Robust Utility in Continuous Time*

Patrick Beissner

Research School of Economics, Australian National University

Fabio Maccheroni

Departmente of Decision Sciencies and IGIER, Università Bocconi

Massimo Marinacci

Departmente of Decision Sciencies and IGIER, Università Bocconi

Sujoy Mukerji

School of Economics and Finance, Queen Mary University of London

April 17, 2023

Abstract

We study a general class of utility processes $V(\mathbf{c}) = (V_t(\mathbf{c}))$, where $V_t(\mathbf{c})$, a dynamic utility operator, is a decision criterion that quantifies a decision maker's evaluation of uncertain consumption streams $\mathbf{c}$. We call this dynamic utility operator *robust* and its distinctiveness is that it features the diffusion of the process $V(\mathbf{c})$, i.e., the utility is affected by *its own* variability. A main result of this paper is to identify a general class of robust dynamic utility operators that are monotone and, yet, irreducibly depend on the utility variability. A principal motivation for studying such robust dynamic operators is that, by incorporating utility variability into the decision criterion, they bring a facility required to adapt models of ambiguity sensitive preferences to Brownian environments. In particular, those preference models which permit flexibility in ambiguity attitudes. We demonstrate this facility by obtaining continuous-time extensions of two prominent ambiguity aversion frameworks which incorporate variable ambiguity attitude, the smooth ambiguity model and the α-maxmin expected utility.

*We thank several seminar audiences as well as Anne Balter, Masaki Fukasawa, Lars Peter Hansen, Chiaki Hara, Jianjun Miao, Yuri Sannikov, Todd Sarver, Jun Sekine and Bruno Strulovici for useful comments.

1 Introduction

Model ambiguity, that is, uncertainty about the data generating process, may evolve over time in ways that enrich and enliven decision environments. While learning may diminish perceived ambiguity, it does not necessarily do so. News that can make a decision maker suspect that the data generating process relevant for the near future is likely to be different from that (believed to have been) operating in the recent past, may well increase ambiguity. The unfolding of the financial and economic crises over the past decade, signals of global warming and the more recent pandemic provide obvious, macro-level examples. Indeed, several of these have registered as events of "uncertainty shock" on uncertainty indices (e.g., Jurado et al., 2014). Arrival of such information, from multiple sources and agencies, is pretty relentless and near-continuous. Studying dynamics of decision-making sensitive to (possibly) continuous innovations in ambiguity in the decision environment, is of interest.

Recursions We study a general class of utility processes $V(\mathbf{c}) = (V_t(\mathbf{c}))$ that evaluate a consumption stream $\mathbf{c} = (\mathbf{c}_t)$ and are recursively defined as maximal bounded solutions of a forward recursion:

$$V_t(\mathbf{c}) = \mathrm{E}_t\left[\int_t^T f(s, \mathbf{c}_s, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c})))\mathrm{d}s\right] \tag{1}$$

where $\sigma_s(V(\mathbf{c}))$ is the diffusion of the utility process that indexes its variability and the random map f, a temporal *aggregator*, gives the random change in continuation utility at time s. In other words, $V_t(\mathbf{c})$, a dynamic utility operator, is a decision criterion that quantifies the decision maker's evaluation of uncertain consumption streams $\mathbf{c}$. We call this dynamic utility operator *robust* and its distinctiveness is that it features the diffusion of process $V(\mathbf{c})$, i.e., the utility is affected by *its own* variability. A principal motivation for studying robust dynamic operators is that, by incorporating utility variability into the decision criterion, they bring a facility required to adapt models of ambiguity sensitive preferences that permit flexibility in ambiguity attitude, originally constructed for time-less environments, to Brownian environments. We demonstrate this facility by obtaining continuous-time extensions of two prominent ambiguity aversion frameworks which incorporate variable ambiguity attitude, the smooth ambiguity model and the α-maxmin expected utility. As will be shown, robustness adjustment incorporated in the dynamic utility is an ambiguity sensitive decision maker's response to *model ambiguity*, i.e., to ambiguous uncertainty over models.[1]

