *partially reducible uncertainty* to describe "situations in which there is a limit to what we can deduce about the underlying phenomena generating the data. Examples include data-generating processes that exhibit: (1) stochastic or time-varying parameters that vary too frequently to be estimated accurately; (2) nonlinearities too complex to be captured by existing models, techniques and datasets; (3) nonstationarities and non-ergodicities that render useless the Law of Large Numbers, Central Limit Theorem, and other methods of statistical inference and approximation; and (4) the dependence on relevant but unknown and unknowable conditioning information." Lo and Mueller do not offer a model. One can view this paper as an attempt to introduce some of their concerns into continuous-time modeling and particularly into formal asset pricing theory.

The natural question is whether and in what form the cornerstones of received asset pricing theory extend to a framework with ambiguous volatility. Some initial steps in answering this question are provided in Section 3.[3] A notable finding is that both equilibrium and "no-arbitrage" asset prices can be characterized by means of "state prices" even though the analysis cannot be undertaken in a probability space framework (which precludes talking about state price *densities* or about *equivalent* martingale, or risk-neutral, measures).

1.2 Discrete-time trees

Time varies over $\{0,h,2h,\ldots,(n-1)h,nh\}$, where $0<h<1$ scales the period length and n is a positive integer with $nh=T$. Uncertainty is driven by the colors of balls drawn from a sequence of urns. It is known that each urn contains 100 balls that are either red (R), green (G), or yellow (Y), and that the urns are constructed independently (informally speaking). A ball is drawn from each urn, and the colors drawn determine the evolution of the state variable $B=(B_t)$ according to: $B_0=0$ and, for $t=h,\ldots,nh$,

$$dB_t \equiv B_t - B_{t-h} = \begin{cases} h^{1/2} & \text{if } R_t \\ -h^{1/2} & \text{if } G_t \\ 0 & \text{if } Y_t \end{cases}$$

We describe three alternative assumptions regarding the additional information available about the urns. They provide intuition for continuous-time models where (respectively) the driving process is (i) a standard Brownian motion, (ii) a Brownian motion modified by ambiguous drift, and (iii) a Brownian motion modified by ambiguous volatility. The first two are included in order to provide perspective on the third.

Scenario 1: You are told further that $Y=0$ and that $R=G$ for each urn (thus all urns are known to have the identical composition). The state process (B_t)

[3] Early work on the pricing of derivative securities when volatility is ambiguous includes Lyons (1995) and Avellaneda, Levy, and Paras (1995). See Section 3 for the relation to our analysis and for additional references.

can be described equivalently in terms of the measure $p_0 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$ and its i.i.d. product that induces a measure P_0 on trajectories of B. Thus, we have a random walk that, by Donsker's theorem, converges weakly to a standard Brownian motion in the continuous-time limit as $h \to 0$ (see Billingsley 1999, for example).

Scenario 2: You are told again that $Y = 0$ for every urn. However, you are given less information than previously about the other colors. Specifically, you are told that for each urn the proportion of R lies in the interval $\left[\frac{1}{2} - \frac{1}{2}\kappa h^{1/2}, \frac{1}{2} + \frac{1}{2}\kappa h^{1/2}\right]$, for some fixed $\kappa > 0$. Thus the composition of the urn at any time t could be given by a measure of the form $p^{\mu_t} = \left(\frac{1}{2} + \frac{1}{2}\mu_t h^{1/2}, \frac{1}{2} - \frac{1}{2}\mu_t h^{1/2}, 0\right)$, for some μ_t satisfying $|\mu_t| \leq \kappa$. The increment dB_t has mean and variance under p^{μ_t} given by[4]

$$E(dB_t) = \mu_t h \qquad \text{and} \qquad var\,(dB_t) = h - (\mu_t h)^2 = h + o(h).$$

Accordingly, the weaker information about the composition of urns implies ambiguity about drift per unit time, but up to the $o(h)$ approximation, it does not affect the corresponding one-step-ahead variance.

The preceding is the building block of the Chen and Epstein (2002) continuous-time model of ambiguity about drift.[5] The transition from discrete to continuous time amounts to a minor variation of the convergence result noted for Scenario 1 (see Skiadas 2011 for some details). The sets $\{p^{\mu_t} : |\mu_t| \leq \kappa\}$, $t = 0, h, 2h, \ldots, (n-1)h, nh$ of one-step-ahead measures can be combined to construct a set $\mathcal{P}$ of priors over the set Ω of possible trajectories for B. It is not difficult to see that the priors are mutually *equivalent*—that is, they all agree on which events are null (have zero probability). Further, as described in Section 2.1, equivalence holds in the continuous-time limit. Consequently, the model with ambiguous drift can be formulated within a probability space framework with ambient probability measure P_0 according to which B is a standard Brownian motion. Alternative hypotheses about the true probability law can be expressed via densities with respect to P_0.

Scenario 3: Turn now to a model having ambiguity only about volatility. You are told that $R = G$, thus eliminating uncertainty about the relative composition of R versus G. However, the information about Y is weakened, and you are told only that $Y \leq 20$.[6]

Any probability measure over trajectories consistent with these facts makes B a martingale. In that sense, there is certainty that B is a martingale. However, the one-step-ahead variance $\sigma_t^2 h$ depends on the number of yellow balls and thus is ambiguous; it equals $p_t h$, where we know only that $0 \leq 1 - p_t \leq 0.2$, or

$$.8 = \underline{\sigma}^2 \leq \sigma_t^2 \leq \overline{\sigma}^2 = 1. \qquad (2)$$

4 $o(h)$ represents a function such that $o(h)/h \to 0$ as $h \to 0$.

5 It corresponds to the special case of their model called κ-ignorance.

6 This scenario is adapted from Levy, Avellaneda, and Paras (1998).

Because urns are perceived to be independent, they may differ in actual composition. Therefore, any value for σ_t in the interval $[\underline{\sigma},\overline{\sigma}]$ could apply at any time. Independence implies also that past draws do not reveal anything about the future, and ambiguity is left undiminished. This is an extreme case that is a feature of this example and is a counterpart of the assumption of i.i.d. increments in the binomial tree.

By a generalization of Donsker's theorem (Yuan 2011), the above trinomial model converges weakly (or "in distribution") to a continuous-time model on the interval $[0,T]$ as the time period length h goes to 0.[7]

The limiting continuous-time model inherits from the discrete-time trinomial the interpretation that it models certainty that the driving process $B=(B_t)$ is a martingale, whereas volatility is known only up to the interval $[\underline{\sigma},\overline{\sigma}]$. To be more precise about the meaning of volatility, let the quadratic variation process of B be defined by

$$\langle B\rangle_t(\omega)=\lim_{\triangle t_k\to 0}\sum_{t_k\le t}\left|B_{t_{k+1}}(\omega)-B_{t_k}(\omega)\right|^2 \tag{3}$$

where $0=t_1<\ldots<t_n=t$ and $\triangle t_k=t_{k+1}-t_k$.[8] Then the volatility (σ_t) of B is defined by

$$d\langle B\rangle_t=\sigma_t^2 dt.$$

Therefore, the interval constraint on volatility can be written also in the form

$$\underline{\sigma}^2 t\le\langle B\rangle_t\le\overline{\sigma}^2 t. \tag{4}$$

The preceding defines the stochastic environment. Consumption and other processes are defined accordingly (for example, they are required to be adapted to the natural filtration generated by B). We emphasize that our model is much more general than suggested by this outline. Importantly, the interval $[\underline{\sigma},\overline{\sigma}]$ can be time- and state-varying, and the dependence on history of the interval at time t is unrestricted, thus *permitting any model of how ambiguity varies with observation (that is, learning) to be accommodated.* In addition, we admit multi-dimensional driving processes and also ambiguity about both drift and volatility.

As noted earlier, ambiguity about volatility leads to a set of *nonequivalent* priors—that is, to disagreement between priors as to what events are possible

[7] To clarify the meaning of weak convergence, consider the set Ω of continuous trajectories on $[0,T]$ that begin at the origin. For each period length h, identify any discrete-time trajectory with a continuous path on $[0,T]$ obtained by linear interpolation. Then each hypothesis about the compositions of all urns implies a probability measure on Ω. By varying over all hypotheses consistent with the above description of the urns, one obtains a set $\mathcal{P}^h$ of probability laws on Ω. Suppose that $\mathcal{P}$ is a given set of measures on Ω. Then say that $\mathcal{P}^h$ converges weakly to $\mathcal{P}$ if $\sup_{P\in\mathcal{P}^h}E^P f$ converges to $\sup_{P\in\mathcal{P}}E^P f$ for every function $f:\Omega\to\mathbb{R}$ that is bounded and suitably continuous. The cited result by Yuan implies this convergence for the set $\mathcal{P}$ constructed below corresponding to the special case of our model for which volatility is constrained by (2). See also Dolinsky, Nutz, and Soner (2012) for a related result.

[8] By Follmer (1981) and Karandikar (1995), the above limit exists almost surely for every measure that makes B a martingale. Because there is certainty that B is a martingale, this limited universality is all we need.

(or to ambiguity about what is possible). To see this, let B be a Brownian motion under P_0 and denote by $P^{\underline{\sigma}}$ and $P^{\overline{\sigma}}$ the probability distributions over continuous paths induced by the two processes $(\underline{\sigma} B_t)$ and $(\overline{\sigma} B_t)$. Given the ambiguity described by (4), $P^{\underline{\sigma}}$ and $P^{\overline{\sigma}}$ are two alternative hypotheses about the probability law driving uncertainty. It is apparent that they are mutually singular, and hence not equivalent, because[9]

$$P^{\underline{\sigma}}\left(\left\{\langle B\rangle_T = \underline{\sigma}^2 T\right\}\right) = 1 = P^{\overline{\sigma}}\left(\left\{\langle B\rangle_T = \overline{\sigma}^2 T\right\}\right). \tag{5}$$

We caution against a possible conceptual misinterpretation of (5). If $P^{\underline{\sigma}}$ and $P^{\overline{\sigma}}$ were the only two hypotheses being considered, then ambiguity could be eliminated quickly because one can approximate volatility locally as in (3) and thus use observations on a short time interval to differentiate between the two hypotheses. This is possible because of the tight connection between past and future volatility imposed in each of $P^{\underline{\sigma}}$ and $P^{\overline{\sigma}}$. As discussed earlier, this is not the kind of ambiguity we have in mind. The point of (5) is only to illustrate nonequivalence as simply as possible. Importantly, such nonequivalence of priors is a feature also of the more complex and interesting cases at which the model is directed.

An objection to modeling ambiguity about possibility might take the form: "If distinct priors (or models) are not equivalent, then one can discriminate between them readily. Therefore, when studying nontransient ambiguity there is no loss in restricting priors to be equivalent." The connection between past and future is again the core issue (as in the preceding subsection). Consider the individual at time t and her beliefs about the future. The source of ambiguity is her concern with locally time-varying and poorly understood factors. This limits her confidence in predictions about the immediate future, or "next step," to a degree that depends on history but that is not eliminated by the retrospective empirical discrimination between models. At a formal level, Epstein and Schneider (2003) show that when backward induction reasoning is added to multiple priors utility, then the individual behaves *as if* the set of conditionals entertained at any time t and state does not vary with marginal prior beliefs on time t measurable uncertainty. (They call this property rectangularity.) Thus, looking back on past observations at t, even though the individual might be able to dismiss some priors or models as being inconsistent with the past, this is unimportant for prediction because the set of conditional beliefs about the future is unaffected.

2. Utility

Many components of the formal setup are typical in continuous-time asset pricing. Time t varies over the finite horizon $[0,T]$. Paths or trajectories of the

[9] Two measures P and P' on Ω are singular if there exists $A \subset \Omega$ such that $P(A)=1$ and $P'(A)=0$. They are equivalent, if for every A, $P(A)=0$ if and only if $P'(A)=0$. Thus, P and P' singular implies that they are not equivalent, but the converse is false.

driving process are assumed to be continuous and thus are modeled by elements of $C^d([0,T])$, the set of all $\mathbb{R}^d$-valued continuous functions on $[0,T]$, endowed with the sup norm. The generic path is $\omega=(\omega_t)_{t\in[0,T]}$, where we write ω_t instead of $\omega(t)$. All relevant paths begin at 0, and thus we define the canonical state space to be

$$\Omega=\left\{\omega=(\omega_t)\in C^d([0,T]):\omega_0=0\right\}.$$

The coordinate process (B_t), where $B_t(\omega)=\omega_t$, is denoted by B. Information is modeled by the filtration $\mathcal{F}=\{\mathcal{F}_t\}$ generated by B. Let P_0 be the Wiener measure on Ω so that B is a Brownian motion under P_0.

Consumption processes c take values in C, a convex subset of $\mathbb{R}^\ell$. The objective is to formulate a suitable utility function on a domain D of consumption processes.

2.1 Recursive utility under equivalence

For perspective, we begin by outlining the Chen-Epstein model, where there is ambiguity only about drift. This is the continuous-time counterpart of Scenario 2 in Section 1.2.

If P_0 describes the individual's beliefs, then following Duffie and Epstein (1992b), utility may be defined by:[10]

$$V_t^{P_0}(c)=E^{P_0}\left[\int_t^T f(c_s,V_s^{P_0})ds\,\middle|\,\mathcal{F}_t\right],\quad 0\le t\le T. \tag{6}$$

Here $V_t^{P_0}$ gives the utility of the continuation $(c_s)_{s\ge t}$ and $V_0^{P_0}$ is the utility of the entire process c. The function f is a primitive of the specification, called an *aggregator*. The most commonly used aggregator has the form

$$f(c_t,v)=u(c_t)-\beta v,\quad \beta\ge 0, \tag{7}$$

which delivers the expected utility specification

$$V_t(c)=E^{P_0}\left[\int_t^T e^{-\beta(s-t)}u(c_s)ds\,\middle|\,\mathcal{F}_t\right]. \tag{8}$$

The use of more general aggregators permits a partial separation of risk aversion from intertemporal substitution.

To admit a concern with model uncertainty, Chen and Epstein replace the single measure P_0 by a set $\mathcal{P}^\Theta$ of measures equivalent to P_0. This is done by specifying a suitable set of densities. For each well-behaved $\mathbb{R}^d$-valued process $\theta=(\theta_t)$, called a density generator, let

$$z_t^\theta\equiv\exp\left\{-\frac{1}{2}\int_0^t|\theta_s|^2ds-\int_0^t\theta_s^\top dB_s\right\},\quad 0\le t\le T,$$

10 Below we often suppress c and write V_t instead of $V_t(c)$. The dependence on the state ω is also frequently suppressed.

and let P^θ be the probability measure on $(\Omega, \mathcal{F})$ with density z_T^θ, that is,

$$\frac{dP^\theta}{dP_0} = z_T^\theta; \text{ more generally, } \left.\frac{dP^\theta}{dP}\right|_{\mathcal{F}_t} = z_t^\theta \text{ for each } t. \tag{9}$$

Given a set Θ of density generators, the corresponding set of priors is

$$\mathcal{P}^\Theta = \left\{ P^\theta : \theta \in \Theta \text{ and } P^\theta \text{ is defined by (9)} \right\}. \tag{10}$$

By construction, all measures in $\mathcal{P}^\Theta$ are equivalent to P_0. Because the role of P_0 is only to define null events, any other member of $\mathcal{P}^\Theta$ could equally well serve as the reference measure.

Continuation utilities are defined by:

$$V_t = \inf_{P \in \mathcal{P}^\Theta} E^P \left[\int_t^T f(c_s, V_s) ds \,\middle|\, \mathcal{F}_t \right]. \tag{11}$$

An important property of the utility process is dynamic consistency, which follows from the following recursivity: For every c in D,

$$V_t = \min_{P \in \mathcal{P}^\Theta} E_P \left[\int_t^\tau f(c_s, V_s) ds + V_\tau \,\middle|\, \mathcal{F}_t \right], \quad 0 \le t < \tau \le T. \tag{12}$$

Regarding interpretation, by the Girsanov theorem $B_t + \int_0^t \theta_s ds$ is a Brownian motion under P^θ. Thus as θ varies over Θ and P^θ varies over $\mathcal{P}^\Theta$, alternative hypotheses about the drift of the driving process are defined. Accordingly, the infimum suggests that the utility functions V_t exhibit an aversion to ambiguity about the drift. Because B_t has a variance-covariance matrix equal to the identity according to all measures in $\mathcal{P}^\Theta$, there is no ambiguity about volatility. Neither is there any uncertainty about what is possible, because P_0 defines which events are null.

There is a limited sense in which the preceding framework is adequate for modeling also ambiguity about volatility. For example, suppose that the driving process is (X_t), where $dX_t = \sigma_t dB_t$ (where B is a Brownian motion under P_0) and where the volatility is thought to evolve according to

$$d\sigma_t = \theta_t dt + v_t dB_t.$$

Here the drift (θ_t) is ambiguous in the above sense, and the volatility of volatility (v_t) is a fixed stochastic process—for example, it might be constant, as in many stochastic volatility models. Thus, the difficulty of finding a specification for (σ_t) in which one can have complete confidence is moved one level from volatility to its volatility. This constitutes progress if there is greater evidence about vol of vol and if model implications are less sensitive to misspecifications of the latter. We suspect that in many modeling situations neither is true. Moreover, this approach cannot intermediate between, or robustify, the stochastic volatility models that have been used in the empirical literature (Section 1.1).

