

AMBIGUOUS VOLATILITY, POSSIBILITY AND UTILITY IN CONTINUOUS TIME*

Larry G. Epstein Shaolin Ji

September 17, 2013

Abstract

This paper formulates a model of utility for a continuous time framework that captures the decision-maker's concern with ambiguity about both the drift and volatility of the driving process. At a technical level, the analysis requires a significant departure from existing continuous time modeling because it cannot be done within a probability space framework. This is because ambiguity about volatility leads invariably to a set of nonequivalent priors, that is, to priors that disagree about which scenarios are possible.

Key words: ambiguity, recursive utility, G-Brownian motion, undominated measures, quasisure analysis, robust stochastic volatility

*Department of Economics, Boston University, lepstein@bu.edu and Institute of Financial Study, Shandong University, jsl@sdu.edu.cn. We gratefully acknowledge the financial support of the National Science Foundation (awards SES-0917740 and 1216339), National Natural Science Foundation of China (awards No. 11171187 and No. 10921101), Programme of Introducing Talents of Discipline to Universities of China (award No.B12023), Program for New Century Excellent Talents in University of China. We have benefited also from the comments of a referee and from discussions with Shige Peng, Zengjing Chen, Mingshang Hu, Jin Ma, Guihai Zhao and especially Jianfeng Zhang. Epstein is grateful also for the generous hospitality of CIRANO where some of this work was completed during an extended visit. The original version of this paper, first posted March 5, 2011, contained applications to asset pricing which are now contained in Epstein and Ji [8]. Marcel Nutz pointed out an error in a previous version.

1. Introduction

This paper formulates a model of utility for a continuous time framework that captures the decision-maker's concern with ambiguity or model uncertainty. The paper's novelty lies in the range of model uncertainty that is accommodated. Specifically, aversion to ambiguity about both drift and volatility is captured. At a technical level, the analysis requires a significant departure from existing continuous time modeling because it cannot be done within a probability space framework. This is because ambiguity about volatility leads invariably to an undominated set of priors. In fact, priors are typically nonequivalent (not mutually absolutely continuous) - they disagree about which scenarios are possible.

The model of utility is a continuous time version of multiple priors (or maxmin) utility formulated by Gilboa and Schmeidler [13] for a static setting. Related continuous time models are provided by Chen and Epstein [3] and also Hansen, Sargent and coauthors (see Anderson et al. [1], for example).¹ In these papers, ambiguity is modeled so as to retain the property that all priors are equivalent. This universal restriction is driven 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 volatility or in which future events are possible.

Economic motivation for our model, some applications to asset pricing, and also an informal intuitive outline of the construction of utility (without proofs), are provided in a companion paper Epstein and Ji [8].² The present paper provides a mathematically rigorous treatment of utility. Thus, for example, it provides the foundations for an equilibrium analysis of asset markets in the presence of ambiguous volatility. The reader is referred to [8] for further discussion of the

¹The discrete time counterpart of the former is axiomatized in Epstein and Schneider [9].

²Economic motivation is provided in part by 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 often unwarranted and implausible.

framework and its economic rationale.³

The challenge in formulating the model is that it cannot be done within a probability space framework.⁴ Typically, the ambient framework is a probability space (Ω, P_0) , it is assumed that $B = (B_t)$ is a Brownian motion under P_0 , and importantly, P_0 is used to define null events. Thus random variables and stochastic processes are defined only up to the P_0 -almost sure qualification and P_0 is an essential part of the definition of all formal domains. However, ambiguity about volatility implies that nullity (or possibility) cannot be defined by any single probability measure. This is easily illustrated. Let B be a Brownian motion under P_0 and denote by $P^\underline{\sigma}$ and $P^\bar{\sigma}$ the probability distributions over continuous paths induced by P_0 and the two processes $(\underline{\sigma}B_t)$ and $(\bar{\sigma}B_t)$, where, for simplicity, $\underline{\sigma}$ and $\bar{\sigma}$ are constants. Then $P^\underline{\sigma}$ and $P^\bar{\sigma}$ are mutually singular (and hence not equivalent) because

$$P^\underline{\sigma}(\{\langle B \rangle_T = \underline{\sigma}^2 T\}) = 1 = P^\bar{\sigma}(\{\langle B \rangle_T = \bar{\sigma}^2 T\}). \quad (1.1)$$

To overcome the resulting difficulty, we define appropriate domains of stochastic processes by using the entire set of priors to define the almost sure qualification. For example, equality of two random variables will be taken to mean almost sure equality for every prior in the decision maker's set of priors. This so-called *quasi-sure stochastic analysis* was developed by Denis and Martini [6]. See also Soner et al. [34] and Denis et al. [4] for elaboration on why a probability space framework is inadequate and for a comparison of quasisure analysis with related approaches.⁵ Prominent among the latter is the seminal contribution of G-expectation due to Peng [29, 31, 32, 33], wherein a nonlinear expectations operator is defined by a set of undominated measures. We combine elements of both quasisure analysis and G-expectation. Conditioning, or updating, is obviously a crucial ingredient in modeling dynamic preferences. In this respect we adapt the approach in Soner et

³For example, an occasional reaction is that ambiguity about volatility is implausible because one can estimate the law of motion for volatility (of asset prices, for example) extremely well. However, this perspective presumes a stationary environment and relies on a tight connection between the past and future that we relax. For similar reasons we reject the suggestion that one can discriminate readily between nonequivalent laws of motion and hence that there is no loss of empirical relevance in restricting priors to be equivalent. See the companion paper for elaboration on these and other matters of interpretation.

⁴Where the set of priors is finite (or countable), a dominating measure is easily constructed. However, the set of priors in our model is not countable, and a suitable probability space framework does not exist outside of the extreme case where there is no ambiguity about volatility.

⁵Related developments are provided by Bion-Nadal et al. [2] and Soner et al. [35, 36].

al. [37] and Nutz [25] to conditioning undominated sets of measures.⁶ However, these analyses do not apply off-the-shelf because, for example, they permit ambiguity about volatility but not about drift. In particular, accommodating both kinds of ambiguity necessitates a novel construction of the set of priors.⁷

Besides those already mentioned, there are only a few relevant papers in the literature on continuous time utility theory. Denis and Kervarec [5] formulate multiple priors utility functions and study optimization in a continuous-time framework; they do not assume equivalence of measures but they restrict attention to the case where only terminal consumption matters and where all decisions are made at a single ex ante stage. Bion-Nadal and Kervarec [2] study risk measures (which can be viewed as a form of utility function) in the absence of certainty about what is possible.

Section 2 is the heart of the paper and presents the model of utility, beginning with the construction of the set of priors and the definition of (nonlinear) conditional expectation. Proofs are collected in appendices.

2. Utility

2.1. Preliminaries

Time t varies over the finite horizon $[0, T]$. Paths or trajectories of the 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

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

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 . It is a reference measure only in the mathematical sense of facilitating the description of the individual's set of priors; but the latter need not contain P_0 .

Define also the set of paths over the time interval $[0, t]$:

$${}^t\Omega = \{{}^t\omega = ({}^t\omega_s) \in C^d([0, t]) : {}^t\omega_0 = 0\}.$$

⁶Peng [29, 30] provides a related approach to conditioning.

⁷See Section 2.3 for comparison with a related paper (Nutz [26]) written after the first version of this paper was circulated.

Identify ${}^t\Omega$ with a subspace of Ω by identifying any ${}^t\omega$ with the function on $[0, T]$ that is constant at level ${}^t\omega_t$ on $[t, T]$. Note that the filtration \mathcal{F}_t is the Borel σ -field on ${}^t\Omega$. (Below, for any topological space we always adopt the Borel σ -field even where not mentioned explicitly.)

Consumption processes c take values in C , a convex subset of \mathbb{R}^ℓ . The domain of consumption processes is denoted D . Because we are interested in describing dynamic choice, we need to specify not only ex ante utility over D , but a suitable process of (conditional) utility functions.

The key is construction of the set of priors. The primitive is the individual's hypotheses about drift and volatility. These are used to specify the set of priors. Conditioning is treated next. Finally, these components are applied to define a recursive process of utility functions.

2.2. Drift and Volatility Hypotheses

Before moving to the general setup, we outline briefly a special case where $d = 1$ and there is ambiguity only about volatility. Accordingly, suppose that (speaking informally) the individual is certain that the driving process $B = (B_t)$ is a martingale, but that its volatility is known only up to the interval $[\underline{\sigma}, \bar{\sigma}]$. In particular, she does not take a stand on any particular parametric model of volatility dynamics.

To be more precise about the meaning of volatility, recall that the quadratic variation process of (B_t) is defined by

$$\langle B \rangle_t(\omega) = \lim_{\Delta t_k \rightarrow 0} \sum_{t_k \leq t} |B_{t_{k+1}}(\omega) - B_{t_k}(\omega)|^2 \quad (2.1)$$

where $0 = t_1 < \dots < t_n = t$ and $\Delta t_k = t_{k+1} - t_k$. (By Follmer [11] and Karandikar [19], the above limit exists almost surely for every measure that makes B a martingale, thus giving a universal stochastic process $\langle B \rangle$; because the individual is certain that B is a martingale, this limited universality is all we need.) 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

$$\underline{\sigma}^2 t \leq \langle B \rangle_t \leq \bar{\sigma}^2 t. \quad (2.2)$$

The model that follows is much more general. Importantly, the interval $[\underline{\sigma}, \bar{\sigma}]$ 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, the model admits multidimensional driving processes ($d > 1$) and also ambiguity about drift, thus relaxing the assumption of certainty that B is a martingale.

In general, the individual is not certain that the driving process has zero drift and/or unit variance. Accordingly, she entertains a range of alternative hypotheses $X^\theta = (X_t^\theta)$ parametrized by $\theta = (\theta_t)$. 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 $\sigma = (\sigma_t)$. The primitive is the process of correspondences (Θ_t) , where, for each t ,

$$\Theta_t : \Omega \rightsquigarrow \mathbb{R}^d \times \mathbb{R}^{d \times d}.$$

Roughly, $\Theta_t(\omega)$ gives the set of admissible drift and volatility pairs at t along the trajectory ω . The idea is that each θ parametrizes the driving process $X^\theta = (X_t^\theta)$ given by the unique solution to the following stochastic differential equation (SDE) under P_0 :

$$dX_t^\theta = \mu_t(X^\theta)dt + \sigma_t(X^\theta)dB_t, \quad X_0^\theta = 0, \quad t \in [0, T]. \quad (2.3)$$

We assume that only θ 's for which a unique strong solution exists are adopted as hypotheses. Therefore, denote by Θ^{SDE} the set of all processes θ that ensure a unique strong solution X^θ to the SDE,⁸ and define the set Θ of admissible drift and volatility processes by

$$\Theta = \{ \theta \in \Theta^{SDE} : \theta_t(\omega) \in \Theta_t(\omega) \text{ for all } (t, \omega) \in [0, T] \times \Omega \}. \quad (2.4)$$

We impose the following technical regularity conditions on (Θ_t) :

- (i) *Measurability*: The correspondence $(t, \omega) \mapsto \Theta_t(\omega)$ on $[0, s] \times \Omega$ is $\mathcal{B}([0, s]) \times \mathcal{F}_s$ -measurable for every $0 < s \leq T$.
- (ii) *Uniform Boundedness*: There is a compact subset \mathcal{K} in $\mathbb{R}^d \times \mathbb{R}^{d \times d}$ such that $\Theta_t : \Omega \rightsquigarrow \mathcal{K}$ each t .
- (iii) *Compact-Convex*: Each Θ_t is compact-valued and convex-valued.
- (iv) *Uniform Nondegeneracy*: There exists \hat{a} , a $d \times d$ real-valued positive definite matrix, such that for every t and ω , if $(\mu_t, \sigma_t) \in \Theta_t(\omega)$, then $\sigma_t \sigma_t^\top \geq \hat{a}$.

⁸By uniqueness we mean that $P_0(\{\sup_{0 \leq t \leq T} |X_t^\theta - X_t^{\theta'}| > 0\}) = 0$ for any other strong solution $(X_t^{\theta'})$. A Lipschitz condition and boundedness are sufficient for existence of a unique solution, but these properties are not necessary and a Lipschitz condition is not imposed.

- (v) *Uniform Continuity*: The process (Θ_t) is uniformly continuous in the sense defined in Appendix A.
- (vi) *Uniform Interiority*: There exists $\delta > 0$ such that $ri^\delta\Theta_t(\omega) \neq \emptyset$ for all t and ω , where $ri^\delta\Theta_t(\omega)$ is the δ -relative interior of $\Theta_t(\omega)$. (For any $D \subseteq (\mathbb{R}^d \times \mathbb{R}^{d \times d})$ and $\delta > 0$, $ri^\delta D \equiv \{x \in D : (x + B_\delta(x)) \cap (\text{aff } D) \subset D\}$, where $\text{aff } D$ is the affine hull of D and $B_\delta(x)$ denotes the open ball of radius δ .)
- (vii) *Uniform Affine Hull*: The affine hulls of $\Theta_{t'}(\omega')$ and $\Theta_t(\omega)$ are the same for every (t', ω') and (t, ω) in $[0, T] \times \Omega$.

Conditions (i)-(iii) parallel assumptions made by Chen and Epstein [3]. A form of Nondegeneracy is standard in financial economics. The remaining conditions are adapted from Nutz [25] and are imposed in order to accommodate ambiguity in volatility. The major differences from Nutz' assumptions are in (vi) and (vii). Translated into our setting, he assumes that $int^\delta\Theta_t(\omega) \neq \emptyset$, where, for any D , $int^\delta\Theta_t(\omega) = \{x \in D : (x + B_\delta(x)) \subset D\}$. By weakening his requirement to deal with relative interiors, we are able to broaden the scope of the model in important ways (see the first three examples below). Because each $\Theta_t(\omega)$ is convex, if it also has nonempty interior then its affine hull is all of $\mathbb{R}^d \times \mathbb{R}^{d \times d}$. Then $ri^\delta\Theta_t(\omega) = int^\delta\Theta_t(\omega) \neq \emptyset$ and also (vii) is implied. In this sense, (vi)-(vii) are jointly weaker than assuming $int^\delta\Theta_t(\omega) \neq \emptyset$.