Since robust dynamic utility operators (1) extend classical discounted utility,

$$V_t(\mathbf{c}) = \mathrm{E}_t\left[\int_t^T e^{-\beta(s-t)}u(\mathbf{c}_s)\mathrm{d}s\right] = \int_t^T e^{-\beta(s-t)}\mathrm{E}_t\left[u(\mathbf{c}_s)\right]\mathrm{d}s,$$

it might be helpful to understand how the recursion (1) shapes up in this special case. Consider the equation,

$$V_t(\mathbf{c}) = u(\mathbf{c}_t)\,\mathrm{d}t + \delta(\mathrm{d}t)\,\mathcal{M}_t(V_{t+\mathrm{d}t}(\mathbf{c})) \tag{2}$$

The future utility is a random variable from the perspective at time t, so a stochastic operator, $\mathcal{M}_t$, is needed to convert uncertain utils $V_{t+\mathrm{d}t}(\mathbf{c})$ at time $t+\mathrm{d}t$ into equivalent certain utils conditional on information at time t before applying the discount function $\delta(\cdot)$. Setting $\delta(\mathrm{d}t) = (1+\beta\mathrm{d}t)^{-1}$ in (2) we get

$$\mathcal{M}_t(V_{t+\mathrm{d}t}(\mathbf{c})) - V_t(\mathbf{c}) \approx -(u(\mathbf{c}_t) - \beta V_t(\mathbf{c}))\,\mathrm{d}t.$$

[1] Cf. Hansen and Marinacci (2016).

A maintained assumption in this paper, descriptive of our Brownian setting, is that the utility process $V(\mathbf{c})$ is a bounded Ito process,

$$\mathrm{d}V_t(\mathbf{c}) = \mu_t(V(\mathbf{c}))\,\mathrm{d}t + \sigma_t(V(\mathbf{c}))\,\mathrm{d}B_t \tag{3}$$

for all streams $\mathbf{c}$, where $\mu_t(\cdot)$ and $\sigma_t(\cdot)$ denote the drift and the diffusion, respectively. If the decision maker is an expected utility maximizer the stochastic operator $\mathcal{M}_t$ is the basic conditional expectation, $\mathcal{M}_t(V_{t+\mathrm{d}t}(\mathbf{c})) = \mathrm{E}_t[V_{t+\mathrm{d}t}(\mathbf{c})]$. So, in this case

$$\begin{aligned}
\mathcal{M}_t(V_{t+\mathrm{d}t}(\mathbf{c})) - V_t(\mathbf{c}) &= \mathrm{E}_t[V_{t+\mathrm{d}t}(\mathbf{c})] - V_t(\mathbf{c}) = \mathrm{E}_t[V_t(\mathbf{c}) + \mathrm{d}V_t(\mathbf{c})] - V_t(\mathbf{c}) \\
&= \mathrm{E}_t[\mathrm{d}V_t(\mathbf{c})] = \mu_t(V(\mathbf{c}))\,\mathrm{d}t \\
&\Leftrightarrow \mu_t(V(\mathbf{c}))\,\mathrm{d}t \approx -(u(\mathbf{c}_t) - \beta V_t(\mathbf{c}))\,\mathrm{d}t.
\end{aligned}$$

Hence, equation (2) is solved by an Ito value process that satisfies the backward stochastic differential equation (BSDE),

$$\mathrm{d}V_t(\mathbf{c}) = -(u(\mathbf{c}_t) - \beta V_t(\mathbf{c}))\,\mathrm{d}t + \sigma_t(V(\mathbf{c}))\,\mathrm{d}B_t \tag{4}$$

with terminal condition $V_T(\mathbf{c}) = 0$. A key result we establish proves equivalence between forward recursions and quadratic BSDEs. This result is the engine room, as it were, driving the paper's more significant conclusions. An application of this general result shows that the *aggregator* f in (1) corresponding to this special case, discounted expected utility, is simply $u(\mathbf{c}_t) - \beta V_t(\mathbf{c})$. The general class of aggregators we consider is given by

$$f(s, \mathbf{c}_s, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c}))) = u(\mathbf{c}_s) - \beta V_s(\mathbf{c}) + \eta(s, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c}))) \tag{5}$$

So, in the traditional expected utility case just considered we have $\eta = 0$: the aggregator does not include the *variability term*, $\eta(s, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c})))$, which may be quadratic and comes into play when we bring in non-neutral attitudes to ambiguity.