2.2 The set of priors

The objective is to specify beliefs, in the form of a set of priors generalizing (10) that captures ambiguity about both drift and volatility. Another key ingredient is conditioning. The nonequivalence of priors (illustrated by (5)) poses a particular difficulty for updating because of the need to update beliefs conditional on events having zero probability according to some, but not all, priors. Once these steps are completed, continuation utilities can be defined (apart from technical details) as in (11); see the next section.

The construction of the set of priors can be understood by referring back to the binomial and trinomial examples in Section 1.2. In Scenario 2, the composition of all urns is specified by fixing $\mu = (\mu_t)$ with $|\mu_t| \leq \kappa$. Define $X^\mu = \left(X_t^\mu\right)$ by

$$dX_t^\mu = \mu_t h + dB_t, \quad X_0^\mu = 0.$$

Then X^μ and P_0 induce a distribution P^μ over trajectories, and as one varies over all choices of μ, one obtains the set of priors $\mathcal{P}$ described earlier. Thus, beliefs are described indirectly through the set $\{\mu = (\mu_t) : |\mu_t| \leq \kappa\}$ of alternative hypotheses about the drift of the driving process. The case is similar in Scenario 3, where the composition of all urns is specified by by fixing $\sigma = (\sigma_t)$ with $\underline{\sigma} \leq \sigma_t \leq \overline{\sigma}$. If $X^\sigma = \left(X_t^\sigma\right)$ is defined by

$$dX_t^\sigma = \sigma_t dB_t, \quad X_0^\sigma = 0,$$

then X^σ and P_0 induce a distribution P^σ over trajectories, and as one varies over all choices of σ, one obtains the set of priors $\mathcal{P}$ described earlier. Thus beliefs are described indirectly through a set of alternative hypotheses about the volatility of the driving process.

The preceding construction is readily generalized to permit a vector-valued driving process ($d \geq 1$), ambiguity about both drift and volatility, and to allow ambiguity at any time t to depend on history. We describe the corresponding construction in continuous time.[11]

The individual is not certain that the driving process has zero drift and/or unit variance (where $d = 1$). Accordingly, she entertains a range of alternative hypotheses $X^\theta = (X_t^\theta)$ parameterized by θ. Here $\theta_t = (\mu_t, \sigma_t)$ is an $\mathcal{F}$-progressively measurable process with values in $\mathbb{R}^d \times \mathbb{R}^{d \times d}$ that describes a conceivable process for drift $\mu = (\mu_t)$ and for volatility $\sigma = (\sigma_t)$.[12] Available information leads to the constraint on drift and volatility pairs given by

$$\theta_t(\omega) \in \Theta_t(\omega), \qquad \text{for all } (t, \omega), \tag{13}$$

where $\Theta_t(\omega)$ is a subset of $\mathbb{R}^d \times \mathbb{R}^{d \times d}$. The Θ_t's are primitives of the model.[13] In the trinomial model expanded in the obvious way to include also ambiguity

[11] See Epstein and Ji (2013) for a general and mathematically rigorous development.

[12] Write $\theta = (\mu, \sigma)$.

[13] See our companion paper for the technical regularity conditions assumed for (Θ_t).

about drift,

$$\Theta_t(\omega)=[-\kappa,\kappa]\times[\underline{\sigma},\overline{\sigma}], \quad \text{for all } (t,\omega).$$

In general, the dependence of $\Theta_t(\omega)$ on the history corresponding to state ω permits the model to accommodate learning. Moreover, because the form of history dependence is unrestricted (apart from technical regularity conditions), so is the nature of learning. Just as in the Chen-Epstein model, we provide a framework within which additional structure modeling learning can be added.

Two other examples might be helpful. The robust stochastic volatility model described in Section 1.1 corresponds to taking

$$\Theta_t(\omega)=\{0\}\times\left[\underline{\sigma}_t(\omega),\overline{\sigma}_t(\omega)\right],$$

where $\underline{\sigma}_t(\omega)$ and $\overline{\sigma}_t(\omega)$ are given by (1). When $d>1$, one way to robustify is through the restriction

$$\Theta_t(\omega)=\left\{\sigma\in\mathbb{R}^{d\times d}:\sigma_t^1(\omega)\left(\sigma_t^1(\omega)\right)^\intercal\leq\sigma\sigma^\intercal\leq\sigma_t^2(\omega)\left(\sigma_t^2(\omega)\right)^\intercal\right\},$$

though other natural specifications exist in the multidimensional case.

The model is flexible in the way it relates ambiguity about drift and ambiguity about volatility. Thus, as a final example, suppose that drift and volatility are thought to move together. Then, joint ambiguity is captured by specifying

$$\Theta_t(\omega)=\left\{(\mu,\sigma)\in\mathbb{R}^2:\mu=\mu_{\min}+z,\ \sigma^2=\sigma_{\min}^2+2z/\gamma,\ 0\leq z\leq\overline{z}_t(\omega)\right\}, \quad (14)$$

where $\mu_{\min}$, $\sigma_{\min}^2$ and $\gamma>0$ are fixed and known parameters.[14]

Given a hypothesis θ about drift and volatility, the implication for the driving process is that it is given by the unique solution $X^\theta=\left(X_t^\theta\right)$ to the following stochastic differential equation (SDE) under P_0:

$$dX_t^\theta=\mu_t\left(X_\cdot^\theta\right)dt+\sigma_t\left(X_\cdot^\theta\right)dB_t, \quad X_0^\theta=0, t\in[0,T]. \quad (15)$$

We restrict the process θ further so that a unique strong solution X^θ to the SDE exists. Denote by Θ the set of all processes θ satisfying the latter and also (13).

As in the discrete-time examples, X^θ and P_0 induce a probability measure P^θ on $(\Omega,\mathcal{F}_T)$:

$$P^\theta(A)=P_0\left(\left\{\omega:X^\theta(\omega)\in A\right\}\right), \quad A\in\mathcal{F}_T. \quad (16)$$

Therefore, we arrive at the set $\mathcal{P}^\Theta$ of priors on the set of continuous trajectories given by

$$\mathcal{P}^\Theta=\left\{P^\theta:\theta\in\Theta\right\}. \quad (17)$$

Fix Θ and denote the set of priors simply by $\mathcal{P}$. This set of priors is used, as in the Gilboa-Schmeidler model, to define utility and guide choice between consumption processes.

[14] This specification is adapted from Epstein and Schneider (2010).

Remark 1. Here is a recap. The set $\mathcal{P}$ consists of priors over Ω, the space of continuous trajectories for the driving process. B denotes the coordinate process, $B_t(\omega)=\omega_t$. It is a Brownian motion under P_0, which may or may not lie in $\mathcal{P}$, but B is typically not a Brownian motion relative to (other) priors in $\mathcal{P}$. Indeed, different priors P typically imply different conditional expectations $E_t^P\left(B_{t+\triangle t}-B_t\right)$ and $E_t^P\left(\langle B\rangle_{t+\triangle t}-\langle B\rangle_t\right)$. This is the justification for interpreting $\mathcal{P}$ as modeling ambiguity about the drift and volatility of the driving process. The preceding should be viewed as one way to construct $\mathcal{P}$, but not necessarily as a description of the individual's thought processes. The model's objective is to describe behavior that can be thought of "as if" being derived from a maxmin objective function using the above set of priors.

The construction of utility requires that first we show how beliefs, through the set $\mathcal{P}$, lead to natural definitions of "expectation" and "conditional expectation." The former is straightforward. For any random variable ξ on $(\Omega,\mathcal{F}_T)$, if $\sup_{P\in\mathcal{P}} E_P\xi<\infty$, define its (nonlinear) expectation by

$$\hat{E}\xi=\sup_{P\in\mathcal{P}} E_P\xi. \tag{18}$$

Because we will assume that the individual is concerned with worst-case scenarios, below we use the fact that

$$\inf_{P\in\mathcal{P}} E_P\xi=-\hat{E}[-\xi].$$

Conditional beliefs and expectations are not as clear cut because of the need, mentioned above, to update beliefs conditional on events having zero probability according to some priors. A naive approach to defining conditional expectation would be to use the standard conditional expectation $E_P[\xi\mid\mathcal{F}_t]$ for each P in $\mathcal{P}$ and then to take the (essential) supremum over $\mathcal{P}$. Such an approach immediately encounters a roadblock due to the nonequivalence of priors. The conditional expectation $E_P[\xi\mid\mathcal{F}_t]$ is well defined only P-almost surely, while speaking informally, conditional beliefs and expectations must be defined at every node deemed possible by some prior in $\mathcal{P}$. The economic rationale is simple. Suppose that P and P' are two nonequivalent priors held (for example) at time 0, and consider updating at $t>0$. Let A be an event, measurable at time t, such that $P(A)=0<P'(A)$. Then A is conceivable according to the individual's ex ante perception. Consequently, beliefs at time t conditional on A are relevant both for ex post choice and also for ex ante choice—for example, if the individual reasons by backward induction. Therefore, the ex ante perception represented by P should also be updated there, even though A was deemed impossible ex ante according to P.

This difficulty can be overcome because for every admissible hypothesis θ, $\theta_t(\omega)$ is defined for *every* (t,ω)—that is, the primitives specify a hypothesized instantaneous drift-volatility pair everywhere in the tree. This feature of the model resembles the approach adopted in the theory of extensive form games—namely the use of conditional probability systems, whereby conditional beliefs at *every* node are specified as primitives, obviating the need to update.[15] We show that a solution to the updating problem is also available here, though it requires nontrivial mathematical arguments; see Epstein and Ji (2013) for details, rigorous statements and supporting proofs.

To outline it, let $\theta = (\theta_s)$ be a conceivable scenario ex ante and fix a node (t, ω). By definition of θ, the continuation of θ is seen by the individual ex ante as a conceivable continuation from time t along the history ω. We assume that then it is also seen as a conceivable scenario ex post conditionally on (t, ω), thus ruling out surprises or unanticipated changes in outlook. Then, paralleling (15), each such conditional scenario has an implication for the driving process conditionally on (t, ω). The implied process and P_0 induce a probability measure $P_t^{\theta,\omega}$ on Ω, denoted simply by P_t^ω with θ suppressed when it is understood that $P = P^\theta$. The crucial facts are that, for each P in $\mathcal{P}$, (i) P_t^ω is defined for every t and ω, and (ii) P_t^ω is a version of the regular $\mathcal{F}_t$-conditional probability of P.[16] The set of all such conditionals obtained as θ varies over Θ is denoted $\mathcal{P}_t^\omega$, that is,

$$\mathcal{P}_t^\omega = \left\{ P_t^\omega : P \in \mathcal{P} \right\}. \tag{19}$$

We take $\mathcal{P}_t^\omega$ to be the individual's set of priors conditional on (t, ω). Then, the conditional expectation of suitable random variables ξ is defined by

$$\hat{E}[\xi \mid \mathcal{F}_t](\omega) = \sup_{P \in \mathcal{P}_t^\omega} E_P \xi, \quad \text{for every } (t, \omega).$$

This completes the prerequisites regarding beliefs for defining utility.

2.3 The definition of utility

Let D be a domain of consumption processes and defer elaboration until the next section. For each c in D, define its continuation utility $V_t(c)$, or simply V_t, by

$$V_t = -\hat{E}\left[-\int_t^T f(c_s, V_s) ds \middle| \mathcal{F}_t \right]. \tag{20}$$

This definition parallels the Chen and Epstein definition (11). In particular, f is an aggregator that is assumed to satisfy suitable measurability, Lipschitz, and integrability conditions. Under these conditions and for a suitable domain D, there is a unique utility process $(V_t(c))$ solving (20) for each c in D—that is, utility is well defined.

[15] It also resembles the approach in the discrete-time model in Epstein and Schneider (2003), where roughly, conditional beliefs about the next instant for every time and history are adopted as primitives and are pasted together by backward induction to deliver the ex ante set of priors.

[16] For any probability measure P on the canonical space Ω, a corresponding *regular $\mathcal{F}_t$-conditional probability* P_t^ω is defined to be any mapping $P_t^\omega : \Omega \times \mathcal{F}_T \to [0,1]$ satisfying the following conditions: (i) for any ω, P_t^ω is a probability measure on $(\Omega, \mathcal{F}_T)$, (ii) for any $A \in F_T$, $\omega \to P_t^\omega(A)$ is $\mathcal{F}_t$-measurable, and (iii) for any $A \in \mathcal{F}_T$, $E^P[1_A \mid \mathcal{F}_t](\omega) = P_t^\omega(A)$, P-$a.s.$

For the standard aggregator (7), utility admits the closed-form expression

$$V_t(c)=-\hat{E}\left[-\int_t^T u(c_s)e^{-\beta s}ds\big|\mathcal{F}_t\right]. \tag{21}$$

More generally, closed-from expressions are rare.

The following example illustrates the effect of volatility ambiguity.

Example 1 (Closed form). Suppose that there is no ambiguity about the drift, and that ambiguity about volatility is captured by the fixed interval $[\underline{\sigma},\overline{\sigma}]\subset\mathbb{R}_{++}$. Consider consumption processes that are certain and constant, at level 0, for example, on the time interval $[0,1)$, and that yield constant consumption on $[1,T]$ at a level that depends on the state ω_1 at time 1. Specifically, let

$$c_t(\omega)=\psi(\omega_1), \quad \text{for } 1\leq t\leq T,$$

where $\psi:\mathbb{R}^1\to\mathbb{R}^1_+$. For simplicity, suppose further that u is linear. Then time 0 utility evaluated using (21) is, ignoring irrelevant constants,

$$V_0=-\hat{E}[-\psi(\omega_1)].$$

If ψ is a convex function, then[17]

$$V_0(c)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\psi\left(\underline{\sigma}^2 y\right)\exp\left(-\frac{y^2}{2}\right)dy,$$

and if ψ is concave, then

$$V_0(c)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\psi\left(\bar{\sigma}^2 y\right)\exp\left(-\frac{y^2}{2}\right)dy. \tag{22}$$

There is an intuitive interpretation for these formulae. Given risk neutrality, the individual cares only about the expected value of consumption at time 1. The issue is expectation according to which probability law? For simplicity, consider the following concrete specifications:

$$\psi_1(x)=|x-\kappa|, \quad \text{and} \quad \psi_2(x)=-|x-\kappa|.$$

Then ψ_1 is convex and ψ_2 is concave. If we think of the driving process as the price of a stock, then $\psi_1(\cdot)$ can be interpreted as a straddle—the sum of a European put and a European call option on the stock at the common strike price κ and expiration date 1. (We are ignoring nonnegativity constraints.) A straddle pays off if the stock price moves, whether up or down, and thus constitutes a bet on volatility. Accordingly, the worst-case scenario is that the price process has the lowest possible volatility $\underline{\sigma}$. In that case, ω_1 is $N\left(0,\underline{\sigma}^2\right)$ and the indicated expected value of consumption follows. Similarly, $\psi_2(\cdot)$ describes the

[17] See Levy, Avellaneda, and Paras (1998) and Peng (2010).

corresponding short position and amounts to a bet against volatility. Therefore, the relevant volatility for its worst-case evaluation is the largest possible value $\overline{\sigma}$, consistent with the expression for utility given above.

When the function ψ is neither concave nor convex globally, closed-form expressions for utility are available only in extremely special and unrevealing cases. However, a generalization to d-dimensional processes, $d \geq 1$, is available and will be used below. Let there be certainty that the driving process is a martingale and let the volatility matrix σ_t in (15) be restricted to lie in the compact and convex set $\Gamma \subset \mathbb{R}^{d\times d}$ such that, for all σ in Γ, $\sigma\sigma^\top \geq \hat{a}$ for some positive definite matrix $\widehat{a}$. Consumption is as above except that, for some $a \in \mathbb{R}^d$,

$$c_t(\omega) = \psi(a^\top \omega_1), \quad \text{for } 1 \leq t \leq T.$$

Let $\underline{\sigma}$ be any solution to $\min_{\sigma\in\Gamma} tr\left(\sigma\sigma^\top aa^\top\right)$ and let $\overline{\sigma}$ be any solution to $\max_{\sigma\in\Gamma} tr\left(\sigma\sigma^\top aa^\top\right)$. If ψ is convex (concave), then the worst-case scenario is that $\sigma_t = \underline{\sigma}$ $(\overline{\sigma})$ for all t. Closed-form expressions for utility follow immediately.