We illustrate the scope of the model through some examples.

Example 2.1 (Ambiguous drift). *If $\Theta_t(\omega) \subset \mathbb{R}^d \times \{\sigma_t(\omega)\}$ for every t and ω , for some volatility process σ , then there is ambiguity only about drift. If $d = 1$, it is modeled by the random and time varying interval $[\underline{\mu}_t, \bar{\mu}_t]$. The regularity conditions above for (Θ_t) are satisfied if: $\sigma_t^2 \geq a > 0$ and $\bar{\mu}_t - \underline{\mu}_t > 0$ everywhere, and if $\bar{\mu}_t$ and $\underline{\mu}_t$ are continuous in ω uniformly in t . This special case corresponds to the Chen and Epstein [3] model.*

Example 2.2 (Ambiguous volatility). *If $\Theta_t(\omega) \subset \{\mu_t(\omega)\} \times \mathbb{R}^{d \times d}$ for every t and ω , for some drift process μ , then there is ambiguity only about volatility. If $d = 1$, it is modeled by the random and time varying interval $[\underline{\sigma}_t, \bar{\sigma}_t]$. The regularity conditions for (Θ_t) are satisfied if: $\bar{\sigma}_t > \underline{\sigma}_t \geq a > 0$ everywhere, and if $\bar{\sigma}_t$ and $\underline{\sigma}_t$ are continuous in ω uniformly in t .*

A generalization is important. Allow $d \geq 1$ and let $\mu_t = 0$. Then (speaking informally) there is certainty that B is a martingale in spite of uncertainty about

the true probability law. Volatility (σ_t) is a process of $d \times d$ matrices. Let the admissible volatility processes (σ_t) be those satisfying $\sigma_t \in \Gamma$, where Γ is any compact convex subset of $\mathbb{R}^{d \times d}$ such that, for all σ in Γ , $\sigma\sigma^\top \geq \hat{a}$ for some positive definite matrix \hat{a} .⁹ This specification is essentially equivalent to Peng's [31] notion of G -Brownian motion.¹⁰

Example 2.3 (Robust stochastic volatility). *This is a special case of the preceding example but we describe it separately in order to highlight the connection of our model to the stochastic volatility literature. 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) satisfying regularity conditions of the sort given above. The specification of a single process for volatility indicates the individual'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; for instance, they might conform to the Hull and White [17] and Heston [15] parametric forms respectively.¹¹ 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 a context in which stochastic volatility models are used heavily) differ significantly. Faced with these two models, the individual might place probability $\frac{1}{2}$ on each being the true model. But why should she be certain that either one is true? Both (σ_t^1) and (σ_t^2) may fit data well to some approximation, but other approximating models may do as well. An intermediate model such as $(\frac{1}{2}\sigma_t^1 + \frac{1}{2}\sigma_t^2)$ is one alternative, but there are many others that "lie between" (σ_t^1) and (σ_t^2) and that plausibly should be taken into account. Accordingly, (assuming $d = 1$), let*

$$\underline{\sigma}_t(\omega) = \min\{\sigma_t^1(\omega), \sigma_t^2(\omega)\} \text{ and } \bar{\sigma}_t(\omega) = \max\{\sigma_t^1(\omega), \sigma_t^2(\omega)\},$$

and admit all volatility processes with values lying in the interval $[\underline{\sigma}_t(\omega), \bar{\sigma}_t(\omega)]$ for every ω .

⁹Given symmetric matrices A' and A , $A' \geq A$ if $A' - A$ is positive semidefinite.

¹⁰Peng provides generalizations of Itô's Lemma and Itô integration appropriate for G -Brownian motion that we exploit in our companion paper in deriving asset pricing implications.

¹¹However, these functional forms would lead to violation of the Uniform Boundedness regularity condition (ii).

Example 2.4 (Joint ambiguity). *The model is flexible in the way it relates ambiguity about drift and ambiguity about volatility. For example, a form of independence is modeled if (taking $d = 1$)*

$$\Theta_t(\omega) = [\underline{\mu}_t(\omega), \bar{\mu}_t(\omega)] \times [\underline{\sigma}_t(\omega), \bar{\sigma}_t(\omega)], \quad (2.5)$$

the Cartesian product of the intervals described in the preceding two examples.

An alternative hypothesis is that drift and volatility are thought to move together. This is captured by specifying, for example,

$$\Theta_t(\omega) = \{(\mu, \sigma) \in \mathbb{R}^2 : \mu = \mu_{\min} + z, \sigma^2 = \sigma_{\min}^2 + 2z/\gamma, 0 \leq z \leq \bar{z}_t(\omega)\}, \quad (2.6)$$

where μ_{\min} , σ_{\min}^2 and $\gamma > 0$ are fixed and known parameters. The regularity conditions for (Θ_t) are satisfied if \bar{z}_t is positive everywhere and continuous in ω uniformly in t . This specification is adapted from Epstein and Schneider [10].

Example 2.5 (Markovian ambiguity). *Assume that (Θ_t) satisfies:*

$$\omega'_t = \omega_t \implies \Theta_t(\omega') = \Theta_t(\omega).$$

Then ambiguity depends only on the current state and not on history. Note, however, that according to (2.4), the drift and volatility processes deemed possible are not necessarily Markovian - θ_t can depend on the complete history at any time. Thus the individual is not certain that the driving process is Markovian, but the set of processes that she considers possible at any given time is independent of history beyond the prevailing state.

2.3. Priors, expectation and conditional expectation

We proceed to translate the set Θ of hypotheses about drift and volatility into a set of priors. Each θ induces (via P_0) a probability measure P^θ on (Ω, \mathcal{F}_T) given by

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

Therefore, we arrive at the set of priors \mathcal{P}^Θ given by

$$\mathcal{P}^\Theta = \{P^\theta : \theta \in \Theta\}. \quad (2.7)$$

Fix Θ and denote the set of priors \mathcal{P}^Θ simply by \mathcal{P} . This is the set of priors used, as in the Gilboa-Schmeidler model, to define utility and to describe choice between consumption processes.¹²

¹²The set \mathcal{P} is relatively compact in the topology induced by bounded continuous functions (this is a direct consequence of Gihman and Skorohod [12, Theorem 3.10]).

Remark 1. *If all alternative hypotheses display unit variance, then ambiguity is limited to the drift as in the Chen-Epstein model, and one can show that, as in [3], measures in \mathcal{P} are pairwise equivalent. At the other extreme, if they all display zero drift, then ambiguity is limited to volatility and many measures in \mathcal{P} are mutually singular. Nonequivalence of priors prevails in the general model when both drift and volatility are ambiguous.*

Given \mathcal{P} , we define (nonlinear) expectation as follows. For a random variable ξ on (Ω, \mathcal{F}_T) , if $\sup_{P \in \mathcal{P}} E_P \xi < \infty$ define

$$\hat{E}\xi = \sup_{P \in \mathcal{P}} E_P \xi. \quad (2.8)$$

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

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

The crucial remaining ingredient of the model, and the focus of most of the work in the appendices, is conditioning. A naive approach to defining conditional expectation would be to use the standard conditional expectation $E_P[\xi | \mathcal{F}_t]$ 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 $E_P[\xi | \mathcal{F}_t]$ is well defined only P -almost surely, while to be a meaningful object for analysis, a random variable must be well defined from the perspective of every measure in \mathcal{P} . In the following, we say that a property holds *quasisurely* (*q.s.* for short) if it holds P -*a.s.* for every $P \in \mathcal{P}$.¹³ In other words, and speaking very informally, conditional beliefs must be defined at every node deemed possible by some measure in \mathcal{P} . The economic rationale is that even if $P(A) = 0$, for some $A \in \mathcal{F}_t$ and $t > 0$, if also $Q(A) > 0$ for some other prior in \mathcal{P} , then ex ante the individual does not totally dismiss the possibility of A occurring when she formulates consumption plans: if she is guided by the worst case scenario, then $\min_{P' \in \mathcal{P}} P'(\Omega \setminus A) < 1$ implies that she would reject a

¹³Throughout, when Z is a random variable, $Z \geq 0$ quasisurely means that the inequality is valid P -*a.s.* for every P in \mathcal{P} . If $Z = (Z_t)$ is a process, by the statement “ $Z_t \geq 0$ for every t quasisurely (*q.s.*)” we mean that for every t there exists $G_t \subset \Omega$ such that $Z_t(\omega) \geq 0$ for all $\omega \in G_t$ and $P(G_t) = 1$ for all P in \mathcal{P} . If $Z_t = 0$ for every t quasisurely, then $Z = 0$ in $M^2(0, T)$ (because $\hat{E}[\int_0^T |Z_t|^2 dt] \leq \int_0^T \hat{E}[|Z_t|^2] dt$), but the converse is not valid in general.

bet against A that promised a sufficiently poor prize (low consumption stream) if A occurs. Therefore, a model of dynamic choice by a sophisticated and forward-looking individual should specify her consumption plan contingent on arriving at (A, t) .

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. It resembles also the approach in the discrete time model in Epstein and Schneider [9], 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.

To proceed, recall the construction of the set of priors through (2.3) and the set Θ of admissible drift and volatility processes. If $\theta = (\theta_s)$ is a conceivable scenario ex ante, then $(\theta_s(t\omega, \cdot))_{t \leq s \leq T}$ 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 a conceivable scenario ex post conditionally on (t, ω) , thus ruling out surprises or unanticipated changes in outlook. Accordingly, $X^{\theta, t, \omega} = (X_s^{\theta, t, \omega})_{t \leq s \leq T}$ is a conceivable conditional scenario for the driving process if it solves the following SDE under P_0 :

$$\begin{cases} dX_s^{\theta, t, \omega} &= \mu_s(X_s^{\theta, t, \omega})ds + \sigma_s(X_s^{\theta, t, \omega})dB_s, \quad t \leq s \leq T \\ X_s^{\theta, t, \omega} &= \omega_s, \quad 0 \leq s \leq t. \end{cases} \quad (2.9)$$

The solution $X^{\theta, t, \omega}$ induces a probability measure $P_t^{\theta, \omega} \in \Delta(\Omega)$, denoted simply by P_t^ω with θ suppressed when it is understood that $P = P^\theta$. For each P in \mathcal{P} , the measure $P_t^\omega \in \Delta(\Omega)$ is defined for every t and ω , and, importantly, it is a version of the regular \mathcal{F}_t -conditional probability of P (see Lemma B.2).

The set of all such conditionals obtained as θ varies over Θ is denoted \mathcal{P}_t^ω , that is,

$$\mathcal{P}_t^\omega = \{P_t^\omega : P \in \mathcal{P}\}. \quad (2.10)$$

We take \mathcal{P}_t^ω to be the individual's set of priors conditional on (t, ω) .¹⁴

¹⁴The evident parallel with the earlier construction of the ex ante set \mathcal{P} can be expressed more formally because the construction via (2.9) can be expressed in terms of a process of correspondences $(\Theta_s^{t, \omega})_{t \leq s \leq T}$, $\Theta_s^{t, \omega} : C^d([t, T]) \rightsquigarrow \mathbb{R}^d \times \mathbb{R}^{d \times d}$, satisfying counterparts of the regularity conditions (i)-(vii) on the time interval $[t, T]$.

The sets of conditionals in (2.10) lead to the following (nonlinear) conditional expectation on $UC_b(\Omega)$, the set of all bounded and uniformly continuous functions on Ω :

$$\hat{E}[\xi | \mathcal{F}_t](\omega) = \sup_{P \in \mathcal{P}_t^\omega} E_P \xi, \text{ for every } \xi \in UC_b(\Omega) \text{ and } (t, \omega) \in [0, T] \times \Omega. \quad (2.11)$$

Conditional expectation is defined thereby on $UC_b(\Omega)$, but this domain is not large enough for our purposes.¹⁵ For example, the conditional expectation of ξ need not be (uniformly) continuous in ω even if ξ is bounded and uniformly continuous, which is an obstacle to dealing with stochastic processes and recursive modeling. (Similarly, if one were to use the space $C_b(\Omega)$ of bounded continuous functions.)

Thus we consider the larger domain $\widehat{L}^2(\Omega)$, the completion of $UC_b(\Omega)$ under the norm $\|\xi\| \equiv (\hat{E}[|\xi|^2])^{\frac{1}{2}}$.¹⁶ Denis et al. [4] show that a random variable ξ defined on Ω lies in $\widehat{L}^2(\Omega)$ if and only if: (i) ξ is *quasicontinuous* - for every $\epsilon > 0$ there exists an open set $G \subset \Omega$ with $P(G) < \epsilon$ for every P in \mathcal{P} such that ξ is continuous on $\Omega \setminus G$; and (ii) ξ is uniformly integrable in the sense that $\lim_{n \rightarrow \infty} \sup_{P \in \mathcal{P}} E^P(|\xi|^2 \mathbf{1}_{\{|\xi| > n\}}) = 0$. This characterization is inspired by Lusin's Theorem for the classical case which implies that when $\mathcal{P} = \{P\}$, then $\widehat{L}^2(\Omega)$ reduces to the familiar space of P -squared integrable random variables. For general \mathcal{P} , $\widehat{L}^2(\Omega)$ is a proper subset of the set of measurable random variables ξ for which $\sup_{P \in \mathcal{P}} E^P(|\xi|^2) < \infty$.¹⁷ However, it is large in the sense of containing many discontinuous random variables; for example, $\widehat{L}^2(\Omega)$ contains every bounded and lower semicontinuous function on Ω (see the proof of Lemma B.8).

Another aspect of $\widehat{L}^2(\Omega)$ warrants emphasis. Two random variables ξ' and ξ are identified in $\widehat{L}^2(\Omega)$ if and only if $\|\xi' - \xi\| = 0$, which means that $\xi' = \xi$ almost surely with respect to P for *every* P in \mathcal{P} . In that case, say that the equality obtains quasisurely and write $\xi' = \xi$ *q.s.* Thus ξ' and ξ are distinguished whenever they differ with positive probability for *some* measure in \mathcal{P} . Accordingly, the space $\widehat{L}^2(\Omega)$ provides a more detailed picture of random variables than does any single measure in \mathcal{P} .