The domain D of consumption processes, and the ambient space containing utility processes, are defined precisely in our companion paper (see also the next section). Here we mention briefly a feature of these definitions that reveals a great deal about the nature of formal analysis when priors are not equivalent. When the ambient framework is a probability space (Ω, P_0), and thus P_0 is used to define null events, then random variables and stochastic processes are defined only up to the P_0-almost sure qualification. Thus P_0 is an essential part of the definition of all formal domains. However, ambiguity about volatility leads to a set $\mathcal{P}$ of priors that do not agree about which events have zero probability. Therefore, we follow Denis and Martini (2006) and define appropriate domains of stochastic processes by using the entire set of probability measures $\mathcal{P}$. Accordingly, two consumption processes c' and c are identified, and we write $c' = c$, if for every t, $c'_t(\omega) = c_t(\omega)$ for every ω in $G_t \subset \Omega$, where $P(G_t) = 1$ for *all* P in $\mathcal{P}$.[18] We abbreviate the preceding in the form: for every t,

$$c'_t = c_t \quad \mathcal{P}\text{-}a.s.$$

Loosely put, the latter means that the two processes are certain to yield identical consumption levels regardless of which prior in $\mathcal{P}$ is the true law describing the driving process. Put another way, a consumption process as defined formally herein is portrayed in greater detail than if it were seen from the perspective of any single prior in $\mathcal{P}$. For example, if the two priors P_1 and P_2 are singular,

[18] The following perspective may be helpful for nonspecialists in continuous-time analysis. In the classical case of a probability space $(\Omega, \mathcal{F}, P)$ with filtration $\{\mathcal{F}_t\}$, if two processes X and Y satisfy "for each t, $X_t = Y_t$ P-$a.s.$," then Y is called a *modification* of X. If $P(\{\omega : X_t = Y_t \ \forall t \in [0,T]\}) = 1$, then X and Y are said to be *indistinguishable*. These notions are equivalent when restricted to X and Y having a.s. right continuous sample paths, but the second is stronger in general. We point out, however, that the sense in which one constructs a Brownian motion exhibiting continuous sample paths is that a suitable modification exists (see the Kolmogorov-Centsov theorem). A reference for the preceding is Karatzas and Shreve (1991, 2, 53).

then each provides a description of consumption on a subset Ω_i, $i=1,2$, of the set of possible trajectories of the driving process, where Ω_1 and Ω_2 are disjoint, while using the entire set $\mathcal{P}$ yields a description of consumption on $\Omega_1 \cup \Omega_2$ and more.

Remark 2. If every two priors are singular, then the statement $c_t' = c_t$ $\mathcal{P}$-$a.s.$ amounts to standard probability 1 statements on disjoint parts of the state space, and thus is not far removed from a standard equation in random variables. However, singularity is not representative—it can be shown that the set $\mathcal{P}$ typically contains priors that, though not equivalent, are also not mutually singular. Thus, the reader is urged not to be overly influenced by the example of singular priors, which we use often only because it starkly illustrates nonequivalence.

Equations involving processes other than consumption processes are given similar meanings. For example, the equality (20) should be understood to hold $\mathcal{P}$-almost surely for each t. Similar meaning is given also to inequalities.

Utility has a range of natural properties. Most noteworthy is that the process (V_t) satisfies the recursive relation

$$V_t = -\widehat{E}\left[-\int_t^\tau f(c_s, V_s) ds - V_\tau \,\Big|\, \mathcal{F}_t \right], \qquad 0 \le t < \tau \le T. \tag{23}$$

Such recursivity is typically thought to imply dynamic consistency. However, the nonequivalence of priors complicates matters, as we describe next.

The noted recursivity does imply the following weak form of dynamic consistency: For any $0 < \tau < T$, and any two consumption processes c' and c that coincide on $[0, \tau]$,

$$\left[V_\tau\left(c'\right) \ge V_\tau(c) \;\; \mathcal{P}\text{-}a.s. \right] \Longrightarrow V_0\left(c'\right) \ge V_0(c).$$

Typically (see Duffie and Epstein 1992b, for example), dynamic consistency is defined so as to deal also with strict rankings—that is, if also $V_\tau(c') > V_\tau(c)$ on a nonnegligible set of states, then $V_0(c') > V_0(c)$. This added requirement rules out the possibility that c' is chosen ex ante though it is indifferent to c, and yet it is not implemented fully because the individual switches to the conditionally strictly preferable c for some states at time τ. The issue is how to specify "nonnegligible." When all priors are equivalent, then positive probability according to any single prior is the natural specification. However, in the absence of equivalence, a similarly natural specification is unclear. A simple illustration of the consequence is given in the next example.

Example 2 (Weak dynamic consistency). Take $d=1$. Let the endowment process e satisfy (under P_0)

$$d \log e_t = \sigma_t dB_t,$$

or equivalently,

$$de_t/e_t = \frac{1}{2}\sigma_t^2 dt + \sigma_t dB_t \quad P_0\text{-}a.s.$$

Here volatility is restricted only by $0 < \underline{\sigma} \le \sigma_t \le \overline{\sigma}$. Utility is defined, for any consumption process c, by

$$V_0(c) = \inf_{P \in \mathcal{P}} E^P \left[\int_0^T e^{-\beta t} u(c_t) dt \right] = \inf_{P \in \mathcal{P}} \left[\int_0^T e^{-\beta t} E^P u(c_t) dt \right],$$

where

$$u(c_t) = (c_t)^\alpha / \alpha, \quad \alpha < 0.$$

Denote by P^* the prior in $\mathcal{P}$ corresponding to $\sigma_t = \overline{\sigma}$ for all t; that is, P^* is the measure on Ω induced by X^*,

$$X_t^* = \overline{\sigma} B_t, \quad \text{for all } t \text{ and } \omega.$$

Then

$$V_0(e) = E^{P^*} \left[\int_0^T e^{-\beta s} u(e_t) dt \right].$$

(This is because $u(e_t) = \alpha^{-1} \exp(\alpha \log e_t)$ and because $\alpha < 0$ makes $x \longmapsto e^{\alpha x}/\alpha$ concave, so that the argument in Example 1 can be adapted.)

Define the nonnegative continuous function φ on $\mathbb{R}$ by

$$\varphi(x) = \begin{cases} 1 & x \le \underline{\sigma}^2 \\ \dfrac{2x}{\underline{\sigma}^2 - \overline{\sigma}^2} - \dfrac{\underline{\sigma}^2 + \overline{\sigma}^2}{\underline{\sigma}^2 - \overline{\sigma}^2} & \underline{\sigma}^2 < x < \dfrac{\underline{\sigma}^2 + \overline{\sigma}^2}{2} \\ 0 & \dfrac{\underline{\sigma}^2 + \overline{\sigma}^2}{2} \le x \end{cases}$$

Fix $\tau > 0$. Define the event N_τ by $N_\tau = \left\{ \omega : \langle B \rangle_\tau = \underline{\sigma}^2 \tau \right\}$, and the consumption process c by

$$c_t = \begin{cases} e_t & 0 \le t \le \tau \\ e_t + \varphi(\langle B \rangle_\tau / \tau) & \tau \le t \le T \end{cases}$$

Then $V_\tau(c) \ge V_\tau(e)$ $\mathcal{P}$-almost surely and a strict preference prevails on N_τ because $\varphi(\underline{\sigma}^2) = 1$ and $P^{\underline{\sigma}}(N_\tau) = 1$. However, c is indifferent to e at time 0 because $\varphi(\langle B \rangle_\tau / \tau) = \varphi(\overline{\sigma}^2) = 0$ under P^*.

In the asset pricing application below we focus on dynamic behavior (and equilibria) where ex ante optimal plans are implemented for all time $\mathcal{P}$-almost surely. This requires that we examine behavior from conditional perspectives and not only ex ante. Accordingly, if the feasible set in the above example is $\{e, c\}$, the predicted choice would be c.

3. Asset Returns

This section describes some implications of ambiguous volatility for asset pricing theory. First, we describe what can be said about prices based on hedging arguments, without assuming preference maximization or equilibrium. It is well known that ambiguous volatility leads to market incompleteness (see Avellaneda, Levy, and Paras 1995, for example) and hence that perfect hedging is generally impossible. Accordingly, hedging arguments lead only to interval predictions of security prices, which adds to the motivation for considering preferences and equilibrium. We explore such an equilibrium approach by employing the utility functions defined above and a representative agent setup. The main result is a version of the C-CAPM that applies when volatility is ambiguous. As an illustration of the added explanatory power of the model, it can rationalize the well-documented feature of option prices whereby the Black-Scholes implied volatility exceeds the realized volatility of the underlying security.

In received theory, there exist positive "state-prices" that characterize arbitrage-free and equilibrium prices in the familiar way.[19] In the standard setup where null events are defined by a reference (physical or subjective) measure, state prices are often used as densities to define a risk-neutral or martingale measure. Densities do not apply when priors disagree about what is possible. But, surprisingly perhaps, suitable state prices can still be derived.

There is a literature on the pricing of derivative securities when volatility is ambiguous. The problem was first studied by Lyons (1995) and Avellaneda, Levy, and Paras (1995); recent explorations include Denis and Martini (2006), Cont (2006), and Vorbrink (2010).[20] They employ hedging arguments to derive upper and lower bounds on security prices. However, our Theorem 1 is the first to characterize these price bounds in terms of state prices. In addition, we study a Lucas-style endowment economy and thus take the endowment as the basic primitive, while the cited papers take the prices of primitive securities as given. We are not aware of any previous studies of equilibrium in continuous time with ambiguous volatility.

Our asset market analysis is conducted under the assumption that there is ambiguity only about volatility and that the volatility matrix σ_t is restricted by:

$$\sigma_t(\omega) \in \Gamma, \quad \text{for each } t \text{ and } \omega, \tag{24}$$

where $\Gamma \subset \mathbb{R}^{d\times d}$ is compact and convex and, for all σ in Γ, $\sigma\sigma^\top \geq \hat{a}$ for some positive definite matrix $\widehat{a}$. (In the one-dimensional case, $\Gamma = [\underline{\sigma}, \overline{\sigma}]$ with $\underline{\sigma} > 0$; Examples 1 and 2 both use this specification.) The trinomial example in

[19] We do not treat arbitrage formally. However, at an informal level we identify no-arbitrage prices with those produced by Black-Scholes-style hedging arguments because of their intuitive connection and because of the formal connection that is familiar in the standard ambiguity-free model (Duffie 1996).

[20] These papers often refer to uncertain volatility rather than to ambiguous volatility.

Section 1.2 is the discrete time one-dimensional counterpart. The formal model is due to Peng (2006), who calls it *G-Brownian motion*, which is, loosely put, "Brownian motion with ambiGuous volatility." We adopt this terminology and thus refer to the coordinate process B as being a G-Brownian motion under $\mathcal{P}$. Importantly, much of the machinery of stochastic calculus, including generalizations of Itô's lemma and Itô integration, has been extended to the framework of G-Brownian motion (see Appendix A for brief descriptions). This machinery is used in the proofs. However, the proof "ideas" are standard; for example, they exploit Itô's lemma and a martingale representation theorem. The difficulty is only to know when and in precisely what form such tools apply. The statements of results do not rely on this formal material and are easy to understand if one accepts that they differ from standard theory primarily through the use of a new (nonadditive) notion of conditional expectation and the substitution of "$\mathcal{P}$-almost surely" for the usual almost surely qualification.

In order to state the asset pricing results precisely, we need to be more precise about the formal domains for random variables and stochastic processes. They differ from the usual domains (only) because of the nonequivalence of priors. Random payoffs occurring at a single instant are taken to be bounded continuous functions of the state or suitable limits of such functions. Formally, define the space $\widehat{L^2}(\Omega)$ to be the completion, under the norm $\| \xi \| \equiv (\hat{E}[| \xi |^2])^{\frac{1}{2}}$, of the set of all bounded continuous functions on Ω. Then $\widehat{L^2}(\Omega)$ is a subset of the set of measurable random variables ξ for which $\sup_{P \in \mathcal{P}} E^P\left(| \xi |^2\right) < \infty$.[21] For processes, define $M^{2,0}$ to be the class of processes η of the form

$$\eta_t(\omega) = \sum_{i=0}^{N-1} \xi_i(\omega) 1_{[t_i, t_{i+1})}(t),$$

where $\xi_i \in \widehat{L^2}(\Omega)$, $0 \leq i \leq N-1$, and $0 = t_0 < \cdots < t_N = T$. Roughly, each such η is a step function in random variables from $\widehat{L^2}(\Omega)$. For the usual technical reasons, we consider also suitable limits of such processes. Thus, the ambient space for processes, denoted M^2, is taken to be the completion of $M^{2,0}$ under the norm

$$\| \eta \|_{M^2} \equiv \left(\hat{E}\left[\int_0^T | \eta_t |^2 dt \right] \right)^{\frac{1}{2}}.$$

If, for every t, $Z_t = 0$ $\mathcal{P}$-$a.s.$, then $Z = 0$ in M^2 (because $\hat{E}[\int_0^T | Z_t |^2 dt] \leq \int_0^T \hat{E}[| Z_t |^2] dt = 0$), but the converse is not valid in general. The consumption processes (c_t) and utility processes $(V_t(c))$ discussed above, as well as all processes below related to asset markets, are taken to lie in M^2.

The domain M^2 depends on the set of priors $\mathcal{P}$, and hence also on set Γ from (24) that describes volatility ambiguity. If ambiguity about volatility increases

21 It contains many discontinuous random variables. For example, $\widehat{L^2}(\Omega)$ contains every bounded and lower semicontinuous function on Ω (see our companion paper).

in the sense that Γ is replaced by Γ_*, $\Gamma \subset \Gamma_*$, then it is easy to see that the corresponding domain of processes shrinks—that is,

$$\Gamma \subset \Gamma_* \Longrightarrow M_*^2 \subset M^2. \tag{25}$$

The reason is that when ambiguity increases, processes are required to be well behaved (square integrable, for example) with respect to more probability laws.

3.1 Hedging and state prices

Consider the following market environment. There is a single consumption good, a riskless asset with return r_t and d risky securities available in zero net supply. Returns R_t to the risky securities are given by

$$dR_t = b_t dt + s_t dB_t, \tag{26}$$

where s_t is a $d \times d$ invertible volatility matrix. Both (b_t) and (s_t) are known by the investor.[22] Define $\eta_t = s_t^{-1}(b_t - r_t 1)$, the market price of uncertainty (a more appropriate term here than market price of risk). It is assumed henceforth that (r_t) is a bounded process in M^2; a restriction on the market price of risk will be given below.

It is important to understand the significance of the assumption that the returns equation holds $\mathcal{P}$-almost surely, that is, P-$a.s.$ for every prior in $\mathcal{P}$. Because we are excluding ambiguity about drift, each prior P corresponds to an admissible hypothesis (σ_t) for volatility via (15) and (16). Thus, write $P = P^{(\sigma_t)}$. Then, taking $d=1$ for simplicity,

$$\langle R \rangle_t = \int_0^t s_\tau^2 d\langle B \rangle_\tau = \int_0^t s_\tau^2 \sigma_\tau^2 d\tau \quad P^{(\sigma_t)}\text{-}a.s.$$

In general, the prior implied by an alternative hypothesis $\left(\sigma_t'\right)$ is not equivalent to $P^{(\sigma_t)}$, which means that $P^{\left(\sigma_t'\right)}$ and $P^{(\sigma_t)}$ yield different views of the quadratic variation of returns. Consequently, the volatility of returns is ambiguous: it is certain only that $\langle R \rangle_t$ lies in the interval $\left[\underline{\sigma}^2 \int_0^t s_\tau^2 d\tau, \overline{\sigma}^2 \int_0^t s_\tau^2 d\tau\right]$.

Similarly, unless explicitly stated otherwise, the $\mathcal{P}$-almost sure qualification should be understood to apply to all other equations (and inequalities) below even where not stated, and its significance can be understood along the same lines.

Fix the dividend stream denoted (δ, δ_T), where δ_t is the dividend for $0 \leq t < T$ and δ_T is the lumpy dividend paid at the terminal time; formally, $(\delta, \delta_T) \in M^2 \times \widehat{L^2}(\Omega)$. For a given time τ, consider the following law of motion for wealth on $[\tau, T]$:

$$dY_t = \left(r_t Y_t + \eta_t^\top \phi_t - \delta_t\right) dt + \phi_t^\top dB_t, \tag{27}$$

$$Y_\tau = y,$$

[22] b_t and s_t are functions on Ω, the set of possible trajectories for the driving process. It is these *functions* that are known. The trajectory is, of course, uncertain and known only at T.

where y is initial wealth, $\phi_t = Y_t s_t^\top \psi_t$, and (ψ_t) is the trading strategy—that is, ψ_{ti} is the proportion of wealth invested in risky security i. (By nonsingularity of s_t, choice of a trading strategy can be expressed equivalently in terms of choice of (ϕ_t), which reformulation is simplifying.) Denote the unique solution by $Y^{y,\phi,\tau}$. Define the superhedging set

$$\mathcal{U}_\tau = \left\{ y \geq 0 \,\middle|\, \exists \phi \in M^2 \text{ s.t. } Y_T^{y,\phi,\tau} \geq \delta_T \right\},$$

and the *superhedging price* $\overline{S}_\tau = \inf\{y \mid y \in \mathcal{U}_\tau\}$. Similarly define the subhedging set

$$\mathcal{L}_\tau = \left\{ y \geq 0 \,\middle|\, \exists \phi \in M^2 \text{ s.t. } Y_T^{-y,\phi,\tau} \geq -\delta_T \right\}$$

and the *subhedging price* $\underline{S}_\tau = \sup\{y \mid y \in \mathcal{L}_\tau\}$.

The relevance of ambiguity about volatility is apparent once one realizes that the law of motion (27), and also the inequalities at T that define $\mathcal{U}_\tau$ and $\mathcal{L}_\tau$, should be understood to hold $\mathcal{P}$-almost surely. Thus, for example, a superhedging trading strategy must deliver δ_t on $[0,T)$ and also at least δ_T at T for all realizations that are conceivable according to some prior. Speaking loosely, the need to satisfy many nonequivalent priors in this way makes superhedging difficult ($\mathcal{U}_\tau$ small) and the superhedging price large. Similarly, ambiguity reduces the subhedging price. Hence the price interval is made larger by ambiguous volatility. More precisely, by the preceding argument, (25), and using the obvious notation,

$$\Gamma \subset \Gamma_* \Longrightarrow \left[\underline{S}_0, \overline{S}_0\right] \subset \left[\underline{S}_{*0}, \overline{S}_{*0}\right].$$

Thus an increase in volatility ambiguity weakens the implications for price of a hedging argument and naturally bolsters the case for pursuing an equilibrium analysis, which we do after characterizing super and subhedging prices.

We show that both of the above prices can be characterized using appropriately defined state prices. Let

$$v_t = \overline{\lim_{\varepsilon \downarrow 0}} \frac{1}{\varepsilon}(\langle B \rangle_t - \langle B \rangle_{t-\varepsilon}), \tag{28}$$

where $\overline{\lim}$ is taken componentwise.[23] Under P_0, B is a Brownian motion and v_t equals the $d \times d$ identity matrix P_0-$a.s.$ Importantly, we can also describe v_t as seen through the lens of any other prior in $\mathcal{P}$: if $P = P^{(\sigma_t)}$ is a prior in $\mathcal{P}$ corresponding via the SDE (15) to (σ_t), then[24]

$$v_t = \sigma_t \sigma_t^\top \ \ dt \times P^{(\sigma_t)}\text{-}a.s. \tag{29}$$

[23] In this we are following Soner, Touzi, and Zhang (2012). The quadratic variation process $\langle B \rangle$ is defined in (3); $v_t(\omega)$ takes values in $\mathbb{S}_d^{>0}$, the space of all $d \times d$ positive-definite matrices.

[24] Here is the proof: By Soner, Touzi, and Zhang (2012, 4), $\langle B \rangle$ equals the quadratic variation of B $P^{(\sigma_t)}$-$a.s.$; and by Oksendal (2005, 56), the quadratic variation of $\int_0^t \sigma_s dB_s$ equals $\int_0^t \sigma_s \sigma_s^\top ds \quad P^{(\sigma_t)}$-$a.s.$ Thus we have the $P^{(\sigma_t)}$-$a.s.$ equality in processes $\langle B \rangle = (\int_0^t \sigma_s \sigma_s^\top ds)$. Because $\langle B \rangle$ is absolutely continuous, its time derivative exists $a.s.$ on $[0,T]$; indeed, the derivative at t is v_t. Evidently, $\frac{d}{dt}\int_0^t \sigma_s \sigma_s^\top ds = \sigma_t \sigma_t^\top$ for almost every t. Equation (29) follows.

It is assumed henceforth that $(v_t^{-1}\eta_t)$ is a bounded process in M^2.[25]

By the *state price process* we mean the unique solution $\pi=(\pi_t)$ to

$$d\pi_t/\pi_t=-r_t dt-\eta_t^\top v_t^{-1}dB_t,\quad \pi_0=1, \tag{30}$$

which admits a closed-form expression paralleling the classical case:[26]

$$\pi_t=\exp\left\{-\int_0^t r_s ds-\int_0^t \eta_s^\top v_s^{-1}dB_s-\frac{1}{2}\int_0^t \eta_s^\top v_s^{-1}\eta_s ds\right\},\quad 0\le t\le T. \tag{31}$$

We emphasize the important fact that π is "universal" in the sense of being defined almost surely for *every* prior in $\mathcal{P}$. More explicitly, π satisfies: for every t,

$$\pi_t=\exp\left\{-\int_0^t r_s ds-\int_0^t \eta_s^\top\left(\sigma_s\sigma_s^\top\right)^{-1}dB_s-\frac{1}{2}\int_0^t \eta_s^\top\left(\sigma_s\sigma_s^\top\right)^{-1}\eta_s ds\right\},\quad P^{(\sigma_t)}\text{-}a.s.$$

Roughly speaking, this defines $\pi_t(\omega)$ for every trajectory ω of the driving process that is possible according to at least one prior in $\mathcal{P}$.[27]

Our characterization of superhedging and subhedging prices requires an additional, arguably minor, restriction on the security market. To express it, for any $\varepsilon>0$, define

$$\widehat{L^{2+\varepsilon}}(\Omega)=\left\{\xi\in\widehat{L^2}(\Omega):\hat{E}\left[|\xi|^{2+\varepsilon}\right]<\infty\right\}.$$

The restriction is that π and δ satisfy

$$\left(\pi_T\delta_T+\int_0^T \pi_t\delta_t dt\right)\in\widehat{L^{2+\varepsilon}}(\Omega). \tag{32}$$

Theorem 1 (Hedging prices). Fix a dividend stream $(\delta,\delta_T)\in M^2\times\widehat{L^2}(\Omega)$. Suppose that r and $(v_t^{-1}\eta_t)$ are bounded processes in M^2 and that (32) is satisfied. Then the superhedging and subhedging prices at any time τ are given by ($\mathcal{P}$-$a.s.$)

$$\overline{S}_\tau=\hat{E}\left[\int_\tau^T \frac{\pi_t}{\pi_\tau}\delta_t dt+\frac{\pi_T}{\pi_\tau}\delta_T \,\middle|\, \mathcal{F}_\tau\right]$$

and

$$\underline{S}_\tau=-\hat{E}\left[-\int_\tau^T \frac{\pi_t}{\pi_\tau}\delta_t dt-\frac{\pi_T}{\pi_\tau}\delta_T \,\middle|\, \mathcal{F}_\tau\right].$$

25 This restriction will be confirmed below whenever η is taken to be endogenous.

26 Apply Remark 1.3 in Peng (2010, Ch. 5).

27 Still speaking roughly, π_t is defined on the union of the supports of all priors in $\mathcal{P}$. Because these supports need not be pairwise disjoint, it is not obvious that such a "universal" definition exists. But (28) and (29) ensure that π_t is well defined even where supports overlap.

In the special case where the security δ can be perfectly hedged (that is, there exist y and ϕ such that $Y_T^{y,\phi,0} = \delta_T$), then

$$\overline{S}_0 = \underline{S}_0 = \hat{E}\left[\pi_T\delta_T + \int_0^T \pi_t\delta_t dt\right] = -\hat{E}\left[-\pi_T\delta_T - \int_0^T \pi_t\delta_t dt\right].$$

Because this asserts equality of the supremum and infimum of expected values of $\pi_T\delta_T + \int_0^T \pi_t\delta_t dt$ as the measures P vary over $\mathcal{P}$, it follows that $E^P[\pi_T\delta_T + \int_0^T \pi_t\delta_t dt]$ is constant for all such measures P—that is, the hedging price is unambiguous. In the further specialization where there is no ambiguity and P_0 is the single prior, one obtains pricing by an equivalent martingale measure whose density (on $\mathcal{F}_t$) with respect to P_0 is π_t.

Remark 3. Vorbrink (2010) obtains an analogous characterization of hedging prices under the assumption of G-Brownian motion. However, in place of our assumption (32), he adopts the strong assumption that $b_t = r_t$, so that the market price of uncertainty η_t vanishes and $\pi_t = \exp\{-\int_0^t r_s ds\}$.

Example 3 (Closed form hedging prices). We derive the super and subhedging prices of a European call option in a special case and compare the results with the standard Black-Scholes formula.

Let there be one risky security ($d=1$) with price (S_t) satisfying

$$dS_t/S_t = dR_t = b_t dt + s_t dB_t.$$

Suppose further that $r_t \equiv r$, $s_t \equiv 1$ and $b_t - r = bv_t$, where $b > 0$ and r are constants. Thus the market price of uncertainty is given by $\eta_t = bv_t$; that is, using (29),

$$\eta_t = b\sigma_t^2 \quad dt \times P^{(\sigma_t)}\text{-}a.s.$$

It follows that state prices are given, $\mathcal{P}$-almost surely, by

$$\pi_t = \exp\left\{-rt - bB_t - \frac{1}{2}b^2\langle B\rangle_t\right\}, \quad 0 \le t \le T.$$

Consider a European call option on the risky security that matures at date T and has exercise price K. The super and subhedging prices at t can be written in the form $\overline{c}(S_t, t)$ and $\underline{c}(S_t, t)$, respectively. At the maturity date,

$$\overline{c}(S_T, T) = \underline{c}(S_T, T) = \max[0, S_T - K] \equiv \Phi(S_T).$$

By Theorem 1,

$$\overline{c}(S_t, t) = \hat{E}\left[\frac{\pi_T}{\pi_t}\Phi(S_T)\middle|\mathcal{F}_t\right]$$

and

$$\underline{c}(S_t, t) = -\hat{E}\left[-\frac{\pi_T}{\pi_t}\Phi(S_T)\middle|\mathcal{F}_t\right].$$

By the nonlinear Feynman-Kac formula in Peng (2010), we obtain the following Black-Scholes-Barenblatt equations:[28]

$$\partial_t \overline{c} + \sup_{\underline{\sigma} \leq \sigma \leq \overline{\sigma}} \left\{ \frac{1}{2} \sigma^2 S^2 \partial_{SS} \overline{c} \right\} + r S \partial_S \overline{c} - r \overline{c} = 0, \; \overline{c}(S,T) = \Phi(S)$$

and

$$\partial_t \underline{c} - \sup_{\underline{\sigma} \leq \sigma \leq \overline{\sigma}} \left\{ -\frac{1}{2} \sigma^2 S^2 \partial_{SS} \underline{c} \right\} + r S \partial_S \underline{c} - r \underline{c} = 0, \; \underline{c}(S,T) = \Phi(S).$$

Because $\Phi(\cdot)$ is convex, so is $\overline{c}(\cdot, t)$.[29] It follows that the respective suprema in the above equations are achieved at $\overline{\sigma}$ and $\underline{\sigma}$, and we obtain

$$\partial_t \overline{c} + \frac{1}{2} \overline{\sigma}^2 S^2 \partial_{SS} \overline{c} + r S \partial_S \overline{c} - r \overline{c} = 0, \; \overline{c}(S,T) = \Phi(S)$$

and

$$\partial_t \underline{c} + \frac{1}{2} \underline{\sigma}^2 S^2 \partial_{SS} \underline{c} + r S \partial_S \underline{c} - r \underline{c} = 0, \; \underline{c}(S,T) = \Phi(S).$$

Therefore,

$$\overline{c}(S,t) = E^{P^{\overline{\sigma}}} \left[\frac{\pi_T}{\pi_t} \Phi(S_T) \middle| \mathcal{F}_t \right]$$

and

$$\underline{c}(S,t) = E^{P^{\underline{\sigma}}} \left[\frac{\pi_T}{\pi_t} \Phi(S_T) \middle| \mathcal{F}_t \right].$$

In other words, the super and subhedging prices are the Black-Scholes prices with volatilities $\overline{\sigma}$ and $\underline{\sigma}$, respectively.

It is noteworthy that in contrast to this effect of volatility ambiguity, the arbitrage-free price of a European call option is unaffected by ambiguity about drift. This might be expected because the Black-Scholes price does not depend on the drift of the underlying. Nevertheless, some supporting detail may be useful.

Let the security price be given as above by

$$dS_t / S_t = dR_t = (b + r)dt + dB_t, \quad P_0\text{-}a.s.$$

Model drift ambiguity by a set of priors as described in Section 2.1. Each alternative hypothesis $\theta = ((\mu_t), 1)$ about the drift generates a prior P^θ

[28] They reduce to the standard Black-Scholes equation if $\underline{\sigma} = \overline{\sigma}$.

[29] The argument is analogous to that in the classical Black-Scholes analysis.

constructed as in (15) and (16). By the Girsanov theorem, $X_t^\theta = \int_0^\tau (\mu_t dt + dB_t)$ is a standard Brownian motion under probability P^θ. Therefore,

$$dS_t/S_t = (b+r)dt + dX_t^\theta, \quad P^\theta\text{-}a.s.$$

and the security price follows the identical geometric Brownian motion under P^θ. Similarly, the counterpart of the wealth accumulation equation (27) gives

$$dY_t = (rY_t + \eta\phi_t)dt + \phi_t dX_t^\theta, \quad P^\theta\text{-}a.s.$$

where $\eta = b - r$. Together, the latter two equations imply that the identical Black-Scholes price would prevail for the option regardless if P_0 or P^θ were the true probability law. Because θ is an arbitrary hypothesis about drift, it can be shown that the option price is unaffected by ambiguity about drift.

In fact, the irrelevance of ambiguity about drift is valid much more generally. Let there be d risky securities whose prices $S_t \in \mathbb{R}^d$ solve a stochastic differential equation of the form

$$dS_t = \widehat{b_t}(S_t)dt + \widehat{s_t}(S_t)dB_t,$$

where $\widehat{b_t}$ and $\widehat{s_t}$ are given $\mathbb{R}^d$-valued suitably well-behaved functions (for example, each $\widehat{s_t}(\cdot)$ is everywhere invertible). Let the instantaneous return to the riskless security be $r_t(S_t)$. Finally, let the continuous function $\psi : \mathbb{R}^d \to \mathbb{R}$ determine the payoff $\psi(S_T)$ at time T of a derivative security. Under P_0, when B is standard Brownian motion, the arbitrage-free price of the derivative is defined by the Black-Scholes PDE (see Chapter 5 in Duffie 1996, for example). The fact that the drift $\widehat{b_t}$ does not enter into the PDE suggests that ambiguity about drift does not affect the price of the derivative. Further intuition follows as above from the Girsanov theorem.[30] This argument covers all the usual European options.[31] In contrast, as illustrated by the example of a European call option, ambiguity about volatility does matter (the price interval in Theorem 1 is typically nondegenerate). This is not surprising given the known importance of volatility in option pricing. More formally, the difference from the case of drift arises because for an alternative hypothesis $\theta = (0, (\sigma_t))$ for volatility, $\left(X_t^\theta\right)$ satisfying $dX_t^\theta = \sigma_t dB_t$ is *not* a standard Brownian motion under P^θ.

---

30 For any alternative hypothesis $\theta = ((\mu_t), 1)$ and corresponding prior P^θ, the security price process satisfies

$$dS_t = \widehat{b_t}(S_t)dt + \widehat{s_t}(S_t)dX_t^\theta, \; P^\theta\text{-}a.s.,$$

where $X_t^\theta = \int_0^\tau (\mu_t dt + dB_t)$ is a standard Brownian motion under P^θ. Therefore, all conceivable truths P^θ imply the identical price for the derivative, and ambiguity about drift has no effect. Further, the corresponding hedging strategy is also unaffected. A rigorous proof is readily constructed. We do not provide it because ambiguity about drift alone is not our focus.

31 It is not difficult to show by a similar argument that the arbitrage-free price of Asian options is also unaffected by ambiguity about drift.

3.2 Equilibrium

Here we use state prices to study equilibrium in a representative agent economy with sequential security markets.

In the sequel, we limit ourselves to scalar consumption at every instant so that $C \subset \mathbb{R}_+$. At the same time we generalize utility to permit lumpy consumption at the terminal time. Thus we use the utility functions V_t given by

$$V_t(c,\xi) = -\hat{E}\left[-\int_t^T f(c_s, V_s(c,\xi))ds - u(\xi)\,\middle|\,\mathcal{F}_t\right], \quad 0 \le t \le T,$$

where (c,ξ) varies over a subset D of $M^2 \times \widehat{L^2}(\Omega)$. Here c denotes the absolutely continuous component of the consumption process and ξ is the lump of consumption at T. Our analysis of utility extends to this larger domain in a straightforward way.