¹⁵The definition is restricted to $UC_b(\Omega)$ in order to ensure the measurability of $\hat{E}[\xi | \mathcal{F}_t](\cdot)$ and proof of a suitable form of the law of iterated expectations. Indeed, Appendix B, specifically (B.2), shows that conditional expectation has a different representation when random variables outside $UC_b(\Omega)$ are considered.

¹⁶It coincides with the completion of $C_b(\Omega)$; see [4].

¹⁷For example, if \mathcal{P} is the set of all Dirac measures with support in Ω , then $\widehat{L}^2(\Omega) = C_b(\Omega)$.

The next theorem (proven in Appendix B) shows that conditional expectation admits a suitably unique and well behaved extension from $UC_b(\Omega)$ to all of $\widehat{L}^2(\Omega)$. Accordingly, the sets \mathcal{P}_t^ω determine the conditional expectation for all random variables considered in the sequel.

Theorem 2.6 (Conditioning). *The mapping $\hat{E}[\cdot | \mathcal{F}_t]$ on $UC_b(\Omega)$ defined in (2.11) can be extended uniquely to a 1-Lipschitz continuous mapping $\hat{E}[\cdot | \mathcal{F}_t] : \widehat{L}^2(\Omega) \rightarrow \widehat{L}^2(\Omega)$, where 1-Lipschitz continuity means that*

$$\| \hat{E}[\xi' | \mathcal{F}_t] - \hat{E}[\xi | \mathcal{F}_t] \|_{\widehat{L}^2} \leq \| \xi' - \xi \|_{\widehat{L}^2} \quad \text{for all } \xi', \xi \in \widehat{L}^2(\Omega).$$

Moreover, the extension satisfies, for all ξ and η in $\widehat{L}^2(\Omega)$ and for all $t \in [0, T]$,

$$\hat{E}[\hat{E}[\xi | \mathcal{F}_t] | \mathcal{F}_s] = \hat{E}[\xi | \mathcal{F}_s], \quad \text{for } 0 \leq s \leq t \leq T, \quad (2.12)$$

and:

- (i) If $\xi \geq \eta$, then $\hat{E}[\xi | \mathcal{F}_t] \geq \hat{E}[\eta | \mathcal{F}_t]$.
- (ii) If ξ is \mathcal{F}_t -measurable, then $\hat{E}[\xi | \mathcal{F}_t] = \xi$.
- (iii) $\hat{E}[\xi | \mathcal{F}_t] + \hat{E}[\eta | \mathcal{F}_t] \geq \hat{E}[\xi + \eta | \mathcal{F}_t]$ with equality if η is \mathcal{F}_t -measurable.
- (iv) $\hat{E}[\eta\xi | \mathcal{F}_t] = \eta^+ \hat{E}[\xi | \mathcal{F}_t] + \eta^- \hat{E}[-\xi | \mathcal{F}_t]$, if η is \mathcal{F}_t -measurable.

The Lipschitz property is familiar from the classical case of a single prior, where it is implied by Jensen's inequality (see the proof of Lemma B.8). The law of iterated expectations (2.12) is intimately tied to dynamic consistency of the preferences discussed below. The nonlinearity expressed in (iii) reflects the non-singleton nature of the set of priors. Other properties have clear interpretations.

Remark 2. *In a paper written after ours was posted and widely circulated, Nutz [26, Theorem 5.1] also constructs a sublinear conditional expectation accommodating ambiguity about both drift and volatility. He induces priors via an SDE of the form*

$$dX_t = \mu_t(X_t, v_t)dt + \sigma_t(X_t, v_t)dB_t, \quad X_0 = x, t \in [0, T],$$

where the control v_t takes values in a fixed nonempty Borel subset U of R^m . Then, when v_t changes, so does $\theta_t = (\mu_t, \sigma_t)$. We induce priors via (2.3) and we require only that (θ_t) falls in the set Θ of processes. A second difference is that we assume directly that θ has the "feedback" form, i.e., θ is a progressively measurable functional of X^θ . Nutz's Assumption 2.1 is more general, but it is difficult to verify. As he points out in Remark 2.2, the usual way to guarantee Assumption 2.1 would be to assume that θ has the "feedback" form—our assumption.

2.4. The definition of utility

Because we turn to consideration of processes, we define $M^{2,0}(0, T)$, 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 $\xi_i \in \widehat{L}^2({}^{t_i}\Omega)$, $0 \leq i \leq N-1$, and $0 = t_0 < \dots < t_N = T$.¹⁸ Roughly, each such η is a step function in random variables from the spaces $\widehat{L}^2({}^{t_i}\Omega)$. For the usual technical reasons, we consider also suitable limits of such processes. Thus define $M^2(0, T)$ to be the completion of $M^{2,0}(0, T)$ under the norm

$$\| \eta \|_{M^2(0, T)} \equiv \left(\hat{E} \left[\int_0^T | \eta_t |^2 dt \right] \right)^{\frac{1}{2}}.$$

Consumption at every time takes values in C , a convex subset of \mathbb{R}_+^d . Consumption processes $c = (c_t)$ lie in D , a subset of $M^2(0, T)$.

For each c in D , we define a utility process $(V_t(c))$, where $V_t(c)$ is the utility of the continuation $(c_s)_{0 \leq s \leq t}$ and $V_0(c)$ is the utility of the entire process c . We often suppress the dependence on c and write simply (V_t) . We define utility following Duffie and Epstein [7]. This is done in order that our model retain the flexibility to partially separate intertemporal substitution from other aspects of preference (here uncertainty, rather than risk by which we mean ‘probabilistic uncertainty’).

Let Θ and $\mathcal{P} = \mathcal{P}^\Theta$ be as above. The other primitive component is the aggregator $f : C \times \mathbb{R}^1 \rightarrow \mathbb{R}^1$. It is assumed to satisfy:

- (i) f is Borel measurable.
- (ii) Uniform Lipschitz for aggregator: There exists a positive constant K such that

$$| f(c, v') - f(c, v) | \leq K | v' - v |, \text{ for all } (c, v', v) \in C \times \mathbb{R}^2.$$

- (iii) $(f(c_t, v))_{0 \leq t \leq T} \in M^2(0, T)$ for each $v \in R$ and $c \in D$.

We define V_t by

$$V_t = -\hat{E} \left[- \int_t^T f(c_s, V_s) ds \mid \mathcal{F}_t \right]. \quad (2.13)$$

¹⁸The space ${}^{t_i}\Omega$ was defined in Section 2.1.

This definition of utility generalizes both Duffie and Epstein [7], where there is no ambiguity, and Chen and Epstein [3], where ambiguity is confined to drift. Formally, they use different sets of priors: a singleton set in the former paper and a suitable set of equivalent priors in the latter paper.

Our main result follows (see Appendix C for a proof).

Theorem 2.7 (Utility). *Let (Θ_t) and f satisfy the above assumptions. Fix $c \in D$. Then:*

- (a) *There exists a unique process (V_t) in $M^2(0, T)$ solving (2.13).*
- (b) *The process (V_t) is the unique solution in $M^2(0, T)$ to $V_T = 0$ and*

$$V_t = -\hat{E}\left[-\int_t^\tau f(c_s, V_s) ds - V_\tau \mid \mathcal{F}_t\right], \quad 0 \leq t < \tau \leq T. \quad (2.14)$$

Part (a) proves that utility is well defined by (2.13). Recursivity is established in (b).

The most commonly used aggregator has the form

$$f(c_t, v) = u(c_t) - \beta v, \quad \beta \geq 0, \quad (2.15)$$

in which case utility admits the closed-form expression

$$V_t = -\hat{E}\left[-\int_t^T u(c_s)e^{-\beta s} ds \mid \mathcal{F}_t\right]. \quad (2.16)$$

More generally, closed form expressions are rare. The following example illustrates the effect of volatility ambiguity.

Example 2.8 (Closed form). *Consider the consumption process c satisfying (under P_0)*

$$d \log c_t = s^\top \sigma_t dB_t, \quad c_0 > 0 \text{ given}, \quad (2.17)$$

where s is constant and the volatility matrix σ_t is restricted only to lie in the compact and convex set Γ , as in Example 2.2, corresponding to Peng's [31] notion of G -Brownian motion. Utility is defined by the standard aggregator,

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

where the felicity function u is given by

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

Then the conditional utilities $V_t(c)$ can be expressed in closed form. To do so, define $\underline{\sigma}$ and $\bar{\sigma}$ as the respective solutions to

$$\min_{\sigma \in \Gamma} \text{tr}(\sigma \sigma^\top s s^\top) \quad \text{and} \quad \max_{\sigma \in \Gamma} \text{tr}(\sigma \sigma^\top s s^\top). \quad (2.18)$$

(If $d = 1$, then Γ is a compact interval and $\underline{\sigma}$ and $\bar{\sigma}$ are its left and right endpoints.)

Let P^* be the measure on Ω induced by P_0 and X^* , where

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

define P^{**} similarly using $\underline{\sigma}$ and X^{**} . They are worst-case (or minimizing) measures in that¹⁹

$$V_0(c) = \begin{cases} E^{P^*} \left[\int_0^T \alpha^{-1} (c_\tau)^\alpha e^{-\beta\tau} d\tau \right] & \text{if } \alpha < 0 \\ E^{P^{**}} \left[\int_0^T \alpha^{-1} (c_\tau)^\alpha e^{-\beta\tau} d\tau \right] & \text{if } \alpha > 0 \end{cases} \quad (2.19)$$

This follows from Levy et al. [21] and Peng [33], because $u(c_t) = \alpha^{-1} \exp(\alpha \log c_t)$ and $x \mapsto e^{\alpha x} / \alpha$ is concave if $\alpha < 0$ and convex if $\alpha > 0$. From (2.17), almost surely with respect to $P^{(\sigma_\tau)}$,

$$\alpha^{-1} c_t^\alpha = \alpha^{-1} c_0^\alpha \exp \left\{ \alpha \int_0^t s^\top \sigma_\tau dB_\tau \right\}.$$

It follows that utility can be computed as if the volatility σ_t were constant and equal to $\bar{\sigma}$ (if $\alpha < 0$) or $\underline{\sigma}$ (if $\alpha > 0$).

For $t > 0$, employ the regular conditionals of P^* and P^{**} , which have a simple form. For example, following (2.9), for every (t, ω) , $(P^*)^\omega_t$ is the measure on Ω induced by the SDE

$$\begin{cases} dX_\tau = \bar{\sigma} dB_\tau, & t \leq \tau \leq T \\ X_\tau = \omega_\tau, & 0 \leq \tau \leq t \end{cases}$$

Thus under $(P^*)^\omega_t$, $B_\tau - B_t$ is $N(0, \bar{\sigma} \bar{\sigma}^\top (\tau - t))$ for $t \leq \tau \leq T$. Further, $(P^*)^\omega_t$ is the worst case measure in \mathcal{P}_t^ω if $\alpha < 0$; similarly, $(P^{**})^\omega_t$ is the worst case measure in \mathcal{P}_t^ω if $\alpha > 0$. Repeat the argument used above for $t = 0$ to obtain

$$V_t(c) = \begin{cases} E^{(P^*)^\omega_t} \left[\int_t^T \alpha^{-1} (c_\tau)^\alpha e^{-\beta(\tau-t)} d\tau \right] & \text{if } \alpha < 0 \\ E^{(P^{**})^\omega_t} \left[\int_t^T \alpha^{-1} (c_\tau)^\alpha e^{-\beta(\tau-t)} d\tau \right] & \text{if } \alpha > 0 \end{cases}$$

¹⁹That 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.

and, by computing the expectations, that

$$V_t(c) = \begin{cases} \alpha^{-1} c_t^\alpha \bar{\gamma}^{-1} (1 - e^{-\bar{\gamma}(T-t)}) & \text{if } \alpha < 0 \\ \alpha^{-1} c_t^\alpha \underline{\gamma}^{-1} (1 - e^{-\underline{\gamma}(T-t)}) & \text{if } \alpha > 0 \end{cases}$$

where

$$\bar{\gamma} \triangleq \beta - \frac{1}{2} \alpha^2 s^\top \overline{\sigma \sigma}^\top s, \quad \underline{\gamma} \triangleq \beta - \frac{1}{2} \alpha^2 s^\top \underline{\sigma \sigma}^\top s.$$

Utility has a range of natural properties. Most noteworthy is that the process (V_t) satisfies the recursive relation (2.14). However, though such recursivity is typically thought to imply dynamic consistency, the nonequivalence of priors complicates matters. The noted recursivity implies the following weak form of dynamic consistency: For any $0 < t < T$, and any two consumption processes c' and c that coincide on $[0, t]$,

$$[V_t(c') \geq V_t(c) \quad q.s.] \implies V_0(c') \geq V_0(c).$$

Typically, (see Duffie and Epstein [7, p. 373] for example), dynamic consistency is defined so as to deal also with strict rankings, that is, if also $V_t(c') > V_t(c)$ on a “non-negligible” 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 t . The issue is how to specify “non-negligible”. When all priors are equivalent, then positive probability according to any single prior is the natural specification. In the absence of equivalence a similarly natural specification is unclear. In particular, as illustrated in [8], it is possible that c' and c be indifferent ex ante and yet that: (i) $c' \geq c$ on $[0, T]$ and they coincide on $[0, t]$; (ii) there exists $t > 0$, an event $N_t \in \mathcal{F}_t$, with $P(N_t) > 0$ for some $P \in \mathcal{P}$, such that $c'_\tau > c_\tau$ for $t < \tau \leq T$ and $\omega \in N_\tau$. Then monotonicity of preference would imply that c' be weakly preferable to c at t and strictly preferable conditionally on (t, N_t) , contrary to the strict form of dynamic consistency.²⁰ The fact that utility is recursive but not strictly so suggests that, given an optimization problem, though not every time 0 optimal plan may be pursued subsequently at all relevant nodes, under suitable regularity conditions there will exist at least one time 0 optimal plan that will be implemented. This is the case for an example in [8], but a general analysis remains to be done.

²⁰The reason that c' and c are indifferent ex ante is that N_t is ex ante null according to the worst-case measure for c .