When considering (c,ξ), it is without loss of generality to restrict attention to versions of c for which $c_T = \xi$. With this normalization, we can abbreviate $(c,\xi) = (c, c_T)$ by c and identify c with an element of $M^2 \times \widehat{L^2}(\Omega)$. Accordingly, write $V_t(c,\xi)$ more simply as $V_t(c)$, where[32]

$$V_t(c) = -\hat{E}\left[-\int_t^T f(c_s, V_s(c))ds - u(c_T)\,\middle|\,\mathcal{F}_t\right], \quad 0 \le t \le T. \tag{33}$$

The agent's endowment is given by the process e. Define $\pi^e = \left(\pi_t^e\right)$, called a *supergradient at e*, by

$$\pi_t^e = \exp\left(\int_0^t f_v(e_s, V_s(e))ds\right) f_c(e_t, V_t(e)), \quad 0 \le t < T, \tag{34}$$

$$\pi_T^e = \exp\left(\int_0^T f_v(e_s, V_s(e))ds\right) u_c(e_T).$$

Securities, given by (26) and available in zero net supply, are traded in order to finance deviations from the endowment process e. Denote the trading strategy by (ψ_t), where ψ_{ti} is the proportion of wealth invested in risky security i. Then wealth Y_t evolves according to the equation

$$dY_t = \left(r_t Y_t + \eta_t^\top \phi_t - (c_t - e_t)\right)dt + \phi_t^\top dB_t, \tag{35}$$

$$Y_0 = 0, \quad c_T = e_T + Y_T \ge 0,$$

where $\eta_t = s_t^{-1}(b_t - r_t 1)$ and $\phi_t = Y_t s_t^\top \psi_t$. We remind the reader that, by nonsingularity of s_t, choice of a trading strategy can be expressed equivalently in terms of choice of (ϕ_t).

32 We assume the following conditions for f and u. (1) f and u are continuously differentiable and concave. (2) There exists $\kappa > 0$ such that $|u_c(c)| < \kappa(1+c)$ for all $c \in C$, and $\sup\{|f_c(c,V)|, |f(c,0)|\} < \kappa(1+c)$ for all $(c,V) \in C \times \mathbb{R}$. A consequence is that if $c \in M^2 \times \widehat{L^2}(\Omega)$, then $u(c_T), u_c(c_T) \in \widehat{L^2}(\Omega)$ and $f(c_t, 0),\ f_c(c_t, V_t(c)) \in M^2$.

Refer to c as being *feasible* if $c \in D$ and there exists ϕ in M^2 such that (35) is satisfied. More generally, for any $0 \le \tau \le T$, consider an individual with initial wealth Y_τ who trades securities and consumes during the period $[\tau, T]$. Say that *c is feasible on $[\tau, T]$ given initial wealth Y_τ* if (35) is satisfied on $[\tau, T]$ and wealth at τ is Y_τ. Because (35) should be understood as being satisfied P-almost surely for every prior in $\mathcal{P}$, greater ambiguity about volatility tightens the feasibility restriction (paralleling the discussion in the previous section).

State prices can be used to characterize feasible consumption plans as described next.

Theorem 2 (State prices). Define $\pi \in M^2$ by (31) and let $0 \le \tau < T$.

(i) If c is feasible on $[\tau, T]$ given initial wealth Y_τ, then, $\mathcal{P}$-$a.s.$,

$$Y_\tau = \hat{E}\left[\int_\tau^T \frac{\pi_t}{\pi_\tau}(c_t - e_t)dt + \frac{\pi_T}{\pi_\tau}(c_T - e_T)\big|\mathcal{F}_\tau\right] \tag{36}$$

$$= -\hat{E}\left[-\int_\tau^T \frac{\pi_t}{\pi_\tau}(c_t - e_t)dt - \frac{\pi_T}{\pi_\tau}(c_T - e_T)\big|\mathcal{F}_\tau\right].$$

(ii) Conversely, suppose that (36) is satisfied and that $c_T \ge 0$. Then c is feasible on $[\tau, T]$ given initial wealth Y_τ.

When relevant processes are ambiguity-free diffusions, Cox and Huang (1989) show that state prices can be used to transform a dynamic process of budget constraints into a single static budget constraint. The theorem provides a counterpart for our setting: for any given τ and initial wealth Y_τ, feasibility on $[\tau, T]$ may be described by the "expected" expenditure constraint (36). Because both (35) and (36) must be understood to hold $\mathcal{P}$-almost surely, speaking loosely, the equivalence between the dynamic and static budget constraints is satisfied simultaneously for all hypotheses (σ_t) satisfying (24).

Perhaps surprisingly, (36) contains two (generalized) expectations. Their conjunction can be interpreted as in the discussion following Theorem 1: in expected value terms any feasible consumption plan *unambiguously* (that is, for every prior) exhausts initial wealth when consumption is priced using π. Such an interpretation is evident when $\tau = 0$; a similar interpretation can be justified when $\tau > 0$.

Turn to equilibrium. Say that $(e, (r_t, \eta_t))$ is a *sequential equilibrium* if for every c: For each τ, $\mathcal{P}$-almost surely,

$$c \in \Upsilon_\tau(0) \Longrightarrow V_\tau(c) \le V_\tau(e).$$

Thus equilibrium requires not only that the endowment e be optimal at time 0, but also that it remain optimal at any later time given that e has been followed to that point.

The main result of this section follows.

Theorem 3 (Sequential Equilibrium I). Define $\pi, \pi^e \in M^2$ by (31) and (34) respectively, and assume that $\mathcal{P}$-almost surely,

$$\pi_t^e / \pi_0^e = \pi_t. \tag{37}$$

Then $(e, (r_t, \eta_t))$ is a sequential equilibrium.

Condition (37) is in the spirit of the Duffie and Skiadas (1994) approach to equilibrium analysis (see also Skiadas 2008 for a comprehensive overview of this approach). Speaking informally, the process π^e/π_0^e describes marginal rates of substitution at e, while π describes tradeoffs offered by the market. Their equality relates the riskless rate and the market price of risk to consumption and continuation utility through the equation

$$d\pi_t^e / \pi_t^e = -r_t dt - \eta_t^\top v_t^{-1} dB_t.$$

To be more explicit, suppose that e satisfies

$$de_t / e_t = \mu_t^e dt + \left(s_t^e\right)^\top dB_t.$$

Consider also two specific aggregators. For the standard aggregator (7),

$$\pi_t^e = \exp(-\beta t) u_c(e_t), \quad 0 \leq t \leq T. \tag{38}$$

Then Itô's lemma for G-Brownian motion (Appendix A) and (29) imply that

$$b_t - r_t \mathbf{1} = s_t \eta_t = -\left(\frac{e_t u_{cc}(e_t)}{u_c(e_t)}\right) s_t \sigma_t \sigma_t^\top s_t^e, \quad P^{(\sigma_t)}\text{-}a.s. \tag{39}$$

which is a version of the C-CAPM for our setting.[33] From the perspective of the measure P_0 according to which the coordinate process B is a Brownian motion, $\sigma_t \sigma_t^\top$ is the identity matrix and one obtains the usual C-CAPM. However, speaking informally, the equation (39) relates excess returns to consumption also along trajectories that are consistent with alternative hypotheses (σ_t) about the nature of the driving process.

The other case is the so-called Kreps-Porteus aggregator (Duffie and Epstein 1992b). Let

$$f(c, v) = \frac{c^\rho - \beta(\alpha v)^{\rho/\alpha}}{\rho(\alpha v)^{(\rho-\alpha)/\alpha}}, \tag{40}$$

where $\beta \geq 0$ and $0 \neq \rho, \alpha \leq 1$; $(1-\alpha)$ is the measure of relative risk aversion and $(1-\rho)^{-1}$ is the elasticity of intertemporal substitution. To evaluate terminal lumpy consumption as in (33), we take

$$u(c_t) = (c_t)^\alpha / \alpha, \quad 0 \neq \alpha < 1.$$

[33] Equality here (and in similar equations below) means that the two processes $(b_t - r_t 1)$ and $(-\left(\frac{e_t u_{cc}(e_t)}{u_c(e_t)}\right) s_t \sigma_t \sigma_t^\top s_t^e)$ are equal as processes in M^2. Notice also that $v_t^{-1} \eta_t = -\left(\frac{e_t u_{cc}(e_t)}{u_c(e_t)}\right) s_t^e$, yielding a process in M^2 and thus confirming our prior assumption on security markets.

The implied version of the C-CAPM is

$$b_t - r_t 1 = \rho^{-1}\left[\alpha(1-\rho)s_t\sigma_t\sigma_t^{\top}s_t^e + (\rho-\alpha)s_t\sigma_t\sigma_t^{\top}s^M\right], \quad P^{(\sigma_t)}\text{-}a.s. \tag{41}$$

where s^M is the volatility of wealth in the sense that

$$dY_t/Y_t = b^M dt + \left(s^M\right)^{\top} dB_t, \quad \mathcal{P}\text{-}a.s.$$

In the absence of ambiguity where P_0 alone represents beliefs, then (41) reduces to the two-factor model of excess returns derived by Duffie and Epstein (1992a).

Equation (41) is derived in Appendix B.4, which also presents a result for general aggregators.

3.3 Minimizing priors

Equation (37) is a sufficient condition for sequential equilibrium. Here we describe an alternative route to equilibrium that is applicable under an added assumption and that yields an alternative form of C-CAPM.

The intuition for what follows is based on a well-known consequence of the minimax theorem for multiple priors utility in abstract environments. For suitable optimization problems, if a prospect, say e, is feasible and if the set of priors contains a worst-case scenario P^* for e, then e is optimal if and only if it is optimal also for a Bayesian agent that uses the single prior P^*. Moreover, by a form of envelope theorem, P^* suffices to describe marginal rates of substitution at e and hence also supporting shadow prices. This suggests that there exist sufficient conditions for e to be part of an equilibrium in our setup that refer to P^* and less extensively to all other priors in $\mathcal{P}$. We proceed now to explore this direction.

As a first step, define $P^* \in \mathcal{P}$ to be a *minimizing measure for e* if

$$V_0(e) = E^{P^*}\left[\int_0^T f(e_s, V_s(e))ds + u(e_T)\right]. \tag{42}$$

As discussed when defining equilibrium, the fact that only weak dynamic consistency is satisfied requires that one take into account also conditional perspectives. Speaking informally, a minimizing measure P^* as above need not be minimizing conditionally at a later time because of the nonequivalence of priors and the uncertainty about what is possible. (Example 2 is readily adapted to illustrate this.) Thus, to be relevant to equilibrium, a stronger notion of "minimizing" is required.

Recall that for any prior P in $\mathcal{P}$, P_τ^ω is the version of the regular conditional of P; importantly, it is well defined for every (τ,ω). Say that $P^* \in \mathcal{P}$ is a *dynamically minimizing measure for e* if, for all τ, $\mathcal{P}$-$a.s.$,

$$V_\tau(e) = E^{(P^*)_\tau^\omega}\left[\int_\tau^T f(e_s, V_s(e))ds + u(e_T)\right]. \tag{43}$$

Next relax the equality (37) and assume instead: For every τ and $\mathcal{P}$-almost surely in ω,

$$\pi_t^e / f_c(e_0, V_0(e)) = \pi_t \quad \text{on} \quad [\tau, T] \ (P^*)_\tau^\omega\text{-}a.s. \tag{44}$$

Note that equality is assumed not only ex ante P^*-$a.s.$ but also conditionally, even conditioning on events that are P^*-null but that are possible according to other priors in $\mathcal{P}$.

Theorem 4 (Sequential equilibrium II). Let P^* be a dynamic minimizer for e and assume (44). Then $(e, (r_t, \eta_t))$ is a sequential equilibrium.

The counterparts of the C-CAPM relations (39) and (41) are:[34] For every τ, $\mathcal{P}$-$a.s.$,

$$b_t - r_t \mathbf{1} = -\left(\frac{e_t u_{cc}(e_t)}{u_c(e_t)}\right) s_t \left(\sigma_t^* \sigma_t^{*\top}\right) s_t^e \quad \text{on} \quad [\tau, T] \ (P^*)_\tau^\omega\text{-}a.s. \tag{45}$$

and

$$b_t - r_t 1 = \rho^{-1}\left[\alpha(1-\rho) s_t \left(\sigma_t^* \sigma_t^{*\top}\right) s_t^e + (\rho - \alpha) s_t \left(\sigma_t^* \sigma_t^{*\top}\right) s^M\right] \quad \text{on} \quad [\tau, T] \ (P^*)_\tau^\omega\text{-}a.s. \tag{46}$$

Here $P^* = P^{(\sigma_t^*)}$ is induced by the process $\left(\sigma_t^*\right)$ as in (15).

There are several "nonstandard" features of these relations that we interpret in the following example where the equations take on a more concrete form. However, it may be useful to consider the general forms briefly. For simplicity, consider (45) corresponding to the standard aggregator. One effect of ambiguous volatility is that the relevant instantaneous covariance between asset returns and consumption is modified from $s_t s_t^e$ to $s_t(\sigma_t^* \sigma_t^{*\top}) s_t^e$, where $\left(\sigma_t^*\right)$ is the worst-case hypothesis for volatility. This adjustment reflects a conservative attitude and confidence only that volatility (σ_t) lies everywhere in Γ rather than in any single hypothesis, such as $\sigma_t \equiv 1$, satisfying this constraint.

Compare (45) also with the C-CAPM relation derived assuming ambiguity about drift only. In that case, Chen and Epstein (2002) show that, instead of (45), mean excess returns satisfy

$$b_t - r_t \mathbf{1} = -\left(\frac{e_t u_{cc}(e_t)}{u_c(e_t)}\right) s_t s_t^e + s_t \mu_t^* \quad P_0\text{-}a.s. \tag{47}$$

where $\left(\mu_t^*\right)$ is the worst-case hypothesis for drift.[35] It is difficult to compare these two alternative adjustments for ambiguity in general qualitative terms. Presumably, each kind of ambiguity matters in some contexts (though recall that

[34] Assume that there exists a dynamic minimizer P^* for e. Then one can show that (39) implies (45) and (41) implies (46).

[35] More precisely, in the notation of Section 2.2, $P^{((\mu_t^*),1)}$ is a minimizer in the utility calculation $V_0(e) = \inf_{P \in \mathcal{P}} E^P[\int_0^T u(e_s) e^{-\beta s} ds]$.

drift ambiguity has no effect in European options markets). Because both kinds of ambiguity may matter simultaneously, one obviously would like to establish a version of C-CAPM that accommodates both. However, that would require extensions of the machinery described in Appendix A that to our knowledge is currently available only for environments described by G-Brownian motion.

Further interpretation and comparisons are discussed in the context of a final example.

3.4 A final example

Theorem 3 begs the question whether or when dynamic minimizers exist. We have no general answers at this point, but they exist in the following example. Its simplicity also helps to illustrate the effects of ambiguous volatility on asset returns.

We build on previous examples. Let $d \geq 1$. The endowment process e satisfies (under P_0)

$$d\log e_t = (s^e)^\top \sigma_t dB_t, \quad e_0 > 0 \text{ given}, \tag{48}$$

where s^e is constant and B is a G-Brownian motion (thus the volatility matrix σ_t is restricted only to lie in Γ). We assume that P_0 lies in $\mathcal{P}$, that is, Γ admits the constant $d \times d$ identity matrix. Utility is defined, for any consumption process c, by the following special case of (33):

$$V_t(c) = -\hat{E}\left[-\int_t^T u(c_s)e^{-\beta(s-t)}ds - e^{-\beta(T-t)}u(c_T) \,\middle|\, \mathcal{F}_t\right],$$

where the felicity function u is

$$u(c_t) = (c_t)^\alpha/\alpha, \quad 0 \neq \alpha < 1.$$

There exists a dynamic minimizer for e that depends on the sign of α. Compute that

$$u(e_t) = \alpha^{-1}e_t^\alpha = \alpha^{-1}e_0^\alpha \exp\left\{\alpha \int_0^t (s^e)^\top \sigma_s dB_s\right\}$$

Let $\underline{\sigma}$ and $\overline{\sigma}$ solve, respectively,

$$\min_{\sigma\in\Gamma} tr\left(\sigma\sigma^\top s^e(s^e)^\top\right) \quad \text{and} \quad \max_{\sigma\in\Gamma} tr\left(\sigma\sigma^\top s^e(s^e)^\top\right). \tag{49}$$

If $d = 1$, then Γ is a compact interval and $\underline{\sigma}$ and $\overline{\sigma}$ are its left and right endpoints. Let P^* be the measure on Ω induced by P_0 and X^*, where

$$X_t^* = \overline{\sigma}^\top B_t, \quad \text{for all } t \text{ and } \omega;$$

define P^{**} similarly using $\underline{\sigma}$ and X^{**}. Then, by a slight extension of the observation in Example 1, P^* is a dynamic minimizer for e if $\alpha < 0$ and P^{**} is a dynamic minimizer for e if $\alpha > 0$.

Remark 4. That the minimizing measure corresponds to constant volatility is a feature of this example. More generally, the minimizing measure in $\mathcal{P}$ defines a specific stochastic volatility model. It is interesting to note that when volatility is modeled by robustifying the Hull-White and Heston parametric forms, for example, the minimizing measure does not lie in either parametric class. Rather it corresponds to pasting the two alternatives together endogenously—that is, in a way that depends on the endowment process and on α.