A. Appendix: Uniform Continuity

Define the shifted canonical space by

$$\Omega^t \equiv \{\omega \in C^d([t, T]) \mid \omega_t = 0\}.$$

Denote by B^t the canonical process on Ω^t , (a shift of B), by P_0^t the probability measure on Ω^t such that B^t is a Brownian motion, and by $\mathcal{F}^t = \{\mathcal{F}_\tau^t\}_{t \leq \tau \leq T}$ the filtration generated by B^t .

Fix $0 \leq s \leq t \leq T$. For $\omega \in \Omega^s$ and $\tilde{\omega} \in \Omega^t$, the concatenation of ω and $\tilde{\omega}$ at t is the path

$$(\omega \otimes_t \tilde{\omega})_\tau \triangleq \omega_\tau 1_{[s, t)}(\tau) + (\omega_t + \tilde{\omega}_\tau) 1_{[t, T]}(\tau), \quad s \leq \tau \leq T.$$

Given an \mathcal{F}_T^s -measurable random variable ξ on Ω^s and $\omega \in \Omega^s$, define the shifted random variable $\xi^{t, \omega}$ on Ω^t by

$$\xi^{t, \omega}(\tilde{\omega}) \triangleq \xi(\omega \otimes_t \tilde{\omega}), \quad \tilde{\omega} \in \Omega^t.$$

For an \mathcal{F}^s -progressively measurable process $(X_\tau)_{s \leq \tau \leq T}$, the shifted process $(X_\tau^{t, \omega})_{t \leq \tau \leq T}$ is \mathcal{F}^t -progressively measurable.

Let $(\Theta_s)_{0 \leq s \leq T}$ be a process of correspondences as in Section 2.2. For each (t, ω) in $[0, T] \times \Omega$, define a new process of correspondences $(\Theta_s^{t, \omega})_{t \leq s \leq T}$ by:

$$\Theta_s^{t, \omega}(\tilde{\omega}) \triangleq \Theta_s(\omega \otimes_t \tilde{\omega}), \quad \tilde{\omega} \in \Omega^t.$$

Then

$$\Theta_s^{t, \omega} : \Omega^t \rightsquigarrow \mathbb{R}^d \times \mathbb{R}^{d \times d}.$$

The new process inherits conditions (i)-(iv) and (vi). The same is true for (v), which we define next.

The following definition is adapted from Nutz [25, Defn. 3.2]. Say that (Θ_t) is *uniformly continuous* if for all $\delta > 0$ and $(t, \omega) \in [0, T] \times \Omega$ there exists $\epsilon(t, \omega, \delta) > 0$ such that if $\sup_{0 \leq s \leq t} |\omega_s - \omega'_s| \leq \epsilon$, then

$$ri^\delta \Theta_s^{t, \omega}(\tilde{\omega}) \subseteq ri^\epsilon \Theta_s^{t, \omega'}(\tilde{\omega}) \text{ for all } (s, \tilde{\omega}) \in [t, T] \times \Omega^t.$$

The process (Θ_t) , and hence also Θ , are fixed throughout the appendices. Thus we write \mathcal{P} instead of \mathcal{P}^Θ . Define $\mathcal{P}^0 \subset \mathcal{P}$ by

$$\mathcal{P}^0 \equiv \{P \in \mathcal{P} : \exists \delta > 0 \theta_t(\omega) \in ri^\delta \Theta_t(\omega) \text{ for all } (t, \omega) \in [0, T] \times \Omega\}. \quad (\text{A.1})$$

B. Appendix: Conditioning

Theorem 2.6 is proven here.

In fact, we prove more than is stated in the theorem because we prove also the following representation for conditional expectation. For each $t \in [0, T]$ and $P \in \mathcal{P}$, define

$$\mathcal{P}(t, P) = \{P' \in \mathcal{P} : P' = P \text{ on } \mathcal{F}_t\}. \quad (\text{B.1})$$

Then for each $\xi \in \widehat{L^2}(\Omega)$, $t \in [0, T]$ and $P \in \mathcal{P}$,²¹

$$\widehat{E}[\xi \mid \mathcal{F}_t] = \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E_{P'}[\xi \mid \mathcal{F}_t], \quad P\text{-a.e.} \quad (\text{B.2})$$

Perspective on this representation follows from considering the special case where all measures in \mathcal{P} are equivalent. Fix a measure P_0 in \mathcal{P} . Then the condition (B.2) becomes

$$\widehat{E}[\xi \mid \mathcal{F}_t] = \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E_{P'}[\xi \mid \mathcal{F}_t], \quad P_0\text{-a.e., for every } P \in \mathcal{P}.$$

Accordingly, the random variable on the right side is (up to P_0 -nullity) independent of P . Apply $\cup_{P \in \mathcal{P}} \mathcal{P}(t, P) = \mathcal{P}$ to conclude that²²

$$\widehat{E}[\xi \mid \mathcal{F}_t] = \operatorname{ess\,sup}_{P' \in \mathcal{P}} E_{P'}[\xi \mid \mathcal{F}_t], \quad P_0\text{-a.s.}$$

In other words, conditioning amounts to applying the usual Bayesian conditioning to each measure in \mathcal{P} and taking the upper envelope of the resulting expectations. This coincides with the prior-by-prior Bayesian updating rule in the Chen-Epstein model (apart from the different convention there of formulating expectations using infima rather than suprema). When the set \mathcal{P} is undominated, different measures in \mathcal{P} typically provide different perspectives on any random variable. Accordingly, (B.2) describes $\widehat{E}[\xi \mid \mathcal{F}_t]$ completely by describing how it appears when seen through the lens of every measure in \mathcal{P} .

²¹Because all measures in $\mathcal{P}(t, P)$ coincide with P on \mathcal{F}_t , essential supremum is defined as in the classical case (see He et al. [14, pp. 8-9], for example). Thus the right hand side of (B.2) is defined to be any random variable ξ^* satisfying: (i) ξ^* is \mathcal{F}_t -measurable, $E_{P'}[\xi \mid \mathcal{F}_t] \leq \xi^*$ P -a.e. and (ii) $\xi^* \leq \xi^{**}$ P -a.e. for any other random variable ξ^{**} satisfying (i).

²²The P_0 -null event can be chosen independently of P by He et al. [14, Theorem 1.3]: Let \mathcal{H} be a non-empty family of random variables on any probability space. Then the essential supremum exists and there is a countable number of elements (ξ_n) of \mathcal{H} such that $\operatorname{ess\,sup} \mathcal{H} = \bigvee_n \xi_n$.

Our proof adapts arguments from Nutz [25]. He constructs a time consistent sublinear expectation in a setting with ambiguity about volatility but not about drift. Because of this difference and because we use a different approach to construct the set of priors, his results do not apply directly.

For any probability measure P on the canonical space Ω , a corresponding *regular conditional probability* P_t^ω is defined to be any mapping $P_t^\omega : \Omega \times \mathcal{F}_T \rightarrow [0, 1]$ satisfying the following conditions:

- (i) for any ω , P_t^ω is a probability measure on (Ω, \mathcal{F}_T) .
- (ii) for any $A \in \mathcal{F}_T$, $\omega \rightarrow P_t^\omega(A)$ is \mathcal{F}_t -measurable.
- (iii) for any $A \in \mathcal{F}_T$, $E^P[1_A | \mathcal{F}_t](\omega) = P_t^\omega(A)$, P -a.e.

Of course, P_t^ω is not defined uniquely by these properties. We will fix a version defined via (2.9) after proving in Lemma B.2 that P_t^ω defined there satisfies the conditions characterizing a regular conditional probability. This explains our use of the same notation P_t^ω in both instances.

If P is a probability on Ω^s and $\omega \in \Omega^s$, for any $A \in \mathcal{F}_T^t$ we define

$$P^{t,\omega}(A) \triangleq P_t^\omega(\omega \otimes_t A),$$

where $\omega \otimes_t A \triangleq \{\omega \otimes_t \tilde{\omega} \mid \tilde{\omega} \in A\}$.

For each $(t, \omega) \in [0, T] \times \Omega$, let

$$\theta_s(\tilde{\omega}) = (\mu_s(\tilde{\omega}), \sigma_s(\tilde{\omega})) \in ri^\delta \Theta_s^{t,\omega}(\tilde{\omega}) \text{ for all } (s, \tilde{\omega}) \in [t, T] \times \Omega^t,$$

where $\delta > 0$ is some constant. Let $X^{t,\theta} = (X_s^{t,\theta})$ be the solution of the following equation (under P_0^t)

$$dX_s^{t,\theta} = \mu_s(X_s^{t,\theta})ds + \sigma_s(X_s^{t,\theta})dB_s^t, \quad X_t^{t,\theta} = 0, \quad s \in [t, T].$$

Then $X^{t,\theta}$ and P_0^t induce a probability measure $P^{t,\theta}$ on Ω^t .

Remark 3. *For nonspecialists we emphasize the difference between the preceding SDE and (2.9). The former is defined on the time interval $[t, T]$, and is a shifted version of (2.3), while (2.9) is defined on the full interval $[0, T]$. This difference is reflected also in the difference between the induced measures: the shifted measure $P_0^t \in \Delta(\Omega^t)$ and the conditional measure $P_t^\omega \in \Delta(\Omega)$. Part of the analysis to follow concerns shifted SDE's, random variables and measures and their relation to unshifted conditional counterparts. (See also Appendix A.)*

Let $\mathcal{P}^0(t, \omega)$ be the collection of such induced measures $P^{t, \theta}$. Define $\deg(t, \omega, P^{t, \theta}) = \delta^*/2 > 0$, where δ^* is the supremum of all δ such that

$$\theta_s(\tilde{\omega}) \in ri^\delta \Theta_s^{t, \omega}(\tilde{\omega}) \text{ for all } s \text{ and } \tilde{\omega}.$$

Note that at time $t = 0$, $\mathcal{P}^0(0, \omega)$ does not depend on ω and it coincides with \mathcal{P}^0 .

For each $(t, \omega) \in [0, T] \times \Omega$, let $\mathcal{P}(t, \omega)$ be the collection of all induced measures $P^{t, \theta}$ such that

$$\theta_s(\tilde{\omega}) = (\mu_s(\tilde{\omega}), \sigma_s(\tilde{\omega})) \in \Theta_s^{t, \omega}(\tilde{\omega}) \text{ for all } (s, \tilde{\omega}) \in [t, T] \times \Omega^t.$$

Note that $\mathcal{P}(0, \omega) = \mathcal{P}$.

Now we investigate the relationship between $\mathcal{P}(t, \omega)$ and \mathcal{P}_t^ω . (Recall that for any $P = P^\theta$ in \mathcal{P} , the measure P_t^ω is defined via (2.9) and \mathcal{P}_t^ω is the set of all such measures as in (2.10).)

For any $\theta = (\mu, \sigma) \in \Theta$, $(t, \omega) \in [0, T] \times \Omega$, define the shifted process $\bar{\theta}$ by

$$\bar{\theta}_s(\tilde{\omega}) = (\bar{\mu}_s(\tilde{\omega}), \bar{\sigma}_s(\tilde{\omega})) \triangleq (\mu_s^{t, \omega}(\tilde{\omega}), \sigma_s^{t, \omega}(\tilde{\omega})) \text{ for } (s, \tilde{\omega}) \in [t, T] \times \Omega^t. \quad (\text{B.3})$$

Then $\bar{\theta}_s(\tilde{\omega}) \in \Theta_s^{t, \omega}(\tilde{\omega})$. Consider the equation

$$\begin{cases} d\bar{X}_s = \bar{\mu}_s(\bar{X}.)ds + \bar{\sigma}_s(\bar{X}.)dB_s^t, & s \in [t, T], \\ \bar{X}_t = 0. \end{cases} \quad (\text{B.4})$$

Under P_0^t , the solution \bar{X} induces a probability measure $P^{t, \bar{\theta}}$ on Ω^t . By the definition of $\mathcal{P}(t, \omega)$, $P^{t, \bar{\theta}} \in \mathcal{P}(t, \omega)$.

Lemma B.1. $\{(P')^{t, \omega} : P' \in \mathcal{P}_t^\omega\} = \mathcal{P}(t, \omega)$.

Proof. \subset : For any $\theta = (\mu, \sigma) \in \Theta$, $P = P^\theta$ and $(t, \omega) \in [0, T] \times \Omega$, we claim that

$$(P_t^\omega)^{t, \omega} = P^{t, \bar{\theta}}$$

where $P^{t, \bar{\theta}}$ is defined through (B.4). Because B has independent increments under P_0 , the shifted solution $(X_s^{t, \omega})_{t \leq s \leq T}$ of (2.9) has the same distribution as does $(\bar{X}_s)_{t \leq s \leq T}$. This proves the claim.

\supset : Prove that for any $P^{t, \bar{\theta}} \in \mathcal{P}(t, \omega)$, there exists $\theta \in \Theta$ such that (where $P = P^\theta$)

$$P_t^\omega \in \mathcal{P}_t^\omega \text{ and } (P_t^\omega)^{t, \omega} = P^{t, \bar{\theta}}.$$

Each $P^{t,\bar{\theta}}$ in $\mathcal{P}(t, \omega)$ is induced by the solution \bar{X} of (B.4), where $\bar{\theta}(\tilde{\omega})$, defined in (B.3), lies in $\Theta_s^{t,\omega}(\tilde{\omega})$ for every $\tilde{\omega} \in \Omega^t$. For any $\theta \in \Theta$ such that

$$\theta_s^{t,\omega}(\tilde{\omega}) = (\mu_s^{t,\omega}(\tilde{\omega}), \sigma_s^{t,\omega}(\tilde{\omega})) = (\bar{\mu}_s(\tilde{\omega}), \bar{\sigma}_s(\tilde{\omega})), s \in [t, T],$$

consider the equation (2.9). Under P_0 , the solution X induces a probability P_t^ω on Ω . Because B has independent increments under P_0 , we know that $(P_t^\omega)^{t,\omega} = P^{t,\bar{\theta}}$. This completes the proof. \blacksquare

Lemma B.2. *For any $\theta \in \Theta$ and $P = P^\theta$, $\{P_t^\omega : (t, \omega) \in [0, T] \times \Omega\}$ is a version of the regular conditional probability of P .*

Proof. Firstly, for any $0 < t_1 < \dots < t_n \leq T$ and bounded, continuous functions φ and ψ , we prove that

$$E^P[\varphi(B_{t_1 \wedge t}, \dots, B_{t_n \wedge t})\psi(B_{t_1}, \dots, B_{t_n})] = E^P[\varphi(B_{t_1 \wedge t}, \dots, B_{t_n \wedge t})\psi_t] \quad (\text{B.5})$$

where $t \in [t_k, t_{k+1})$ and, for any $\hat{\omega} \in \Omega$,

$$\psi_t(\hat{\omega}) \triangleq E^{P^{t,\bar{\theta}}}[\psi(\hat{\omega}(t_1), \dots, \hat{\omega}(t_k), \hat{\omega}(t) + B_{t_{k+1}}^t, \dots, \hat{\omega}(t) + B_{t_n}^t)].$$

Note that $P^{t,\bar{\theta}}$ on Ω^t is induced by \bar{X} (see (B.3)) under P_0^t , and $P = P^\theta$ on Ω is induced by $X = X^\theta$ (see (2.3)) under P_0 . Then,

$$\begin{aligned} \psi_t(X(\omega)) &= \\ E^{P_0^t}[\psi(X_{t_1}(\omega), \dots, X_{t_k}(\omega), X_t(\omega) + \bar{X}_{t_{k+1}}(\tilde{\omega}), \dots, X_t(\omega) + \bar{X}_{t_n}(\tilde{\omega}))]. \end{aligned} \quad (\text{B.6})$$

Because B has independent increments under P_0 , the shifted regular conditional probability

$$P_0^{t,\omega} = P_0^t, P_0\text{-a.e.}$$

Thus (B.6) holds under probability $P_0^{t,\omega}$.

Because $P_0^{t,\omega}$ is the shifted probability of $(P_0)_t^\omega$, we have

$$\begin{aligned} &\psi_t(X(\omega)) \\ &= E^{(P_0)_t^\omega}[\psi(X_{t_1}(\omega), \dots, X_{t_k}(\omega), X_{t_{k+1}}(\omega), \dots, X_{t_n}(\omega))] \\ &= E^{P_0}[\psi(X_{t_1}(\omega), \dots, X_{t_k}(\omega), X_{t_{k+1}}(\omega), \dots, X_{t_n}(\omega)) \mid \mathcal{F}_t](\omega), P_0\text{-a.e.} \end{aligned}$$

Further, because P is induced by X and P_0 ,

$$\begin{aligned}
& E^P[\varphi(B_{t_1 \wedge t}, \dots, B_{t_n \wedge t})\psi_t] \\
&= E^{P_0}[\varphi(X_{t_1 \wedge t}, \dots, X_{t_n \wedge t})\psi_t(X)] \\
&= E^{P_0}[\varphi(X_{t_1 \wedge t}, \dots, X_{t_n \wedge t})E^{P_0}[\psi(X_{t_1}, \dots, X_{t_n}) \mid \mathcal{F}_t]] \\
&= E^{P_0}[\varphi(X_{t_1 \wedge t}, \dots, X_{t_n \wedge t})\psi(X_{t_1}, \dots, X_{t_n})] \\
&= E^P[\varphi(B_{t_1 \wedge t}, \dots, B_{t_n \wedge t})\psi(B_{t_1}, \dots, B_{t_n})].
\end{aligned}$$

Secondly, note that (B.5) is true and φ and (t_1, \dots, t_n) are arbitrary. Then by the definition of the regular conditional probability, for P -a.e. $\hat{\omega} \in \Omega$ and $t \in [t_k, t_{k+1})$,

$$\psi_t(\hat{\omega}) = E^{\tilde{P}^{t, \hat{\omega}}}[\psi(\hat{\omega}(t_1), \dots, \hat{\omega}(t_k), \hat{\omega}(t) + B_{t_{k+1}}^t, \dots, \hat{\omega}(t) + B_{t_n}^t)], \quad (\text{B.7})$$

where $\tilde{P}^{t, \hat{\omega}}$ is the shift of the regular conditional probability of P given $(t, \hat{\omega}) \in [0, T] \times \Omega$.

By standard approximating arguments, there exists a set M such that $P(M) = 0$ and for any $\omega \notin M$, (B.7) holds for all continuous bounded function ψ and (t_1, \dots, t_n) . This means that for $\omega \notin M$ and for all bounded \mathcal{F}_T^t -measurable random variables ξ

$$E^{\tilde{P}^{t, \omega}} \xi = E^{P^{t, \bar{\theta}}} \xi.$$

Then $\tilde{P}^{t, \omega} = P^{t, \bar{\theta}}$ P -a.e. By Lemma B.1, $\tilde{P}^{t, \omega} = P^{t, \bar{\theta}} = (P_t^\omega)^{t, \omega}$, P -a.e. Thus P_t^ω is a version of the regular conditional probability for P . \blacksquare

In the following, we always use P_t^ω defined by (2.9) as the fixed version of regular conditional probability for $P \in \mathcal{P}$. Thus

$$E^{P_t^\omega} \xi = E^P[\xi \mid \mathcal{F}_t](\omega), \quad P\text{-a.e.}$$

Because we will want to consider also dynamics beginning at arbitrary s , let $0 \leq s \leq T$, $\bar{\omega} \in \Omega$, and $P \in \mathcal{P}(s, \bar{\omega})$. Then given $(t, \omega) \in [s, T] \times \Omega^s$, we can fix a version of the regular conditional probability, also denoted P_t^ω (here a measure on Ω^s), which is constructed in a similar fashion via a counterpart of (2.9). Define

$$\mathcal{P}_t^\omega(s, \bar{\omega}) = \{P_t^\omega : P \in \mathcal{P}(s, \bar{\omega})\} \text{ and}$$

$$\mathcal{P}_t^{0, \omega}(s, \bar{\omega}) = \{P_t^\omega : P \in \mathcal{P}^0(s, \bar{\omega})\}.$$

In each case, the obvious counterpart of the result in Lemma B.1 is valid.

The remaining arguments are divided into three steps. First we prove that if $\xi \in UC_b(\Omega)$ and if the set \mathcal{P} is replaced by \mathcal{P}^0 defined in (A.1), then the counterparts of (B.2) and (2.12) are valid. Then we show that \mathcal{P}^0 (resp. $\mathcal{P}^0(t, \omega)$) is dense in \mathcal{P} (resp. $\mathcal{P}(t, \omega)$). In Step 3 the preceding is extended to apply to all ξ in the completion of $UC_b(\Omega)$.

Step 1

Given $\xi \in UC_b(\Omega)$, define

$$v_t^0(\omega) \triangleq \sup_{P \in \mathcal{P}^0(t, \omega)} E^P \xi^{t, \omega}, \quad (t, \omega) \in [0, T] \times \Omega. \quad (\text{B.8})$$

(The notation $\xi^{t, \omega}$ is defined in Appendix A.)

Lemma B.3. *Let $0 \leq s \leq t \leq T$ and $\bar{\omega} \in \Omega$. Given $\epsilon > 0$, there exist a sequence $(\hat{\omega}^i)_{i \geq 1}$ in Ω^s , an \mathcal{F}_t^s -measurable countable partition $(H^i)_{i \geq 1}$ of Ω^s , and a sequence $(P^i)_{i \geq 1}$ of probability measures on Ω^t such that*

- (i) $\|\omega - \hat{\omega}^i\|_{[s, t]} \equiv \sup_{s \leq \tau \leq t} |\omega_\tau - \hat{\omega}_\tau^i| \leq \epsilon$ for all $\omega \in H^i$;
- (ii) $P^i \in \mathcal{P}^0(t, \bar{\omega} \otimes_s \omega)$ for all $\omega \in H^i$ and $\inf_{\omega \in H^i} \deg(t, \bar{\omega} \otimes_s \omega, P^i) > 0$;
- (iii) $v_t^0(\bar{\omega} \otimes_s \hat{\omega}^i) \leq E^{P^i}[\xi^{t, \bar{\omega} \otimes_s \hat{\omega}^i}] + \epsilon$.

Proof. Given $\epsilon > 0$ and $\hat{\omega} \in \Omega^s$, by (B.8) there exists $P(\hat{\omega}) \in \mathcal{P}^0(t, \bar{\omega} \otimes_s \hat{\omega})$ such that

$$v_t^0(\bar{\omega} \otimes_s \hat{\omega}) \leq E^{P(\hat{\omega})}[\xi^{t, \bar{\omega} \otimes_s \hat{\omega}}] + \epsilon.$$

Because (Θ_t) is uniformly continuous, there exists $\epsilon(\hat{\omega}) > 0$ such that

$$P(\hat{\omega}) \in \mathcal{P}^0(t, \bar{\omega} \otimes_s \omega') \text{ for all } \omega' \in B(\epsilon(\hat{\omega}), \hat{\omega}) \text{ and } \inf_{\omega' \in B(\epsilon(\hat{\omega}), \hat{\omega})} \deg(t, \bar{\omega} \otimes_s \omega', P(\hat{\omega})) > 0$$

where $B(\epsilon, \hat{\omega}) \triangleq \{\omega' \in \Omega^s \mid \|\omega' - \hat{\omega}\|_{[s, t]} < \epsilon\}$ is the open $\|\cdot\|_{[s, t]}$ ball. Then $\{B(\epsilon(\hat{\omega}), \hat{\omega}) \mid \hat{\omega} \in \Omega^s\}$ forms an open cover of Ω^s . There exists a countable subcover because Ω^s is separable. We denote the subcover by

$$B^i \triangleq B(\epsilon(\hat{\omega}^i), \hat{\omega}^i), \quad i = 1, 2, \dots$$

and define a partition of Ω^s by

$$H^1 \triangleq B^1, \quad H^{i+1} \triangleq B^{i+1} \setminus (H^1 \cup \dots \cup H^i), \quad i \geq 1.$$

Set $P^i \triangleq P(\hat{\omega}^i)$. Then (i)-(iii) are satisfied. ■

For any $A \in \mathcal{F}_T^s$, define

$$A^{t, \omega} = \{\tilde{\omega} \in \Omega^t \mid \omega \otimes_t \tilde{\omega} \in A\}.$$

Lemma B.4. Let $0 \leq s \leq t \leq T$, $\bar{\omega} \in \Omega$ and $P \in \mathcal{P}^0(s, \bar{\omega})$. Let $(H^i)_{0 \leq i \leq N}$ be a finite \mathcal{F}_t^s -measurable partition of Ω^s . For $1 \leq i \leq N$, assume that $\bar{P}^i \in \mathcal{P}^0(t, \bar{\omega} \otimes_s \omega)$ for all $\omega \in H^i$ and that $\inf_{\omega \in H^i} \deg(t, \bar{\omega} \otimes_s \omega, P^i) > 0$. Define \bar{P} by

$$\bar{P}(A) \triangleq P(A \cap H^0) + \sum_{i=1}^N E^P[P^i(A^{t, \omega}) 1_{H^i}(\omega)], \quad A \in \mathcal{F}_T^s.$$

Then: (i) $\bar{P} \in \mathcal{P}^0(s, \bar{\omega})$.

(ii) $\bar{P} = P$ on \mathcal{F}_t^s .

(iii) $\bar{P}^{t, \omega} = P^{t, \omega}$ P -a.e. $[\omega]$ on H^0 .

(iv) $\bar{P}^{t, \omega} = P^i$ P -a.e. $[\omega]$ on H^i , $1 \leq i \leq N$.

Proof. (i) Let θ (resp. θ^i) be the \mathcal{F}^s (resp. \mathcal{F}^t) -measurable process such that $P = P^\theta$ (resp. $P^i = P^{\theta^i}$). Define $\bar{\theta}$ by

$$\bar{\theta}_\tau(\omega) \triangleq \theta_\tau(\omega) 1_{[s, t)}(\tau) + [\theta_\tau(\omega) 1_{H^0}(\omega) + \sum_{i=1}^N \theta_\tau^i(\omega^t) 1_{H^i}(\omega)] 1_{[t, T]}(\tau) \quad (\text{B.9})$$

for $(\tau, \omega) \in [s, T] \times \Omega^s$. Then $P^{\bar{\theta}} \in \mathcal{P}^0(s, \bar{\omega})$ and $\bar{P} = P^{\bar{\theta}}$ on \mathcal{F}_T^s .

(ii) Let $A \in \mathcal{F}_t^s$. We prove $\bar{P}(A) = P(A)$. Note that for $\omega \in \Omega^s$, if $\omega \in A$, then $A^{t, \omega} = \Omega^t$; otherwise, $A^{t, \omega} = \emptyset$. Thus, $P^i(A^{t, \omega}) = 1_A(\omega)$ for $1 \leq i \leq N$, and

$$\bar{P}(A) = P(A \cap H^0) + \sum_{i=1}^N E^P[1_A(\omega) 1_{H^i}(\omega)] = \sum_{i=0}^N P(A \cap H^i) = P(A).$$

(iii)-(iv) Recall the definition of P_t^ω by (2.9). Note that $\bar{P} = P^{\bar{\theta}}$ where $\bar{\theta}$ is defined by (B.9). Then it is easy to show that the shifted regular conditional probability $\bar{P}^{t, \omega}$ satisfies (iii)-(iv). \blacksquare

The technique used in proving Nutz [25, Theorem 4.5] can be adapted to prove the following dynamic programming principle.