We describe further implications assuming $\alpha < 0$; the corresponding statements for $\alpha > 0$ will be obvious to the reader. Interpretation of the sign of α is confounded by the dual role of α in the additive expected utility model. However, the example can be generalized to the Kreps-Porteus aggregator (40), and then the same characterization of the worst-case volatility is valid, with $1-\alpha$ interpretable as the measure of relative risk aversion. Therefore, the intuition is clear for the pricing results that follow: only the largest (in the sense of (49)) volatility $\overline{\sigma}$ is relevant, assuming $\alpha < 0$, because it represents the worst-case scenario given a large (greater than 1) measure of relative risk aversion.

Corresponding regular conditionals have a simple form. For example, $(P^*)^{\omega}_{\tau}$ is the measure on Ω induced by the stochastic differential equation (under P_0)

$$\begin{cases} dX_t &= \overline{\sigma} dB_t, \tau \le t \le T \\ X_t &= \omega_t, 0 \le t \le \tau \end{cases}$$

Thus under $(P^*)^{\omega}_{\tau}$, the increment $B_t - B_\tau$ is $N\left(0, \overline{\sigma\sigma}^\top (t-\tau)\right)$ for $\tau \le t \le T$.

The C-CAPM (45) takes the form (assuming $\alpha < 0$): For every τ, $\mathcal{P}$-almost surely in ω,

$$b_t - r_t \mathbf{1} = (1-\alpha) s_t \left(\overline{\sigma\sigma}^\top\right) s^e \quad \text{on } [\tau, T] \ \ (P^*)^{\omega}_{\tau}\text{-}a.s.$$

For comparison purposes, it is convenient to express this equation partially in terms of P_0. The measures P_0 and P^* differ only via the change of variables defined via the SDE (15). Therefore, we arrive at the following equilibrium condition: *For every τ, $\mathcal{P}$-almost surely in ω,*

$$\widehat{b_t} - \widehat{r_t} \mathbf{1} = (1-\alpha) \widehat{s_t} \left(\overline{\sigma\sigma}^\top\right) s^e \quad \text{on } [\tau, T] \ \ (P_0)^{\omega}_{\tau}\text{-}a.s. \tag{50}$$

where $\widehat{b_t} = b_t\left(X^{\overline{\sigma}}_{\cdot}\right)$, $\widehat{r_t} = r_t\left(X^{\overline{\sigma}}_{\cdot}\right)$ and $\widehat{s_t} = s_t\left(X^{\overline{\sigma}}_{\cdot}\right)$, corresponding to the noted change of variables under which $B_t \longmapsto X^{\overline{\sigma}}_t = \overline{\sigma}^\top B_t$. Note that the difference between random variables with and without hats is ultimately not important because they follow identical distributions under $(P_0)^{\omega}_{\tau}$ and $(P^*)^{\omega}_{\tau}$, respectively.

The impact of ambiguous volatility is most easily seen by comparing with the standard C-CAPM obtained assuming complete confidence in the single probability law P_0, which renders B a standard Brownian motion. Then the prediction for asset returns is

$$b_t - r_t \mathbf{1} = (1-\alpha) s_t s^e \quad \text{on } [0, T] \ \ P_0\text{-}a.s.$$

or equivalently: *For every* τ *and* P_0*-almost surely in* ω,

$$b_t - r_t \mathbf{1} = (1-\alpha) s_t s^e \quad \text{on } [\tau, T] \;\; (P_0)^{\omega}_{\tau}\text{-}a.s. \tag{51}$$

There are two differences between the latter and the equilibrium condition (50) for our model. First, the "instantaneous covariance" between asset returns and consumption is modified from $s_t s^e$ to $\widehat{s_t}(\overline{\sigma}\overline{\sigma}^{\top}) s^e$, reflecting the fact that $\overline{\sigma}$ is the worst-case volatility scenario for the representative agent. Such an effect, whereby ambiguity leads to standard equilibrium conditions except that the reference measure is replaced by the worst-case measure, is familiar from the literature. The second difference is new. Condition (51) refers to the single measure P_0 only and events that are null under P_0 are irrelevant.[36] In contrast, the condition (50) is required to hold $\mathcal{P}$-almost surely in ω because, as described in Example 2, dynamic consistency requires that possibility be judged according to all priors in $\mathcal{P}$.

Turn to a brief consideration of corresponding equilibrium prices. Fix a dividend stream $(\delta, \delta_T) \in M^2 \times \widehat{L^2}(\Omega)$ where the security is available in zero net supply. Then its equilibrium price $S^{\delta} = \left(S^{\delta}_{\tau}\right)$ is given by: For all τ, $\mathcal{P}$-$a.s.$ in ω,

$$S^{\delta}_{\tau} = E^{(P^*)^{\omega}_{\tau}} \left[\int_{\tau}^{T} \frac{\pi^e_t}{\pi^e_{\tau}} \delta_t dt + \frac{\pi^e_T}{\pi^e_{\tau}} \delta_T \right], \tag{52}$$

which lies between the hedging bounds in Theorem 1 by (44).[37] It is interesting to compare this equilibrium pricing rule with the price bounds derived from hedging arguments (Theorem 1). Suppose the security in question is an option on an underlying. Under the conditions of Example 3, the volatilities used to define the upper and lower price bounds depend on whether the terminal payoff is (globally) convex or concave as a function of the price of the underlying. In contrast, the volatility used for equilibrium pricing is the same for all options (and other securities) and depends only on the endowment and the preference parameter α. This difference is further illustrated below.

If $\delta = e$, then elementary calculations yield the time τ price of the endowment stream in the form

$$S^e_{\tau} = A_{\tau} e_{\tau} = A_{\tau} e_0 \exp\left((s^e)^{\top} B_{\tau} \right),$$

where $A_{\tau} > 0$ is deterministic and $A_T = 1$. Thus $\log\left(S^e_{\tau}/A_{\tau}\right) = \log e_{\tau}$ and the logarithm of (deflated) price is also a G-Brownian motion. We can also price an option on the endowment. Thus let $\delta_t = 0$ for $0 \le t < T$ and $\delta_T = \psi\left(S^e_T\right)$. Denote its price process by S^{ψ}. From (52), any such derivative is priced in equilibrium as though σ_t were constant at $\overline{\sigma}$ (or at $\underline{\sigma}$ if $\alpha > 0$). In particular, for a European call option where $\delta_T = \left(S^e_T - \kappa\right)^+$, its equilibrium price at τ is

36 Similarly for the C-CAPM (47) when only drift is ambiguous, because then all priors are equivalent to P_0.

37 The proof is analogous to that of Lemma 4, particularly surrounding (B8).

$BS_\tau\left((s^e)^\top\overline{\sigma}, T, \kappa\right)$, where the latter term denotes the Black-Scholes price at τ for a call option with strike price κ and expiry time T when the underlying security price process is geometric Brownian motion with volatility $(s^e)^\top\overline{\sigma}$. Thus the Black-Scholes implied variance is $tr\left(\overline{\sigma\sigma}^\top s^e(s^e)^\top\right)$, which exceeds every conceivable realized variance $tr\left(\sigma\sigma^\top s^e(s^e)^\top\right)$, $\sigma\in\Gamma$, consistent with a documented empirical feature of option prices.

4. Concluding Remarks

We have described a model of utility over continuous-time consumption streams that can accommodate ambiguity about volatility. Such ambiguity necessitates dropping the assumption that a single measure defines null events, which is a source of considerable technical difficulty. The economic motivation provided for confronting the technical challenge is the importance of stochastic volatility modeling in both financial economics and macroeconomics, the evidence that the dynamics of volatility are complicated and difficult to pin down empirically, and the presumption that complete confidence in any single parametric specification is unwarranted and implausible. (Recall, for example, the quote in Section 1.1 from Carr and Lee (2009).) These considerations suggest the potential usefulness of "robust stochastic volatility" models (Section 1.1). We have shown that important elements of representative agent asset pricing theory extend to an environment with ambiguous volatility. We also provided one example of the added explanatory power of ambiguous volatility; it gives a way to understand the documented feature of option prices whereby the Black-Scholes implied volatility exceeds the realized volatility of the underlying security. However, a question that remains to be answered more broadly and thoroughly is "does ambiguity about volatility and possibility matter empirically?" In particular, it remains to determine the empirical content of the derived C-CAPM relations. The contribution of this paper has been to provide a theoretical framework within which one could address such questions.

There are also several extensions at the theoretical level that seem worth pursuing. The utility formulation should be generalized to environments with jumps, particularly in light of the importance attributed to jumps for understanding options markets. The asset market analysis should be extended to permit ambiguity specifications more general than G-Brownian motion. Extension to heterogeneous agent economies is important and intriguing. The nonequivalence of measures raises questions about existence of equilibrium and about the nature of no-arbitrage pricing (for reasons discussed in Willard and Dybvig 1999).

Two further questions that merit attention are more in the nature of refinements, albeit nontrivial ones and beyond the scope of this paper. First, the fact that utility is recursive but not strictly so suggests that although not every time 0 optimal plan may be pursued subsequently at all relevant nodes, one might expect that (under suitable regularity conditions) there exists at least one time 0 optimal plan that will be implemented. (This is the case in Example 2 and also in the asset market example in Section 3.4.) Sufficient conditions for such existence should be explored. Second, Sections 3.3 and 3.4 demonstrated the significance of worst-case scenarios in the form of dynamic minimizing measures. Their existence and characterization pose important questions.

In terms of applications, we note that the model (slightly modified) can be interpreted in terms of investor sentiments. Replace all infima by suprema and vice versa. Then, the consumer may be described as an ambiguity lover, or alternatively in terms of optimism and overconfidence. For example, in a recent study of how the pricing kernel is affected by sentiment, Barone-Adesi, Mancini, and Shefrin (2012) subdivide the latter and define optimism as occurring when the investor overestimates mean returns and overconfidence as occurring when return volatility is underestimated. This fits well with the distinction we have emphasized at a formal modeling level between ambiguity about drift and ambiguity about volatility. In a continuous-time setting, ambiguity about drift, or optimism, can be modeled in a probability space framework, but not so ambiguity about volatility, or overconfidence.[38]

We mention one more potential application. Working in a discrete-time setting, Epstein and Schneider (2008) point to ambiguous volatility as a way to model signals with ambiguous precision. This leads to a new way to measure information quality that has interesting implications for financial models (see also Illeditsch 2011). The utility framework that we provide should permit future explorations of this dimension of information quality in continuous-time settings.

Appendix

A. G-Brownian Motion

Peng (2006) introduced G-Brownian motion using PDEs (specifically, a nonlinear heat equation). Further contributions are due to Denis, Hu, and Peng (2011) and Soner, Touzi, and Zhang (2011a). For the convenience of the reader, in this appendix we outline some key elements of the theory of G-Brownian motion in terms of the specifics of our model.

Itô Integral and Quadratic Variation Process: For each $\eta \in M^2$, we can consider the usual Itô integral $\int_0^T \eta_t^{\intercal} dB_t$, which lies in $\widehat{L^2}(\Omega)$. Each $P \in \mathcal{P}$ provides a different perspective on the integral; a comprehensive view requires that one consider all priors. The quadratic variation process $\langle B \rangle$ also agrees with the usual quadratic variation process $\mathcal{P}$-$a.s$. In Section 3.1 we defined a universal process v (via (28)) and proved that

$$\langle B \rangle = \left(\int_0^t v_s ds : 0 \le t \le T \right).$$

[38] The applied finance literature has not used sets of priors in modeling sentiment. The use of sets gives a best scenario, or subjective prior, that depends on the portfolio being evaluated. Thus, optimism can be exhibited for every portfolio, as one might expect of an investor who has an optimistic nature. In contrast, when the subjective prior is fixed, then a high estimated return for a security implies pessimism when the agent considers going short.

The following properties are satisfied for any $\lambda, \eta \in M^2$, $X \in \widehat{L^2}(\Omega_T)$, and constant α:

$$\hat{E}[B_t] = 0,$$

$$\hat{E}\left[\int_0^T \eta_t^\top dB_t\right] = 0,$$

$$\hat{E}\left[\left(\int_0^T \eta_t^\top dB_t\right)^2\right] = \hat{E}\left[\int_0^T \eta_t^\top v_t \eta_t dt\right],$$

$$\int_0^T \left(\alpha\eta_t^\top + \lambda_t^\top\right) dB_t = \alpha \int_0^T \eta_t^\top dB_t + \int_0^T \lambda_t^\top dB_t,$$

$$\hat{E}\left[X + \int_s^T \eta_t^\top dB_t \,\middle|\, \mathcal{F}_s\right] = \hat{E}[X \mid \mathcal{F}_s] + \hat{E}\left[\int_s^T \eta_t^\top dB_t \,\middle|\, \mathcal{F}_s\right] = \hat{E}[X \mid \mathcal{F}_s]$$

For the one-dimensional case ($\Gamma = [\underline{\sigma}, \bar{\sigma}]$, $\underline{\sigma} > 0$), we have

$$\underline{\sigma}^2 t \leq \hat{E}\left[(B_t)^2\right] \leq \bar{\sigma}^2 t,$$

$$\underline{\sigma}^2 \hat{E}\left[\int_0^T \eta_t^2 dt\right] \leq \hat{E}\left[\left(\int_0^T \eta_t dB_t\right)^2\right] \leq \bar{\sigma}^2 \hat{E}\left[\int_0^T \eta_t^2 dt\right].$$

Itô's formula: Consider

$$X_t = X_0 + \int_0^t \alpha_s ds + \int_0^t \gamma_s dB_s$$

where α and γ are in $M^2(\mathbb{R}^d)$ and $M^2(\mathbb{R}^{d\times d})$, respectively. (Define $M^2(\mathbb{R}^{\ell\times k})$ similarly to M^2 for $\mathbb{R}^{\ell\times k}$-valued processes.) We adapt Itô's formula from Li and Peng (2011, Theorem 5.4) or Soner, Touzi, and Zhang (2011b, Proposition 6.7) and rewrite it in our context. Let $0 \leq \tau \leq t \leq T$; define $v = (v^{ij})$ by (28). Then, for any function $f : \mathbb{R}^d \to \mathbb{R}$ with continuous second-order derivatives, we have

$$f(X_t) - f(X_\tau) = \int_\tau^t (f_x(X_s))^\top \gamma_s dB_s + \int_\tau^t (f_x(X_s))^\top \alpha_s ds + \frac{1}{2}\int_\tau^t tr\left[\gamma_s^\top f_{xx}(X_s) v_s \gamma_s\right] ds.$$

Consider the special case $f(x_1, x_2) = x_1 x_2$ and

$$X_t^i = X_0^i + \int_0^t \alpha_s^i ds + \int_0^t \gamma_s^i dB_s, \quad i = 1, 2,$$

where $\alpha^i \in M^2$ and $\gamma^i \in M^2(\mathbb{R}^d)$, $i = 1, 2$. Then

$$X_t^1 X_t^2 - X_\tau^1 X_\tau^2 = \int_\tau^t X_s^1 dX_s^2 + \int_\tau^t X_s^2 dX_s^1 + \int_\tau^t \gamma_s^1 v_s \left(\gamma_s^2\right)^\top ds.$$

Formal rules: As in the classical Itô formula, if $dX_t = \alpha_t dt + \gamma_t dB_t$, then we can compute $(dX_t)^2 = (dX_t) \cdot (dX_t)$ by the following formal rules:

$$dt \cdot dt = dt \cdot dB_t = dB_t \cdot dt = 0, \quad dB_t \cdot dB_t = v_t dt.$$

Martingale representation theorem: An $\mathcal{F}$-progressively measurable $\widehat{L^2}(\Omega)$-valued process X is called a *G-martingale* if and only if for any $0 \leq \tau < t$, $X_\tau = \hat{E}[X_t \mid \mathcal{F}_\tau]$. We adapt the martingale representation theorem from Song (2011) and Soner, Touzi, and Zhang (2011a). For any $\xi \in \widehat{L^{2+\varepsilon}}(\Omega)$ and $\varepsilon > 0$, if $X_t = \hat{E}[\xi \mid \mathcal{F}_t]$, $t \in [0, T]$, then we have the following unique decomposition:

$$X_t = X_0 + \int_0^t Z_s dB_s - K_t,$$

where $Z \in M^2$, K is a continuous nondecreasing process with $K_0 = 0$, $K_T \in \widehat{L^2}(\Omega)$ and where $-K$ is a G-martingale.

B. Proofs for Asset Returns

B.1 Proof of Theorem 1

Lemma 1. Consider the following backward stochastic differential equation (BSDE) driven by G-Brownian motion:

$$d\tilde{Y}_t = \left(r_t \tilde{Y}_t + \eta_t^\top \tilde{\phi}_t - \delta_t\right) dt - dK_t + \tilde{\phi}_t^\top dB_t,$$
$$\tilde{Y}_T = \delta_T.$$

Denote by $I(0,T)$ the space of all continuous nondecreasing processes $(K_t)_{0\leq t\leq T}$ with $K_0=0$ and $K_T \in \widehat{L^2}(\Omega)$. Then there exists a unique triple

$$\left(\tilde{Y}_t, \tilde{\phi}_t, K_t\right) \in M^2 \times M^2 \times I(0,T),$$

satisfying the BSDE such that $K_0=0$ and where $-K_t$ is a G-martingale.