Proposition B.5. Let $0 \leq s \leq t \leq T$, $\xi \in UC_b(\Omega)$ and define v_t^0 by (B.8). Then

$$v_s^0(\bar{\omega}) = \sup_{P' \in \mathcal{P}^0(s, \bar{\omega})} E^{P'}[(v_t^0)^{s, \bar{\omega}}] \text{ for all } \bar{\omega} \in \Omega, \quad (\text{B.10})$$

$$v_s^0 = \text{ess sup}_{P' \in \mathcal{P}^0(s, P)} E^{P'}[v_t^0 \mid \mathcal{F}_s] \text{ } P\text{-a.e. for all } P \in \mathcal{P}^0, \quad (\text{B.11})$$

and

$$v_s^0 = \text{ess sup}_{P' \in \mathcal{P}^0(s, P)} E^{P'}[\xi \mid \mathcal{F}_s] \text{ } P\text{-a.e. for all } P \in \mathcal{P}^0. \quad (\text{B.12})$$

Proof. *Proof of (B.10):* First prove \leq . Let $\bar{\omega} \in \Omega$ and $\omega \in \Omega^s$, by Lemma B.1,

$$\{(P')^{t,\omega} \mid P' \in \mathcal{P}_t^{0,\omega}(s, \bar{\omega})\} = \mathcal{P}^0(t, \bar{\omega} \otimes_s \omega).$$

For $P \in \mathcal{P}^0(s, \bar{\omega})$,

$$\begin{aligned} & E^{P^{t,\omega}}[(\xi^{s,\bar{\omega}})^{t,\omega}] = E^{P^{t,\omega}}[\xi^{t,\bar{\omega} \otimes_s \omega}] \\ & \leq \sup_{P' \in \mathcal{P}^0(t, \bar{\omega} \otimes_s \omega)} E^{P'}[\xi^{t,\bar{\omega} \otimes_s \omega}] \\ & = v_t^0(\bar{\omega} \otimes_s \omega). \end{aligned} \tag{B.13}$$

Note that $v_t^0(\bar{\omega} \otimes_s \omega) = (v_t^0)^{s,\bar{\omega}}(\omega)$. Taking the expectation under P on both sides of (B.13) yields

$$E^P \xi^{s,\bar{\omega}} \leq E^P[(v_t^0)^{s,\bar{\omega}}].$$

The desired result holds by taking the supremum over $P \in \mathcal{P}^0(s, \bar{\omega})$.

Prove \geq . Let $\delta > 0$. Because ${}^t\Omega$ (defined in section 2.1) is a Polish space and $(v_t^0)^{s,\bar{\omega}}$ is \mathcal{F}_t^s -measurable, by Lusin's Theorem there exists a compact set $G \in \mathcal{F}_t^s$ with $P(G) > 1 - \delta$ and such that $(v_t^0)^{s,\bar{\omega}}$ is uniformly continuous on G .

Let $\epsilon > 0$. By Lemma B.3, there exist a sequence $(\hat{\omega}^i)_{i \geq 1}$ in G , an \mathcal{F}_t^s -measurable partition $(H^i)_{i \geq 1}$ of G , and a sequence $(P^i)_{i \geq 1}$ of probability measures such that

- (a) $\|\omega - \hat{\omega}^i\|_{[s,t]} \leq \epsilon$ for all $\omega \in H^i$;
- (b) $P^i \in \mathcal{P}^0(t, \bar{\omega} \otimes_s \omega)$ for all $\omega \in H^i$ and $\inf_{\omega \in H^i} \deg(t, \bar{\omega} \otimes_s \omega, P^i) > 0$;
- (c) $v_t^0(\bar{\omega} \otimes_s \hat{\omega}^i) \leq E^{P^i}[\xi^{t,\bar{\omega} \otimes_s \hat{\omega}^i}] + \epsilon$.

Let

$$A_N \triangleq H^1 \cup \dots \cup H^N, \quad N \geq 1.$$

For $P \in \mathcal{P}^0(s, \bar{\omega})$, define

$$\mathcal{P}^0(s, \bar{\omega}, t, P) \triangleq \{P' \in \mathcal{P}^0(s, \bar{\omega}) : P' = P \text{ on } \mathcal{F}_t^s\}.$$

Apply Lemma B.4 to the finite partition $\{H^1, \dots, H^N, A_N^c\}$ of Ω^s to obtain a measure $P_N \in \mathcal{P}^0(s, \bar{\omega})$ such that $P_N \in \mathcal{P}^0(s, \bar{\omega}, t, P)$ and

$$P_N^{t,\omega} = \begin{cases} P^{t,\omega} & \text{for } \omega \in A_N^c, \\ P^i & \text{for } \omega \in H^i, \quad 1 \leq i \leq N \end{cases} \tag{B.14}$$

Because $(v_t^0)^{s,\bar{\omega}}$ and ξ are uniformly continuous on G , there exist moduli of continuity $\rho^{((v_t^0)^{s,\bar{\omega}}|G)}(\cdot)$ and $\rho^{(\xi)}(\cdot)$ such that

$$\begin{aligned} & |(v_t^0)^{s,\bar{\omega}}(\omega) - (v_t^0)^{s,\bar{\omega}}(\omega')| \leq \rho^{((v_t^0)^{s,\bar{\omega}}|G)}(\|\omega - \omega'\|_{[s,t]}), \\ & |\xi^{t,\bar{\omega} \otimes_s \omega} - \xi^{t,\bar{\omega} \otimes_s \omega'}| \leq \rho^{(\xi)}(\|\omega - \omega'\|_{[s,t]}). \end{aligned}$$

Let $\omega \in H^i$ for some $1 \leq i \leq N$. Then

$$\begin{aligned}
& (v_t^0)^{s, \bar{\omega}}(\omega) \\
& \leq (v_t^0)^{s, \bar{\omega}}(\hat{\omega}^i) + \rho^{((v_t^0)^{s, \bar{\omega}}|G)}(\epsilon) \\
& \leq E^{P^i}[\xi^{t, \bar{\omega} \otimes_s \hat{\omega}^i}] + \epsilon + \rho^{((v_t^0)^{s, \bar{\omega}}|G)}(\epsilon) \\
& \leq E^{P^i}[\xi^{t, \bar{\omega} \otimes_s \omega}] + \rho^{(\xi)}(\epsilon) + \epsilon + \rho^{((v_t^0)^{s, \bar{\omega}}|G)}(\epsilon) \\
& = E^{P_N^{t, \omega}}[\xi^{t, \bar{\omega} \otimes_s \omega}] + \rho^{(\xi)}(\epsilon) + \epsilon + \rho^{((v_t^0)^{s, \bar{\omega}}|G)}(\epsilon) \\
& = E^{P_N}[\xi^{s, \bar{\omega}} | \mathcal{F}_t^s](\omega) + \rho^{(\xi)}(\epsilon) + \epsilon + \rho^{((v_t^0)^{s, \bar{\omega}}|G)}(\epsilon) \text{ for } P\text{-a.e. } \omega \in H^i.
\end{aligned} \tag{B.15}$$

These inequalities are due respectively to uniform continuity of v_t^0 , Lemma B.3(iii), uniform continuity of ξ , equation (B.14), and the fact that $P_N \in \mathcal{P}(t, P)$. Because $P = P_N$ on \mathcal{F}_t^s , taking the P -expectation on both sides yields

$$E^P[(v_t^0)^{s, \bar{\omega}} 1_{A_N}] \leq E^{P_N}[\xi^{s, \bar{\omega}} 1_{A_N}] + \rho^{(\xi)}(\epsilon) + \epsilon + \rho^{((v_t^0)^{s, \bar{\omega}}|G)}(\epsilon).$$

Note that $\xi \in UC_b(\Omega)$ and $P_N(G \setminus A_N) = P(G \setminus A_N) \rightarrow 0$ as $N \rightarrow \infty$. Let $N \rightarrow \infty$ and $\epsilon \rightarrow 0$ to obtain that

$$E^P[(v_t^0)^{s, \bar{\omega}} 1_G] \leq \sup_{P' \in \mathcal{P}^0(s, \bar{\omega}, t, P)} E^{P'}[\xi^{s, \bar{\omega}} 1_G].$$

Because $\delta > 0$ is arbitrary, similar arguments show that

$$E^P[(v_t^0)^{s, \bar{\omega}}] \leq \sup_{P' \in \mathcal{P}^0(s, \bar{\omega}, t, P)} E^{P'} \xi^{s, \bar{\omega}} \leq \sup_{P' \in \mathcal{P}^0(s, \bar{\omega})} E^{P'} \xi^{s, \bar{\omega}} = v_s^0(\bar{\omega}).$$

But $P \in \mathcal{P}^0(s, \bar{\omega})$ is arbitrary. This completes the proof of (B.10).

Proof of (B.11): Fix $P \in \mathcal{P}^0$. First we prove that

$$v_t^0 \leq \text{ess sup}_{P' \in \mathcal{P}^0(t, P)} E^{P'}[\xi | \mathcal{F}_t] \quad P\text{-a.e.} \tag{B.16}$$

Argue as in the second part of the preceding proof, specialized to $s = 0$. Conclude that there exists $P_N \in \mathcal{P}^0(t, P)$ such that, as a counterpart of (B.15),

$$v_t^0(\omega) \leq E^{P_N}[\xi | \mathcal{F}_t](\omega) + \rho^{(\xi)}(\epsilon) + \epsilon + \rho^{(v_t^0|G)}(\epsilon) \text{ for } P\text{-a.e. } \omega \in A_N.$$

Because $P = P_N$ on \mathcal{F}_t , as $N \rightarrow \infty$ and $\delta \rightarrow 0$, one obtains (B.16).

Now prove the inequality \leq in (B.11). For any $P' \in \mathcal{P}^0(s, P)$, we know that $(P')^{t, \omega} \in \mathcal{P}^0(t, \omega)$. From (B.10),

$$v_t^0(\omega) \geq E^{(P')^{t, \omega}}[\xi^{t, \omega}] = E^{P'}[\xi | \mathcal{F}_t](\omega) \quad P'\text{-a.e.}$$

Taking the conditional expectation on both sides yields $E^{P'}[\xi | \mathcal{F}_s] \leq E^{P'}[v_t^0 | \mathcal{F}_s]$ P' -a.e., hence also P -a.e. Thus

$$v_s^0 \leq \operatorname{ess\,sup}_{P' \in \mathcal{P}^0(s,P)} E^{P'}[\xi | \mathcal{F}_s] \leq \operatorname{ess\,sup}_{P' \in \mathcal{P}^0(s,P)} E^{P'}[v_t^0 | \mathcal{F}_s] \quad P\text{-a.e.}$$

Thirdly, we prove the converse direction holds in (B.11). For any $P' \in \mathcal{P}^0(s, P)$, by (B.10) we have

$$v_s^0(\omega) \geq E^{(P')^{s,\omega}}[(v_t^0)^{s,\omega}] = E^{P'}[v_t^0 | \mathcal{F}_s](\omega)$$

P' -a.e. on \mathcal{F}_s and hence P -a.e.

Equation (B.12) is implied by (B.11) because $v_T^0 = \xi$. ■

STEP 2

Refer to the topology induced on $\Delta(\Omega)$ by bounded continuous functions as the weak-convergence topology.

Lemma B.6. (a) \mathcal{P}^0 is dense in \mathcal{P} in the weak convergence topology.

(b) For each t and ω , $\mathcal{P}^0(t, \omega)$ is dense in $\mathcal{P}(t, \omega)$ in the weak convergence topology.

Proof. (a) Let $P^{\theta^0} \in \mathcal{P}^0$ and $P^\theta \in \mathcal{P}$, and define

$$\theta^\epsilon = \epsilon\theta^0 + (1 - \epsilon)\theta,$$

where $0 < \epsilon < 1$. By Uniform Interiority for (Θ_t) , there exists $\delta > 0$ such that $\theta_t^{\epsilon^\delta}(\omega) \in \operatorname{ri}^\delta \Theta_t(\omega)$ for all t and ω . Thus $P^{\theta^\epsilon} \in \mathcal{P}^0$.

By the standard approximation of a stochastic differential equation (see Gihman and Skorohod [12, Thm 3.15]), as $\epsilon \rightarrow 0$ there exists a subsequence of X^{θ^ϵ} , which we still denote by X^{θ^ϵ} , such that

$$\sup_{0 \leq t \leq T} |X_t^{\theta^\epsilon} - X_t^\theta| \rightarrow 0 \quad P_0\text{-a.e.}$$

The Dominated Convergence Theorem implies that $P^{\theta^\epsilon} \rightarrow P^\theta$.

(b) The proof is similar. ■

For any $\theta = (\mu, \sigma) \in \Theta$, $(t, \omega) \in [0, T] \times \Omega$, define the shifted process $\bar{\theta}$, the process \bar{X} and the probability measure $P^{t, \bar{\theta}}$ exactly as in (B.3) and (B.4). As noted earlier, $P^{t, \bar{\theta}} \in \mathcal{P}(t, \omega)$.

Recall that for any P in \mathcal{P} , the measure P_t^ω is defined via (2.9); \mathcal{P}_t^ω is the set of all such measures (2.10). By construction,

$$P_t^\omega(\{\bar{\omega} \in \Omega : \bar{\omega}_s = \omega_s \text{ } s \in [0, t]\}) = 1.$$

We show shortly that P_t^ω is a version of the regular conditional probability for P . Given $\xi \in UC_b(\Omega)$, define

$$v_t(\omega) \triangleq \sup_{P \in \mathcal{P}(t, \omega)} E^P \xi^{t, \omega}, \quad (t, \omega) \in [0, T] \times \Omega. \quad (\text{B.17})$$

Lemma B.7. *For any $(t, \omega) \in [0, T] \times \Omega$ and $\xi \in UC_b(\Omega)$, we have*

$$v_t(\omega) = v_t^0(\omega), \quad (\text{B.18})$$

$$v_t(\omega) = \sup_{P \in \mathcal{P}_t^\omega} E^P \xi, \quad (\text{B.19})$$

and

$$v_t = \text{ess sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[\xi \mid \mathcal{F}_t] \text{ } P\text{-a.e. for all } P \in \mathcal{P}. \quad (\text{B.20})$$

Furthermore, for any $0 \leq s \leq t \leq T$,

$$v_s(\omega) = \sup_{P' \in \mathcal{P}(s, \omega)} E^{P'}[(v_t)^{s, \omega}] \text{ for all } \omega \in \Omega, \quad (\text{B.21})$$

and

$$v_s = \text{ess sup}_{P' \in \mathcal{P}(s, P)} E^{P'}[v_t \mid \mathcal{F}_s] \text{ } P\text{-a.e. for all } P \in \mathcal{P}. \quad (\text{B.22})$$

Proof. *Proof of (B.18):* It is implied by the fact that $\mathcal{P}^0(t, \omega)$ is dense in $\mathcal{P}(t, \omega)$.