Proof. Apply Itô's formula to $\pi_t \tilde{Y}_t$ to derive

$$\begin{aligned} d\left(\pi_t \tilde{Y}_t\right) &= \pi_t d\tilde{Y}_t + \tilde{Y}_t d\pi_t - \left\langle \pi_t \tilde{\phi}_t^\top, \eta_t^\top v_t^{-1} d\langle B\rangle_t \right\rangle \\ &= \left(\pi_t \tilde{\phi}_t - \pi_t \tilde{Y}_t \eta_t^\top v_t^{-1}\right) dB_t - \pi_t \delta_t dt - \pi_t dK_t + \left[\pi_t \tilde{\phi}_t^\top \eta_t dt - \left\langle \pi_t \tilde{\phi}_t^\top, \eta_t^\top v_t^{-1} d\langle B\rangle_t \right\rangle\right] \\ &= \left(\pi_t \tilde{\phi}_t - \pi_t \tilde{Y}_t \eta_t^\top v_t^{-1}\right) dB_t - \pi_t \delta_t dt - \pi_t dK_t. \end{aligned}$$

Integrate on both sides to obtain

$$\pi_T \delta_T + \int_\tau^T \pi_t \delta_t dt = \pi_\tau \tilde{Y}_\tau - \int_\tau^T \pi_t dK_t + \int_\tau^T \left(\pi_t \tilde{\phi}_t^\top - \pi_t \tilde{Y}_t \eta_t^\top v_t^{-1}\right) dB_t. \tag{B1}$$

Let

$$X_\tau = \hat{E}\left[\pi_T \delta_T + \int_0^T \pi_t \delta_t dt \middle| \mathcal{F}_\tau\right].$$

Then (X_τ) is a G-martingale. By the martingale representation theorem (Appendix A), there exists a unique pair $(Z_t, \overline{K}_t) \in M^2 \times I(0,T)$ such that

$$X_\tau = \hat{E}\left[\pi_T \delta_T + \int_0^T \pi_t \delta_t dt\right] + \int_0^\tau Z_t dB_t - \overline{K}_t,$$

and such that $-\overline{K}_t$ is a G-martingale. This can be rewritten as

$$X_\tau = X_T - \int_\tau^T Z_t dB_t + \overline{K}_T - \overline{K}_t = \pi_T \delta_T + \int_0^T \pi_t \delta_t dt - \int_\tau^T Z_t dB_t + \overline{K}_T - \overline{K}_\tau.$$

Thus $(\tilde{Y}_t, \tilde{\phi}_t, K_t)$ is the desired solution where

$$\tilde{Y}_\tau = \frac{X_\tau}{\pi_\tau} - \int_0^\tau \frac{\pi_t}{\pi_\tau} \delta_t dt, \quad \tilde{\phi}_\tau^\top = \frac{Z_\tau}{\pi_\tau} + \tilde{Y}_\tau \eta_\tau^\top v_t^{-1}, \quad \text{and} \quad K_\tau = \int_0^\tau \frac{1}{\pi_t} d\overline{K}_t. \quad \blacksquare$$

Turn to proof of the theorem. We prove only the claim related to superhedging. Proof of the other claim is similar.

Step 1: Prove that for any $y \in \mathcal{U}_\tau$,

$$y \geq \hat{E}\left[\int_\tau^T \frac{\pi_t}{\pi_\tau}\delta_t dt + \frac{\pi_T}{\pi_\tau}\delta_T \,\middle|\, \mathcal{F}_\tau\right]$$

If $y \in \mathcal{U}_\tau$, there exists ϕ such that $Y_T^{y,\phi,\tau} \geq \delta_T$. Apply (the G-Brownian version of) Itô's formula to $\pi_t Y_t^{y,\phi,\tau}$ to derive

$$d\left(\pi_t Y_t^{y,\phi,\tau}\right) = \pi_t dY_t^{y,\phi,\tau} + Y_t^{y,\phi,\tau} d\pi_t - \left\langle \pi_t \phi_t^\top, \eta_t^\top v_t^{-1} d\langle B\rangle_t \right\rangle$$

$$= \left(\pi_t\phi_t - \pi_t Y_t^{y,\phi,\tau}\eta_t^\top v_t^{-1}\right) dB_t - \pi_t\delta_t dt + \left[\pi_t\phi_t^\top \eta_t dt - \left\langle \pi_t\phi_t^\top, \eta_t^\top v_t^{-1} d\langle B\rangle_t\right\rangle\right].$$

Integration on both sides yields

$$\pi_T Y_T^{y,\phi,\tau} + \int_\tau^T \pi_t\delta_t dt = \pi_\tau y + \int_\tau^T \left(\pi_t\phi_t^\top - \pi_t Y_t^{y,\phi,\tau}\eta_t^\top v_t^{-1}\right) dB_t,$$

and taking conditional expectations yields

$$y = \hat{E}\left[\frac{\pi_T}{\pi_\tau} Y_T^{y,\phi,\tau} + \int_\tau^T \frac{\pi_t}{\pi_\tau}\delta_t dt \,\middle|\, \mathcal{F}_\tau\right]$$

$$\geq \hat{E}\left[\frac{\pi_T}{\pi_\tau}\delta_T + \int_\tau^T \frac{\pi_t}{\pi_\tau}\delta_t dt \,\middle|\, \mathcal{F}_\tau\right].$$

Step 2: There exists $\hat{y} \in \mathcal{U}_\tau$ and $\hat{\phi}$ such that $Y_T^{\hat{y},\hat{\phi},\tau} \geq \delta_T$ and

$$\hat{y} = \hat{E}\left[\int_\tau^T \frac{\pi_t}{\pi_\tau}\delta_t dt + \frac{\pi_T}{\pi_\tau}\delta_T \,\middle|\, \mathcal{F}_\tau\right].$$

Apply the preceding lemma. Rewrite equation (B1) as

$$\pi_T\delta_T + \int_\tau^T \pi_t\delta_t dt = \pi_\tau \tilde{Y}_\tau + \int_\tau^T \left(\pi_t\tilde{\phi}_t^\top - \pi_t\tilde{Y}_t\eta_t^\top v_t^{-1}\right) dB_t - \int_\tau^T \pi_t dK_t. \tag{B2}$$

Because π_t is positive and $-K_t$ is a G-martingale, $\hat{E}[-\int_\tau^T \pi_t dK_t \mid \mathcal{F}_\tau] = 0$. Thus,

$$\hat{E}\left[\pi_T\delta_T + \int_\tau^T \pi_t\delta_t dt \,\middle|\, \mathcal{F}_\tau\right] = \pi_\tau \tilde{Y}_\tau$$

Finally, define $\hat{y} = \tilde{Y}_\tau$ and $\hat{\phi} = \tilde{\phi}$. Then $\hat{y} \in \mathcal{U}_\tau$ and

$$\hat{y} = \hat{E}\left[\int_\tau^T \frac{\pi_t}{\pi_\tau}\delta_t dt + \frac{\pi_T}{\pi_\tau}\delta_T \,\middle|\, \mathcal{F}_\tau\right].$$

This completes the proof of Theorem 1.

B.2 Proof of Theorem 2

(i) Apply Itô's formula for G-Brownian motion to derive[39]

$$d(\pi_t Y_t) = \pi_t dY_t + Y_t d\pi_t - \left\langle \pi_t \phi_t^\top, \eta_t^\top v_t^{-1} d\langle B \rangle_t \right\rangle \tag{B3}$$

$$= \left(\pi_t \phi_t - \pi_t Y_t \eta_t^\top v_t^{-1}\right) dB_t - \pi_t (c_t - e_t) dt + \left[\pi_t \phi_t^\top \eta_t dt - \left\langle \pi_t \phi_t^\top, \eta_t^\top v_t^{-1} d\langle B \rangle_t \right\rangle\right].$$

Note that for any $a = (a_t) \in M^2$,

$$\int_\tau^T a_t v_t^{-1} d\langle B \rangle_t = \int_\tau^T a_t dt,$$

and therefore,

$$\int_\tau^T \left\langle \pi_t \phi_t, \eta_t^\top v_t^{-1} d\langle B \rangle_t \right\rangle = \int_\tau^T \pi_t \phi_t^\top \eta_t dt.$$

Accordingly, integration on both sides of (B3) yields,

$$\pi_T Y_T + \int_\tau^T \pi_t (c_t - e_t) dt = \pi_\tau Y_\tau + \int_\tau^T \left(\pi_t \phi_t^\top - \pi_t Y_t \eta_t^\top v_t^{-1}\right) dB_t.$$

Take conditional expectations to obtain

$$\hat{E}\left[\pi_T Y_T + \int_\tau^T \pi_t (c_t - e_t) dt \,\middle|\, \mathcal{F}_\tau\right] = \pi_\tau Y_\tau + \hat{E}\left[\int_\tau^T \left(\pi_t \phi_t^\top - \pi_t Y_t \eta_t^\top v_t^{-1}\right) dB_t \,\middle|\, \mathcal{F}_\tau\right].$$

Because B being G-Brownian motion implies that B is a martingale under every prior in $\mathcal{P}$, we have

$$0 = \hat{E}\left[\int_\tau^T \left(\pi_t \phi_t^\top - \pi_t Y_t \eta_t^\top v_t^{-1}\right) dB_t \,\middle|\, \mathcal{F}_\tau\right] = \hat{E}\left[-\int_\tau^T \left(\pi_t \phi_t^\top - \pi_t Y_t \eta_t^\top v_t^{-1}\right) dB_t \,\middle|\, \mathcal{F}_\tau\right],$$

which gives the desired result.

(ii) We need to find a process ϕ such that, for the given c, the solution (Y_t) to

$$dY_t = \left(r_t Y_t + \eta_t^\top \phi_t - (c_t - e_t)\right) dt + \phi_t^\top dB_t, \quad t \in [\tau, T]$$

$$Y_T = c_T - e_T$$

has time τ wealth equal to the given value Y_τ.

For $\tau \leq s \leq T$, define

$$X_s \equiv \hat{E}\left[\int_\tau^T \frac{\pi_t}{\pi_\tau}(c_t - e_t) dt + \frac{\pi_T}{\pi_\tau}(c_T - e_T) \,\middle|\, \mathcal{F}_s\right].$$

Then $X_s = -\hat{E}[-\int_\tau^T \frac{\pi_t}{\pi_\tau}(c_t - e_t) dt - \frac{\pi_T}{\pi_\tau}(c_T - e_T) \,|\, \mathcal{F}_s]$ and $(X_s)_{\tau \leq s \leq T}$ is a symmetric G-martingale. By Soner, Touzi, and Zhang (2011a) and Song (2011), it admits the unique representation

$$X_s = X_\tau + \int_\tau^s Z_t^\top dB_t,$$

[39] For any d-dimensional (column) vectors x and y, we use $\langle x^\top, y^\top \rangle$ occasionally as alternative notation for the inner product $x^\top y$.

where $Z \in M^2$. Note that

$$X_\tau = \hat{E}\left[\int_\tau^T \frac{\pi_t}{\pi_\tau}(c_t - e_t)dt + \frac{\pi_T}{\pi_\tau}(c_T - e_T)\big|\mathcal{F}_\tau\right] = Y_\tau.$$

Set

$$\bar{Y}_s \equiv X_s - \int_\tau^s \frac{\pi_t}{\pi_\tau}(c_t - e_t)dt, \quad s \in [\tau, T].$$

Then $(\bar{Y}_s)$ satisfies

$$d\bar{Y}_s = -\frac{\pi_s}{\pi_\tau}(c_s - e_s)ds + Z_s^\top dB_s, \quad \bar{Y}_\tau = Y_\tau.$$

Define

$$Y_s \equiv \bar{Y}_s\left(\frac{\pi_s}{\pi_\tau}\right)^{-1}.$$

Note that (π_s) satisfies

$$d\pi_s/\pi_s = -r_s ds - \eta_s^\top v_s^{-1} dB_s, \quad s \in [\tau, T].$$

Apply Itô's formula for G-Brownian motion to derive

$$dY_s = \left[r_s Y_s + \eta_s^\top\left(Y_s\left(v_s^{-1}\right)^\top \eta_s + \frac{\pi_\tau}{\pi_s} Z_s\right) - (c_s - e_s)\right]ds + \left(Y_s \eta_s^\top v_s^{-1} + \frac{\pi_\tau}{\pi_s} Z_s^\top\right) dB_s.$$

Finally, set

$$\phi_s^\top \equiv Y_s \eta_s^\top v_s^{-1} + \frac{\pi_\tau}{\pi_s} Z_s^\top.$$

Then

$$dY_s = \left(r_s Y_s + \eta_s^\top \phi_t - (c_s - e_s)\right)ds + \phi_s^\top dB_s, \quad s \in [\tau, T].$$

This completes the proof. ■

B.3 Proof of Theorem 3

The proof follows from Theorem 2 and the following lemma.

Lemma 2. For every c, we have: For each τ, $\mathcal{P}$-almost surely,

$$\widehat{E}\left[\int_\tau^T \pi_t^e (c_t - e_t)dt + \pi_T^e (c_T - e_T)\big|\mathcal{F}_\tau\right] \leq 0 \Longrightarrow V_\tau(c) \leq V_\tau(e). \tag{B4}$$

Proof. Define δ_t implicitly by

$$f(c_t, V_t(c)) = f_c(e_t, V_t(e))(c_t - e_t) + f_v(e_t, V_t(e))(V_t(c) - V_t(e)) - \delta_t + f(e_t, V_t(e)),$$

for $0 \leq t < T$, and

$$u(c_T) = u_c(e_T)(c_T - e_T) - \delta_T + u(e_T).$$

Because f and u are concave, we have $\delta_t \geq 0$ on $[0, T]$.

Define, for $0 \le t < T$,

$$\beta_t = f_v(e_t, V_t(e))$$

$$\gamma_t = f_c(e_t, V_t(e))(c_t - e_t) + f(e_t, V_t(e)) - \beta_t V_t(e) - \delta_t$$

$$\gamma_T = -u_c(e_T)e_T + u(e_T) - \delta_T$$

$$\zeta_t = f(e_t, V_t(e)) - \beta_t V_t(e).$$

Then

$$V_t(c) = -\hat{E}\left[-(u_c(e_T)c_T + \gamma_T) - \int_t^T (\beta_s V_s(c) + \gamma_s)ds \,\middle|\, \mathcal{F}_t\right].$$

Because this is a linear backward stochastic differential equation, its solution has the form (by Hu and Ji 2010)

$$V_t(c) = -\hat{E}\left[-(u_c(e_T)c_T + \gamma_T)\exp\left\{\int_t^T \beta_s ds\right\} - \int_t^T \gamma_s \exp\left\{\int_t^s \beta_{s'} ds'\right\} ds \,\middle|\, \mathcal{F}_t\right].$$

Similarly for e, we have

$$V_t(e) = -\hat{E}\left[-u(e_T) - \int_t^T (\beta_s V_s(e) + \zeta_s)ds \,\middle|\, \mathcal{F}_t\right],$$

and (by Hu and Ji 2010),

$$V_t(e) = -\hat{E}\left[-u(e_T)\exp\left\{\int_t^T \beta_s ds\right\} - \int_t^T \zeta_s \exp\left\{\int_t^s \beta_{s'} ds'\right\} ds \,\middle|\, \mathcal{F}_t\right].$$

Apply the subadditivity of $\hat{E}[\cdot \mid \mathcal{F}_\tau]$ and the nonnegativity of δ_t to obtain

$$\begin{aligned}
&\exp\left\{\int_0^\tau \beta_s ds\right\}(V_\tau(c) - V_\tau(e)) \\
&= -\hat{E}\left[-(u_c(e_T)c_T + \gamma_T)\exp\left\{\int_0^T \beta_s ds\right\} - \int_\tau^T \gamma_t \exp\left\{\int_0^t \beta_s ds\right\} dt \,\middle|\, \mathcal{F}_\tau\right] \\
&\quad - \left\{-\hat{E}\left[-u(e_T)\exp\left\{\int_0^T \beta_s ds\right\} - \int_\tau^T \zeta_t \exp\left\{\int_0^t \beta_s ds\right\} dt \,\middle|\, \mathcal{F}_\tau\right]\right\} \\
&= \hat{E}\left[-u(e_T)\exp\left\{\int_0^T \beta_s ds\right\} - \int_\tau^T \zeta_t \exp\left\{\int_0^t \beta_s ds\right\} dt \,\middle|\, \mathcal{F}_\tau\right] \\
&\quad - \hat{E}\left[-(u_c(e_T)c_T + \gamma_T)\exp\left\{\int_0^T \beta_s ds\right\} - \int_\tau^T \gamma_t \exp\left\{\int_0^t \beta_s ds\right\} dt \,\middle|\, \mathcal{F}_\tau\right] \\
&\le \hat{E}\left[-u(e_T)\exp\left\{\int_0^T \beta_s ds\right\} - \int_\tau^T \zeta_t \exp\left\{\int_0^t \beta_s ds\right\} dt - (-(u_c(e_T)c_T + \gamma_T)\exp\left\{\int_0^T \beta_s ds\right\}\right. \\
&\quad \left. - \int_\tau^T \gamma_t \exp\left\{\int_0^t \beta_s ds\right\} dt \,\middle|\, \mathcal{F}_\tau\right] \\
&= \hat{E}\left[(u_c(e_T)c_T + \gamma_T)\exp\left\{\int_0^T \beta_s ds\right\} + \int_\tau^T \gamma_t \exp\left\{\int_0^t \beta_s ds\right\} dt - u(e_T)\exp\left\{\int_0^T \beta_s ds\right\}\right. \\
&\quad \left. - \int_\tau^T \zeta_t \exp\left\{\int_0^t \beta_s ds\right\} dt \,\middle|\, \mathcal{F}_\tau\right]
\end{aligned}$$

$$=\hat{E}\left[\exp\left\{\int_0^T \beta_s ds\right\}u_c(e_T)(c_T-e_T)+\int_\tau^T \exp\left\{\int_0^t \beta_s ds\right\} f_c(e_t,V_t(e))(c_t-e_t)dt\right.$$
$$\left.-\exp\left\{\int_0^T \beta_s ds\right\}\delta_T-\int_\tau^T \exp\left\{\int_0^t \beta_s ds\right\}\delta_t dt\big|\mathcal{F}_\tau\right]$$
$$\leq\hat{E}\left[\exp\left\{\int_0^T \beta_s ds\right\}u_c(e_T)(c_T-e_T)+\int_\tau^T \exp\left\{\int_0^t \beta_s ds\right\} f_c(e_t,V_t(e))(c_t-e_t)dt\big|\mathcal{F}_\tau\right]$$
$$=\widehat{E}\left[\pi_T^e(c_T-e_T)+\int_\tau^T \pi_t^e(c_t-e_t)dt\big|\mathcal{F}_\tau\right].$$

This completes the proof. ■

B.4 C-CAPM for General Aggregators

We derive (41) and the corresponding form of the C-CAPM for general aggregators, thus justifying claims made following Theorem 3. Utility is defined by (33).