Proof of (B.19): By the definition of $v_t(\omega)$, we know that

$$v_t(\omega) = \sup_{P \in \mathcal{P}(t, \omega)} E^P \xi^{t, \omega}.$$

By Lemma B.1,

$$\{(P')^{t, \omega} \mid P' \in \mathcal{P}_t^\omega\} = \mathcal{P}(t, \omega).$$

Thus

$$v_t(\omega) = \sup_{\hat{P} \in \mathcal{P}_t^\omega} E^{\hat{P}^{t, \omega}} \xi^{t, \omega} = \sup_{\hat{P} \in \mathcal{P}_t^\omega} E^{\hat{P}} \xi.$$

Proof of (B.20): Fix $P \in \mathcal{P}$. For any $P' \in \mathcal{P}(t, P)$, by Lemmas B.1 and B.2, $(P')^{t, \omega} \in \mathcal{P}(t, \omega)$. By the definition of $v_t(\omega)$,

$$v_t(\omega) \geq E^{(P')^{t, \omega}} \xi^{t, \omega} = E^{P'}[\xi \mid \mathcal{F}_t](\omega)$$

P' -a.e. on \mathcal{F}_t and hence P -a.e. Thus

$$v_t \geq \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[\xi \mid \mathcal{F}_t] \quad P\text{-a.e.}$$

Now we prove the reverse inequality. By (B.18), $v_t(\omega) = v_t^0(\omega)$. Then, using the same technique as in the proof of Proposition B.5 for the special case $s = 0$, there exists $P_N \in \mathcal{P}(t, P)$ such that, as a counterpart of (B.15),

$$\begin{aligned} v_t(\omega) &\leq E^{P_N}[\xi \mid \mathcal{F}_t](\omega) + \rho^{(\xi)}(\epsilon) + \epsilon + \rho^{(v_t|G)}(\epsilon) \\ &\leq \sup_N E^{P_N}[\xi \mid \mathcal{F}_t](\omega) + \rho^{(\xi)}(\epsilon) + \epsilon + \rho^{(v_t|G)}(\epsilon) \end{aligned}$$

for P -a.e. $\omega \in A_N$. Let $N \rightarrow \infty$ to obtain that, for P -a.e. $\omega \in G$,

$$v_t(\omega) \leq \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[\xi \mid \mathcal{F}_t](\omega) + \rho^{(\xi)}(\epsilon) + \epsilon + \rho^{(v_t|G)}(\epsilon).$$

Let $\epsilon \rightarrow 0$ to derive

$$v_t \leq \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[\xi \mid \mathcal{F}_t], \quad P\text{-a.e. on } G.$$

Note that G depends on δ , but not on ϵ . Let $\delta \rightarrow 0$ and conclude that

$$v_t \leq \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[\xi \mid \mathcal{F}_t], \quad P\text{-a.e.}$$

Proof of (B.21) and (B.22): The former is due to (B.10) and the fact that $P^0(t, \omega)$ is dense in $P(t, \omega)$. The proof of (B.22) is similar to the proof of (B.12) in Proposition B.5. ■

Now for any $\xi \in UC_b(\Omega)$, we define conditional expectation by

$$\hat{E}[\xi \mid \mathcal{F}_t](\omega) \triangleq v_t(\omega).$$

STEP 3

Thus far we have defined $\hat{E}[\xi \mid \mathcal{F}_t](\omega)$, for all t and ω , for any $\xi \in UC_b(\Omega)$. Now extend the operator $\hat{E}[\cdot \mid \mathcal{F}_t]$ to the completion of $UC_b(\Omega)$.

Lemma B.8 (Extension). *The mapping $\hat{E}[\cdot \mid \mathcal{F}_t]$ on $UC_b(\Omega)$ can be extended uniquely to a 1-Lipschitz continuous mapping $\hat{E}[\cdot \mid \mathcal{F}_t] : \widehat{L}^2(\Omega) \rightarrow \widehat{L}^2({}^t\Omega)$.*

Proof. Define $\widetilde{L}^2({}^t\Omega)$ to be the space of \mathcal{F}_t -measurable random variables X satisfying

$$\|X\| \triangleq (\hat{E}[|X|^2])^{\frac{1}{2}} = (\sup_{P \in \mathcal{P}} E^P[|X|^2])^{\frac{1}{2}} < \infty.$$

Obviously, $\widehat{L}^2({}^t\Omega) \subset \widetilde{L}^2({}^t\Omega)$.

(i) We prove that $\hat{E}[\cdot \mid \mathcal{F}_t]$ can be uniquely extended to a 1-Lipschitz continuous mapping

$$\hat{E}[\cdot \mid \mathcal{F}_t] : \widehat{L}^2(\Omega) \rightarrow \widetilde{L}^2({}^t\Omega).$$

For any ξ and η in $UC_b(\Omega)$,

$$|\hat{E}[\xi' \mid \mathcal{F}_t] - \hat{E}[\xi \mid \mathcal{F}_t]|^2 \leq \left(\hat{E}[|\xi' - \xi| \mid \mathcal{F}_t] \right)^2 \leq \hat{E}[|\xi' - \xi|^2 \mid \mathcal{F}_t],$$

where the first inequality follows primarily from the subadditivity of $\hat{E}[\cdot \mid \mathcal{F}_t]$, and the second is implied by Jensen's inequality applied to each P in \mathcal{P} . Thus

$$\begin{aligned} & \| \hat{E}[\xi' \mid \mathcal{F}_t] - \hat{E}[\xi \mid \mathcal{F}_t] \| \\ &= \left(\hat{E} \left[|\hat{E}[\xi' \mid \mathcal{F}_t] - \hat{E}[\xi \mid \mathcal{F}_t]|^2 \right] \right)^{1/2} \\ &\leq \left(\hat{E} \left[\hat{E}[|\xi' - \xi|^2 \mid \mathcal{F}_t] \right] \right)^{1/2} \\ &= \left(\hat{E}[|\xi' - \xi|^2] \right)^{1/2} = \| \xi' - \xi \|, \end{aligned}$$

where the second equality is due to the 'law of iterated expectations' for integrands in $UC_b(\Omega)$ proven in Lemma B.7.

As a consequence, $\hat{E}[\xi \mid \mathcal{F}_t] \in \widetilde{L}^2({}^t\Omega)$ for ξ and η in $UC_b(\Omega)$, and $\hat{E}[\cdot \mid \mathcal{F}_t]$ extends uniquely to a 1-Lipschitz continuous mapping from $\widehat{L}^2(\Omega)$ into $\widetilde{L}^2({}^t\Omega)$.

(ii) Now prove that $\hat{E}[\cdot \mid \mathcal{F}_t]$ maps $\widehat{L}^2(\Omega)$ into $\widehat{L}^2({}^t\Omega)$.

First we show that if $\xi \in UC_b(\Omega)$, then $\hat{E}[\xi \mid \mathcal{F}_t]$ is lower semicontinuous: Fix $\omega \in \Omega$. Since $\xi \in UC_b(\Omega)$, there exists a modulus of continuity $\rho^{(\xi)}$ such that for all $\omega' \in \Omega$ and $\tilde{\omega} \in \Omega^t$

$$\begin{aligned} |\xi(\omega) - \xi(\omega')| &\leq \rho^{(\xi)}(\|\omega - \omega'\|_{[0,T]}), \text{ and} \\ |\xi^{t,\omega}(\tilde{\omega}) - \xi^{t,\omega'}(\tilde{\omega})| &\leq \rho^{(\xi)}(\|\omega - \omega'\|_{[0,t]}). \end{aligned}$$

Consider a sequence (ω_n) such that $\|\omega - \omega_n\|_{[0,t]} \rightarrow 0$. For any $P \in \mathcal{P}(t, \omega)$, by the Uniform Continuity assumption for (Θ_t) , we know that $P \in \mathcal{P}(t, \omega_n)$ when n is large enough. Thus

$$\begin{aligned} &\liminf_{n \rightarrow \infty} v_t(\omega_n) \\ &= \liminf_{n \rightarrow \infty} \sup_{P' \in \mathcal{P}(t, \omega_n)} E^{P'} \xi^{t, \omega_n} \\ &\geq \liminf_{n \rightarrow \infty} \left(\left(\sup_{P' \in \mathcal{P}(t, \omega_n)} E^{P'} \xi^{t, \omega} \right) - \rho^{(\xi)}(\|\omega - \omega_n\|_{[0,t]}) \right) \\ &= \sup_{P' \in \mathcal{P}(t, \omega_n)} E^{P'} \xi^{t, \omega} \geq E^P \xi^{t, \omega}. \end{aligned}$$

Because $P \in \mathcal{P}(t, \omega)$ is arbitrary, this proves that $\liminf_{n \rightarrow \infty} v_t(\omega_n) \geq v_t(\omega)$, which is the asserted lower semicontinuity.

Next prove that any bounded lower semicontinuous function f on ${}^t\Omega$ is in $\widehat{L}^2({}^t\Omega)$: Because ${}^t\Omega$ is Polish, there exists a uniformly bounded sequence $f_n \in C_b({}^t\Omega)$ such that $f_n \uparrow f$ for all $\omega \in \Omega$. By Gihman and Skorohod [12, Theorem 3.10], \mathcal{P} is relatively compact in the weak convergence topology. Therefore, by Tietze's Extension Theorem (Mandelkern [24]), $C_b({}^t\Omega) \subset \widehat{L}^2({}^t\Omega)$, and by Denis et al. [4, Theorem 12], $\sup_{P \in \mathcal{P}} E^P (|f - f_n|^2) \rightarrow 0$. Thus $f \in \widehat{L}^2({}^t\Omega)$.

Combine these two results to deduce that $\hat{E}[\xi \mid \mathcal{F}_t] \in \widehat{L}^2({}^t\Omega)$ if $\xi \in UC_b(\Omega)$. From (i), $\{\hat{E}[\xi \mid \mathcal{F}_t] : \xi \in \widehat{L}^2(\Omega)\}$ is contained in the $\|\cdot\|$ -closure of $\{\hat{E}[\xi \mid \mathcal{F}_t] : \xi \in UC_b(\Omega)\}$. But $\{\hat{E}[\xi \mid \mathcal{F}_t] : \xi \in UC_b(\Omega)\}$ is contained in $\widehat{L}^2({}^t\Omega)$, which is complete under $\|\cdot\|$. This completes the proof. \blacksquare

Proof of (B.2): Fix $P \in \mathcal{P}$ and $X \in \widehat{L}^2(\Omega)$. Given $\epsilon > 0$, there exists $\bar{X} \in UC_b(\Omega)$ such that

$$\|\hat{E}[X \mid \mathcal{F}_t] - \hat{E}[\bar{X} \mid \mathcal{F}_t]\| \leq \|X - \bar{X}\| \leq \epsilon.$$

For any $P' \in \mathcal{P}(t, P)$,

$$\begin{aligned} & E^{P'}[X | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t] \\ = & E^{P'}[X - \bar{X} | \mathcal{F}_t] + (E^{P'}[\bar{X} | \mathcal{F}_t] - \hat{E}[\bar{X} | \mathcal{F}_t]) + (\hat{E}[\bar{X} | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t]). \end{aligned} \quad (\text{B.23})$$

From Karatzas and Shreve [20, Theorem A.3], derive that there exists a sequence $P_n \in \mathcal{P}(t, P)$ such that

$$\operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[\bar{X} | \mathcal{F}_t] = \lim_{n \rightarrow \infty} E^{P_n}[\bar{X} | \mathcal{F}_t] \quad P\text{-a.e.} \quad (\text{B.24})$$

where $P\text{-a.e.}$ the sequence on the right is increasing in n . Then by Lemma B.7,

$$\hat{E}[\bar{X} | \mathcal{F}_t] = \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[\bar{X} | \mathcal{F}_t] = \lim_{n \rightarrow \infty} E^{P_n}[\bar{X} | \mathcal{F}_t] \quad P\text{-a.e.} \quad (\text{B.25})$$

Denote $L^2(\Omega, \mathcal{F}_T, P)$ by $L^2(P)$. Compute $L^2(P)$ -norms on both sides of (B.23) to obtain, for every n ,

$$\begin{aligned} & \| E^{P_n}[X | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t] \|_{L^2(P)} \\ \leq & \| X - \bar{X} \|_{L^2(P_n)} + \| E^{P_n}[\bar{X} | \mathcal{F}_t] - \hat{E}[\bar{X} | \mathcal{F}_t] \|_{L^2(P)} + \| \hat{E}[\bar{X} | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t] \|_{L^2(P)} \\ \leq & \| E^{P_n}[\bar{X} | \mathcal{F}_t] - \hat{E}[\bar{X} | \mathcal{F}_t] \|_{L^2(P)} + 2\epsilon \end{aligned}$$

By (B.25),

$$\limsup_{n \rightarrow \infty} \| E^{P_n}[X | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t] \| \leq 2\epsilon.$$

Note that ϵ is arbitrary. Therefore, there exists a sequence $\hat{P}_n \in \mathcal{P}(t, P)$ such that $E^{\hat{P}_n}[X | \mathcal{F}_t] \rightarrow \hat{E}[X | \mathcal{F}_t]$, $P\text{-a.e.}$, which implies that

$$\hat{E}[X | \mathcal{F}_t] \leq \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[X | \mathcal{F}_t]. \quad (\text{B.26})$$

As in (B.24), there exists a sequence $P'_n \in \mathcal{P}(t, P)$ such that

$$\operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[X | \mathcal{F}_t] = \lim_{n \rightarrow \infty} E^{P'_n}[X | \mathcal{F}_t] \quad P\text{-a.e.}$$

with the sequence on the right being increasing in n ($P\text{-a.e.}$). Set

$A_n \triangleq \{E^{P'_n}[X | \mathcal{F}_t] \geq \hat{E}[X | \mathcal{F}_t]\}$. By (B.26), $P\text{-a.e.}$

$$0 \leq (E^{P'_n}[X | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t])1_{A_n} \nearrow \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[X | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t].$$

By (B.23) and (B.25), P -a.e.

$$E^{P'_n}[X | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t] \leq E^{P'_n}[X - \bar{X} | \mathcal{F}_t] + (\hat{E}[\bar{X} | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t])$$

Take $L^2(P)$ -norms to derive

$$\begin{aligned} & \left\| \operatorname{ess\,sup}_{P' \in \mathcal{P}(t, P)} E^{P'}[X | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t] \right\|_{L^2(P)} \\ &= \lim_{n \rightarrow \infty} \left\| (E^{P'_n}[X | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t]) 1_{A_n} \right\|_{L^2(P)} \\ &\leq \limsup_{n \rightarrow \infty} \left\| X - \bar{X} \right\|_{L^2(P'_n)} + \left\| \hat{E}[\bar{X} | \mathcal{F}_t] - \hat{E}[X | \mathcal{F}_t] \right\|_{L^2(P)} \\ &\leq 2\epsilon. \end{aligned}$$

This proves (B.2).