Lemma 3. For given $c\in D$, there is a unique solution V_t to

$$V_t(c)=-\hat{E}\left[-\int_t^T f(c_s,V_s(c))ds-u(c_T)\big|\mathcal{F}_t\right],\quad 0\leq t\leq T.$$

Further, there exist unique $Z\in M^2$ and K (a continuous nondecreasing process with $K_0=0$) such that

$$V_t=u(c_T)+\int_t^T f(c_s,V_s(c))ds+\int_t^T Z_s dB_s-K_T+K_t.$$

Proof. Define

$$U_t=-\hat{E}\left[-\int_0^T f(c_s,V_s(c))ds-u(c_T)\big|\mathcal{F}_t\right],\quad 0\leq t\leq T.$$

Note that

$$U_0=-\hat{E}\left[-\int_0^T f(c_s,V_s(c))ds-u(c_T)\right],$$
$$U_T=\int_0^T f(c_s,V_s(c))ds-u(c_T).$$

Because $-U_t$ is a G-martingale, it has the following unique representation:

$$-U_t=-U_0+\int_0^t Z_s dB_s-K_t.$$

Then

$$V_t=-\hat{E}\left[-\int_t^T f(c_s,V_s(c))ds-u(c_T)\big|\mathcal{F}_t\right]=U_t-\int_0^t f(c_s,V_s(c))ds$$
$$=U_0-\int_0^t Z_s dB_s+K_t-\int_0^t f(c_s,V_s(c))ds.$$

Note that

$$V_T = u(c_T) = U_0 - \int_0^T Z_s dB_s + K_T - \int_0^T f(c_s, V_s(c))ds \Longrightarrow$$

$$V_t - V_T = V_t - u(c_T) = \int_t^T f(c_s, V_s(c))ds + \int_t^T Z_s dB_s - K_T + K_t \Longrightarrow$$

$$V_t = u(c_T) + \int_t^T f(c_s, V_s(c))ds + \int_t^T Z_s dB_s - K_T + K_t.$$

Uniqueness of (V_t) follows by standard contraction mapping arguments (see our companion paper). ■

The preceding representation of utility, combined with Itô's lemma for G-Brownian motion, yields

$$b_t - r_t 1 = -\frac{f_{cc}(e_t, V_t)e_t}{f_c(e_t, V_t)} s_t v_t s_t^e + \frac{f_{cV}(e_t, V_t)}{f_c(e_t, V_t)} s_t v_t Z_t^\top. \tag{B5}$$

For the Kreps-Porteus aggregator (40), this becomes

$$b_t - r_t 1 = (1-\rho)s_t v_t s_t^e + \frac{\alpha - \rho}{\alpha} s_t v_t \frac{Z_t^\top}{V_t}. \tag{B6}$$

Then (41) follows from the following relation (which can be proven as in Chen and Epstein (2002) by exploiting the homogeneity of degree α of utility):

$$Z_t/(\alpha V_t) = \rho^{-1}\left[s^M + (\rho - 1)s_t^e\right].$$

B.5 Proof of Theorem 4

The strategy is to argue that for any $c \in \Upsilon_\tau(0)$,

$$\begin{aligned}
V_\tau(c) - V_\tau(e) &= V_\tau(c) - E^{(P^*)^\omega_\tau}\left[\int_\tau^T f(e_s, V_s(e))ds + u(e_T)\right] \\
&{}^\dagger \leq E^{(P^*)^\omega_\tau}\left[\int_\tau^T f(c_s, V_s(c))ds + u(c_T)\right] \\
&\quad - E^{(P^*)^\omega_\tau}\left[\int_\tau^T f(e_s, V_s(e))ds + u(e_T)\right] \\
&\leq E^{(P^*)^\omega_\tau}\left[\int_\tau^T \pi_t^e(c_t - e_t)dt + \pi_T^e(c_T - e_T)\right] \\
\text{by (44)} &= f_c(e_0, V_0(e)) E^{(P^*)^\omega_\tau}\left[\int_\tau^T \pi_t(c_t - e_t)dt + \pi_T(c_T - e_T)\right] \\
&{}^{\dagger\dagger} \leq f_c(e_0, V_0(e)) \widehat{E}\left[\int_\tau^T \pi_t(c_t - e_t)dt + \pi_T(c_T - e_T)\middle|\mathcal{F}_\tau\right] \leq 0.
\end{aligned}$$

The inequalities marked † and †† are justified in the next lemma.

Lemma 4. For every τ, $\mathcal{P}$-almost surely,

$$V_\tau(c) \leq E^{(P^*)^\omega_\tau}\left[\int_\tau^T f(e_s, V_s(e))ds + u(e_T)\right], \tag{B7}$$

and

$$E^{(P^*)^\omega_\tau}\left[\int_\tau^T \pi_t(c_t - e_t)dt + \pi_T(c_T - e_T)\right] \leq \widehat{E}\left[\int_\tau^T \pi_t(c_t - e_t)ds + \pi_T(c_T - e_T)\middle|\mathcal{F}_\tau\right].$$

Proof. We prove the first inequality. The second is proven similarly.

We claim that for any $P \in \mathcal{P}$ and τ, there exists $\overline{P} \in \mathcal{P}$ such that

$$\overline{P} = P \quad \text{on} \quad \mathcal{F}_\tau \quad \text{and} \quad \overline{P}_t^\omega = (P^*)_t^\omega \quad \text{for all } (t,\omega) \in [\tau, T] \times \Omega. \tag{B8}$$

This follows from the construction of priors in $\mathcal{P}$ via the SDE (15). Let P^* and P be induced by θ^* and θ, respectively, and define $\overline{\theta} \in \Theta$ by

$$\overline{\theta}_t = \begin{cases} \theta_t & 0 \le t \le \tau \\ \theta_t^* & \tau < t \le T. \end{cases}$$

Then $\overline{P} = P^{\overline{\theta}}$ satisfies (B8). It follows from the detailed construction of conditional expectation $\widehat{E}[\cdot \mid \mathcal{F}_\tau]$, that P-$a.e.$

$$\begin{aligned} V_\tau(c) &\le E^{\overline{P}}\left[\int_\tau^T f(c_s, V_s(c))ds + u(c_T) \Big| \mathcal{F}_\tau\right] \\ &= E^{\overline{P}_\tau^\omega}\left[\int_\tau^T f(c_s, V_s(c))ds + u(c_T)\right] \\ &= E^{(P^*)_\tau^\omega}\left[\int_\tau^T f(c_s, V_s(c))ds + u(c_T)\right] \end{aligned}$$

The first equality follows from $\overline{P} = P$ on $\mathcal{F}_\tau$ and properties of regular conditionals (see Yong and Zhou 1999, Propositions 1.9, 1.10). Moreover, the preceding is true for any $P \in \mathcal{P}$. ■

References

Anderson, E., L. P. Hansen, and T. J. Sargent. 2003. A quartet of semigroups for model specification, robustness, prices of risk and model detection. *Journal of European Economic Association* 1:68–123.

Avellaneda, M., A. Levy, and A. Paras. 1995. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance* 2:73–88.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron. 2011. Volatility, the macroeconomy, and asset prices. Working Paper, Duke University.

Bansal, R., D. Kiku, and A. Yaron. 2012. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Rev*iew 1:183–221.

Barone-Adesi, G., L. Mancini, and H. Shefrin. 2012. Behavioral finance and the pricing kernel puzzle: Estimating risk aversion, optimism and overconfidence. Working Paper, Swiss Finance Institute.

Beeler, J., and J. Y. Campbell. 2012. The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Rev*iew 1:141–82.

Billingsley, P. 1999. *Convergence of probability measures*, 2nd ed. New York: John Wiley.

Bloom, N. 2009. The impact of uncertainty shocks. *Econometrica* 77:623–85.

Bollerslev, T., N. Sizova, and G. Tauchen. 2012. Volatility in equilbrium: Asymmetries and dynamic dependencies. *Review of Finance* 16:31–80.

Campbell, J. Y., S. Giglio, C. Polk, and R. Turkey. 2012. An intertemporal CAPM with stochastic volatility. Working Paper, Harvard University.

Carr, P., and R. Lee. 2009. Volatility derivatives. *Annual Review of Financial Economics* 1:319–39.

Chen, Z., and L. G. Epstein. 2002. Ambiguity, risk, and asset returns in continuous time. *Econometrica* 70: 1403–43.

Cont, R. 2006. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance* 16:519–47.

Cox, J. C., and C. F. Huang. 1989. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory* 39:33–83.

Denis, L., M. Hu, and S. Peng. 2011. Function spaces and capacity related to a sublinear expectation: Application to G-Brownian motion paths. *Potential Analysis* 34:139–61.

Denis, L., and C. Martini. 2006. A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *Annals of Applied Probability* 16:827–52.

Dolinsky, Y., M. Nutz, and H. M. Soner. 2012. Weak approximation of G-expectations. *Stochastic Processes and Applications* 122:664–75.

Drechsler, I. Forthcoming. Uncertainty, time-varying fear, and asset prices. *Journal of Finance*.

Duffie, D. 1996. *Dynamic asset pricing theory*, 2nd ed. Princeton, NJ: Princeton University Press.

Duffie, D., and L. G. Epstein. 1992a. Asset pricing with stochastic differential utility. *Review of Financial Studies* 5:411–36.

———. 1992b. Stochastic differential utility. *Econometrica* 60:353–94 (Appendix C with C. Skiadas).

Duffie, D., and C. Skiadas. 1994. Continuous-time security pricing: A utility gradient approach. *Journal of Mathematical Economics* 23:107–31.

Epstein, L. G., and S. Ji. 2013. Ambiguous volatility, possibility and utility in continuous time. Working Paper, Boston University.

Epstein, L. G., and M. Schneider. 2003. Recursive multiple priors. *Journal of Economic Theory* 113:1–31.

———. 2008. Ambiguity, information quality and asset pricing. *Journal of Finance* 63:197–228.

———. 2010. Ambiguity and asset markets. *Annual Review of Financial Economics* 2:315–46.

Eraker, B., and I. Shaliastovich. 2008. An equilibrium guide to designing affine pricing models. *Mathematical Finance* 18:519–43.

Fernandez-Villaverde, J., P. Guerron-Quintana, J. F. Rubio, and M. Uribe. 2010. Risk matters: The real effects of volatility shocks. *American Economic Review* 101:2530–61.

Follmer, H. 1981. Calcul d'Itô sans probabilities. *Seminar on Probability XV, Lecture Notes in Mathematics* 850:143–50.

Gilboa, I., and D. Schmeidler. 1989. Maxmin expected utility with non-unique priors. *Journal of Mathematical Economics* 18:141–53.

Heston, S. L. 1993. A closed-form solution for options with stochastic volatility with application to bond and currency options. *Review of Financial Studies* 6:293–326.

Hu, M., and S. Ji. 2010. Explicit solutions of linear BSDEs under G-expectations and applications. Working Paper, Shandong University.

Hull, J., and A. White. 1987. The pricing of options on assets with stochastic volatilities. *Journal of Finance* 3:281–300.

Illeditsch, P. K. 2011. Ambiguous information, portfolio inertia and excess volatility. *Journal of Finance* 66: 2213–47.

Ilut, C., and M. Schneider. 2011. Ambiguous business cycles. Working Paper, Duke University.

Karandikar, R. 1995. On pathwise stochastic integration. *Stochastic Processes and Their Applications* 57: 11–18.

Karatzas, I., and S. Shreve. 1991. *Brownian motion and stochastic calculus*. New York: Springer-Verlag.

Levy, A., M. Avellaneda, and A. Paras. 1998. A new approach for pricing derivative securities in markets with uncertain volatilities: A "case study" on the trinomial tree. Working Paper, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=5342.

Li, X., and S. Peng. 2011. Stopping times and related Itô's calculus with G-Brownian motion. *Stochastic Processes and Their Applications* 121:1492–508.

Lo, A., and M. T. Mueller. 2010. WARNING: Physics envy may be hazardous to your wealth! Working Paper, MIT.

Lyons, T. 1995. Uncertain volatility and the risk-free synthesis of derivatives. *Applied Mathematical Finance* 2:117–33.

Mandelkern, M. 1990. On the uniform continuity of Tietze extensions. *Archiv der Mathematik* 55:387–88.

Nelson, D. B., and K. Ramaswamy. 1990. Simple binomial processes as diffusion approximations in financial models. *Review of Financial Studies* 3:393–430.

Oksendal, B. 2005. *Stochastic differential equations*. Berlin: Springer-Verlag.

Peng, S. 2006. G–expectation, G–Brownian motion and related stochastic calculus of Itô's type. In *Stochastic Analysis and Applications*, *The Abel Symposium 2005*, ed. F. E. Benth, G. Di Nunno, T. Lindstrom, B. Oksendal, and T. Zhang. Berlin: Springer-Verlag, 541–567.

———. 2008. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Processes and Their Applications* 118:2223–53.

———. 2010. Nonlinear expectations and stochastic calculus under uncertainty. Working Paper, Shandong University, http://arxiv.org/abs/1002.4546.

Skiadas, C. 2003. Robust control and recursive utility. *Finance and Stochastics* 7:475–89.

———. 2008. Dynamic portfolio theory and risk aversion. In *Handbooks of Operations Research and Management Science: Financial Engineering*, vol. 15, ed. J. R. Birge and V. Linetsky. Amsterdam: North-Holland, 789–843.

———. 2011. Smooth ambiguity aversion toward small risks and continuous-time recursive utility. Working Paper, Northwestern University.

Soner, M., N. Touzi, and J. Zhang. 2011a. Martingale representation theorem under G-expectation. *Stochastic Processes and Their Applications* 121:265–87.

———. 2011b. Quasisure stochastic analysis through aggregation. *Electronic Journal of Probability* 16:1844–79.

———. 2012. Wellposedness of second order backward SDEs. *Probability Theory and Related Fields* 153:149–90.

Song, Y. 2011. Some properties of G-evaluation and its applications to G-martingale decomposition. *Science China-Mathematics* 54:287–300.

Vorbrink, J. 2010. Financial markets with volatility uncertainty. Working Paper, Bielefeld University, http://arxiv.org/abs/1012.1535.

Willard, G. A., and P. H. Dybvig. 1999. Empty promises and arbitrage. *Review of Financial Studies* 12:807–34.

Yong, J., and X. Zhou. 1999. *Stochastic controls*. New York: Springer.

Yuan, C. 2011. The construction of G-Brownian motion and relative financial application. M.A. dissertation, Shandong University.