Proof of (2.12): It is sufficient to prove that, for $0 \leq s \leq t \leq T$, P -a.e.

$$\begin{aligned} & \operatorname{ess\,sup}_{P' \in \mathcal{P}(s, P)} E^{P'}[X | \mathcal{F}_s] \\ &= \operatorname{ess\,sup}_{P' \in \mathcal{P}(s, P)} E^{P'} \left[\operatorname{ess\,sup}_{P'' \in \mathcal{P}(t, P')} E^{P''}[X | \mathcal{F}_t] \mid \mathcal{F}_s \right] \end{aligned} \quad (\text{B.27})$$

The classical law of iterated expectations implies the inequality \leq in (B.27). Next prove the reverse inequality.

As in (B.24), there exists a sequence $P''_n \in \mathcal{P}(t, P')$ such that P' -a.e.

$$\lim_{n \rightarrow \infty} E^{P''_n}[X | \mathcal{F}_t] \uparrow \operatorname{ess\,sup}_{P'' \in \mathcal{P}(t, P')} E^{P''}[X | \mathcal{F}_t]$$

Then

$$\begin{aligned} & E^{P'} \left[\operatorname{ess\,sup}_{P'' \in \mathcal{P}(t, P')} E^{P''}[X | \mathcal{F}_t] \mid \mathcal{F}_s \right] \\ &= \lim_{n \rightarrow \infty} E^{P''_n}[X | \mathcal{F}_s] \\ &\leq \operatorname{ess\,sup}_{\tilde{P} \in \mathcal{P}(s, P)} E^{\tilde{P}}[X | \mathcal{F}_s] \quad P\text{-a.e.} \end{aligned}$$

This proves (2.12).

Proof properties (i)-(iv) is standard and is omitted. This completes the proof of Theorem 2.6. \blacksquare

C. Appendix: Proofs for Utility

Proof of Theorem 2.7: Part (a). Consider the following backward stochastic differential equation under \hat{E} :

$$V_t = \hat{E}[\xi + \int_t^T f(c_s, V_s) ds \mid \mathcal{F}_t], \quad t \in [0, T],$$

where $\xi \in \widehat{L}^2(\Omega_T)$ is terminal utility and $c \in D$. (Equation (2.13) is the special case where $\xi = 0$.) Given $V \in M^2(0, T)$, let

$$\Lambda_t(V) \equiv \hat{E}[\xi + \int_t^T f(c_s, V_s) ds \mid \mathcal{F}_t], \quad t \in [0, T].$$

We need the following regularity property of Λ .

Lemma C.1. Λ is a mapping from $M^2(0, T)$ to $M^2(0, T)$.

Proof. By assumption (ii) there exists a positive constant K such that

$$|f(c_s, V_s) - f(c_s, 0)| \leq K |V_s|, \quad s \in [0, T].$$

Because both V and $(f(c_s, 0))_{0 \leq s \leq T}$ are in $M^2(0, T)$, we have $(f(c_s, V_s))_{0 \leq s \leq T} \in M^2(0, T)$. Thus

$$\begin{aligned} & \hat{E}[|\xi + \int_t^T f(c_s, V_s) ds|^2] \\ & \leq 2\hat{E}[|\xi|^2 + (T-t) \int_t^T |f(c_s, V_s)|^2 ds] \\ & \leq 2\hat{E}[|\xi|^2] + 2(T-t)\hat{E}[\int_t^T |f(c_s, V_s)|^2 ds] < \infty, \end{aligned}$$

which means that $(\xi + \int_t^T f(c_s, V_s) ds) \in \widehat{L}^2(\Omega)$. Argue further that

$$\begin{aligned} \hat{E}[|\Lambda_t(V)|^2] &= \hat{E}[(\hat{E}[\xi + \int_t^T f(c_s, V_s) ds \mid \mathcal{F}_t])^2] \\ &\leq \hat{E}[\hat{E}[|\xi + \int_t^T f(c_s, V_s) ds|^2 \mid \mathcal{F}_t]] \\ &= \hat{E}[|\xi + \int_t^T f(c_s, V_s) ds|^2] < \infty. \end{aligned}$$

Finally, $(\hat{E} \int_0^T |\Lambda_t|^2 dt)^{\frac{1}{2}} \leq (\int_0^T \hat{E}[|\Lambda_t|^2] dt)^{\frac{1}{2}} < \infty$ and $\Lambda \in M^2(0, T)$. ■

For any V and $V' \in M^2(0, T)$, by Theorem 2.6 the following approximation holds:

$$\begin{aligned}
& | \Lambda_t(V) - \Lambda_t(V') |^2 \\
&= (\widehat{E}[\int_t^T (f(c_s, V_s) - f(c_s, V'_s)) ds | \mathcal{F}_t])^2 \\
&\leq \widehat{E}[(\int_t^T |f(c_s, V_s) - f(c_s, V'_s)| ds)^2 | \mathcal{F}_t] \\
&\leq (T - t)K^2 \widehat{E}[\int_t^T |V_s - V'_s|^2 ds | \mathcal{F}_t] \\
&\leq L \widehat{E}[\int_t^T |V_s - V'_s|^2 ds | \mathcal{F}_t],
\end{aligned} \tag{C.1}$$

where K is the Lipschitz constant for the aggregator and $L = TK^2$. (The first inequality is due to the classical Jensen's inequality and (B.2).) Then for each $r \in [0, T]$,

$$\begin{aligned}
& \widehat{E}[\int_r^T | \Lambda_t(V) - \Lambda_t(V') |^2 dt] \\
&\leq L \widehat{E}[\int_r^T \widehat{E}[\int_t^T |V_s - V'_s|^2 ds | \mathcal{F}_t] dt] \\
&\leq L \int_r^T \widehat{E}[\int_t^T |V_s - V'_s|^2 ds] dt \\
&\leq L(T - r) \widehat{E}[\int_r^T |V_s - V'_s|^2 ds].
\end{aligned}$$

Set $\delta = \frac{1}{2L}$ and $r_1 = \max\{T - \delta, 0\}$. Then,

$$\widehat{E}[\int_{r_1}^T | \Lambda_t(V) - \Lambda_t(V') |^2 dt] \leq \frac{1}{2} \widehat{E}[\int_{r_1}^T |V_t - V'_t|^2 dt],$$

which implies that Λ is a contraction mapping from $M^2(r_1, T)$ to $M^2(r_1, T)$ and there exists a unique solution $(V_t) \in M^2(r_1, T)$ to the above BSDE. Because δ is independent of t , we can work backwards in time and apply a similar argument at each step to prove that there exists a unique solution $(V_t) \in M^2(0, T)$.

Part (b). Uniqueness of the solution is due to the contraction mapping property established in the proof of (a).

By Theorem 2.6,

$$\begin{aligned}
& -\widehat{E}\left[-\int_t^T f(c_s, V_s) ds - V_\tau \mid \mathcal{F}_t\right] \\
&= -\widehat{E}\left[-\int_t^T f(c_s, V_s) ds + \widehat{E}\left[-\int_\tau^T f(c_s, V_s) ds \mid \mathcal{F}_\tau\right] \mid \mathcal{F}_t\right] \\
&= -\widehat{E}\left[\widehat{E}\left[-\int_t^T f(c_s, V_s) ds \mid \mathcal{F}_\tau\right] \mid \mathcal{F}_t\right] \\
&= -\widehat{E}\left[-\int_t^T f(c_s, V_s) ds \mid \mathcal{F}_t\right] = V_t. \quad \blacksquare
\end{aligned}$$

References

- [1] Anderson, E., L.P. Hansen, and T.J. Sargent, A quartet of semigroups for model specification, robustness, prices of risk and model detection, *J. Europ. Econ. Assoc.* 1 (2003), 68-123.
- [2] Bion-Nadal, J., and M. Kervarec, Risk measuring under model uncertainty, *Ann. Appl. Probab.* 22 (2012), 213-238.
- [3] Chen, Z., and L.G. Epstein, Ambiguity, risk, and asset returns in continuous time, *Econometrica* 70 (2002), 1403-1443.
- [4] Denis, L., M. Hu, and S. Peng, Function spaces and capacity related to a sublinear expectation: application to G-Brownian motion paths, *Potential Anal.* 34 (2011), 139-161.
- [5] Denis, L., and M. Kervarec, Utility functions and optimal investment in non-dominated models, 2009, <http://hal.archives-ouvertes.fr/hal-00371215/en/>.
- [6] Denis, L., and C. Martini, A theoretical framework for the pricing of contingent claims in the presence of model uncertainty, *Ann. Appl. Probab.* 16 (2006), 827-852.
- [7] Duffie, D., and L.G. Epstein, Stochastic differential utility, *Econometrica* 60 (1992), 353-394 (Appendix C with C. Skiadas).
- [8] Epstein, L.G., and S. Ji, Ambiguous volatility and asset pricing in continuous time, *Rev. Finan. Stud.* 26 (2013), 1740-1786.
- [9] Epstein, L.G., and M. Schneider, Recursive multiple-priors, *J. Econ. Theory* 113 (2003), 1-31.
- [10] Epstein, L.G., and M. Schneider, Ambiguity and asset markets, *Ann. Rev. Finan. Econ.* 2 (2010), 315-346.
- [11] Follmer, H., Calcul d'Ito sans probabilités, *Seminar on Probability XV, Lecture Notes in Mathematics* 850 (1981), 143-150.
- [12] Gihman, I. I., and A. V. Skorohod, *Controlled Stochastic Processes*. New York: Springer-Verlag, 1979.

- [13] Gilboa, I., and D. Schmeidler, Maxmin expected utility with non-unique priors, *J. Math. Econ.* 18 (1989), 141-153.
- [14] He, S., J. Wang and J. Yan, *Semimartingale Theory and Stochastic Calculus*. Beijing: Science Press and CRC Press, 1992.
- [15] Heston, S.L., A closed-form solution for options with stochastic volatility with application to bond and currency options, *Rev. Finan. Stud.* 6 (1993), 293-326.
- [16] Hu, M., and S. Ji, Explicit solutions of linear BSDEs under G-expectations and applications, 2012.
- [17] Hull, J., and A. White, The pricing of options on assets with stochastic volatilities, *J. Finan.* 3 (1987), 281-300.
- [18] Ikeda, N., and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*. New York: Elsevier, 1989.
- [19] Karandikar, R., On pathwise stochastic integration, *Stoch. Process. Appl.* 57 (1995), 11-18.
- [20] Karatzas, I., and S. Shreve, *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag, 1991.
- [21] Levy, A., M. Avellaneda and A. Paras, A new approach for pricing derivative securities in markets with uncertain volatilities: a “case study” on the trinomial tree, 1998,
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=5342.
- [22] Li, X., and S. Peng, Stopping times and related Itô’s calculus with G-Brownian motion, *Stoch. Process. Appl.* 121 (2011), 1492-1508.
- [23] Lipster, R.S., and A.N. Shiryaev, *Statistics of Random Processes I*, Springer, 1978.
- [24] Mandelkern, M., On the uniform continuity of tietze extensions," *Archiv der Mathematik* 55 (1990), 387-388.
- [25] Nutz, M. Random G-expectations, *Ann. Appl. Probab.* forthcoming.

- [26] Nutz, M. A quasi-sure approach to the control of non-Markovian stochastic differential equations, *Electronic J. Probab.* 17 (2012), 1-23.
- [27] Oksendal, B., *Stochastic Differential Equations*. Berlin: Springer, 2005.
- [28] Peng, S., BSDE and stochastic optimizations, in *Topics in Stochastic Analysis (Ch.2, Chinese vers.)*, ed. by Yan, J., Peng, S., Fang, S., and Wu, L.M., Beijing: Science Press, 1997.
- [29] Peng, S., Filtration consistent nonlinear expectations and evaluations of contingent claims, *Acta Mathematicae Applicatae Sinica, English series*, 20 (2004), 1-24.
- [30] Peng, S., Nonlinear expectations and nonlinear Markov chains, *Chinese Annals of Math.* 26B (2005), 159-184.
- [31] Peng, S., G-expectation, G-Brownian motion and related stochastic calculus of Itô's type, in *The Abel Symposium 2005, Abel Symposia 2*, ed. by Benth et. al., Springer-Verlag, (2006), 541-567.
- [32] Peng, S., Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation, *Stoch. Process. Appl.* 118 (2008), 2223-2253.
- [33] Peng, S., Nonlinear expectations and stochastic calculus under uncertainty, 2010, <http://arxiv.org/abs/1002.4546>.
- [34] Soner, M., N. Touzi, and J. Zhang, Martingale representation theorem under G-expectation, *Stoch. Process. Appl.* 121 (2011), 265-287.
- [35] Soner, M., N. Touzi, and J. Zhang, Quasi-sure stochastic analysis through aggregation, *Electron. J. Probab.* 16 (2011), 1844-1879.
- [36] Soner, M., N. Touzi, and J. Zhang, Wellposedness of second order backward SDEs, *Probab. Theory and Related Fields*, forthcoming.
- [37] Soner, M., N. Touzi, and J. Zhang, Dual formulation of second order target problems, *Annals Appl. Probab.*, forthcoming.
- [38] Song, Y., Some properties of G-evaluation and its applications to G-martingale decomposition, *Science China-Mathematics* 54 (2011), 287-300.
- [39] Yong, J., and X. Zhou, *Stochastic Controls*. New York: Springer, 1999.