Ambiguity and variability Ambiguity is embodied in the model uncertainty perceived by the decision maker: they are unsure which would be the right probabilistic model to apply to evaluate (contingent) consumption streams and they keep in consideration a set of alternative probabilistic models. Intuitively, the desired robustness adjustment (incorporated in a robust dynamic utility) arises from an ambiguity averse decision maker's concern that their subjective belief is not accurate. As will be shown, the *revealed* likelihood an ambiguity sensitive decision maker attaches to a particular probabilistic model is interval valued (rather than a point value, as in the case of Savage-type ambiguity neutral decision maker).

Ambiguity aversion in the face of model uncertainty provides one foundation to the variability term that distinctively figures in the robust dynamic utility. It is the foundation we focus on in this paper. The underpinning intuition is that if the value of a (contingent) consumption stream varies considerably depending on which probability model (among the alternatives considered as possible) is applied, then its robust utility evaluation will be adjusted significantly too. In our setting, model uncertainty about possible probability models is formalized as the decision maker being unsure of the drift term in the Ito process (3). Hence, the stochastic operator $\mathcal{M}_t$ in this case needs to contend with an additional layer of uncertainty: multiplicity of expectations of the future utility, $V_{t+\mathrm{d}t}$. Alternative ambiguity aversion frameworks offer alternative ways of aggregating over this second layer of uncertainty giving rise to corresponding, alternative specifications of $\mathcal{M}_t$. When evaluating a contingent consumption stream, common to alternative ambiguity aversion frameworks is the feature that the aggregation will take into account the extent to which expectation of the future utility varies depending on the probability model applied to compute the expectation.[2] In the smooth ambiguity model, for example, the second layer of uncertainty is represented in a Bayesian fashion and ambiguity aversion manifests as a dispreference for spreads in the induced distribution over expected future utilities. As will be shown, this makes the specification of $\mathcal{M}_t$ corresponding to the smooth ambiguity model a decreasing function of the total variance of future utility, obtained by taking both layers of uncertainty into account. This total variance is increasing in $\sigma_s(V(\mathbf{c}))$; which is why the diffusion term appears in the aggregator corresponding to smooth ambiguity.

In our framework the decision maker, when they have no doubts about their subjective belief, behaves as a (discounted) subjective expected utility maximizer, as if facing only risk. This special case corresponds to a dynamic utility operator which features an aggregator f that does not depend on $\sigma_s(V(\mathbf{c}))$, i.e., $\eta = 0$ in (5). The class of aggregators we consider in general assumes that our dynamic utility operators rank deterministic streams via standard discounting. Time and risk, are thus treated, per se, in a conventional way as we are primarily interested in issues that are peculiar to (ambiguous) uncertainty. In comparison, the recursions studied by Duffie and Epstein (1992) correspond to aggregators that have $\eta = 0$, but allow non-separabilities in the treatment of time and risk. Thus there, the term $u(\mathbf{c}_s) - \beta V_s(\mathbf{c})$ takes a more general form $g(s, c_s, V_s(\mathbf{c}))$. Aggregators f that, like ours, depend on a variability variable were first introduced by Lazrak and Quenez (2003). Their insightful analysis requires, however, Lipschitzian conditions that are too restrictive for our general analysis in which the variability component might well enter quadratically in the specifications that we consider.

Monotonicity and variability A key part of our analysis establishes that the robust dynamic utility operator is monotone and satisfies dynamic consistency. A basic primitive in our analysis, alongside standard discounting of deterministic streams, is an atemporal (instantaneous) *utility function* $u : C \to \mathbb{R}$ that ranks and evaluates quantities of a material consequence $c \in C$, say a consumption good. It is cardinal (i.e., unique up to affine transformations), in keeping with the presumption across ambiguity aversion frameworks we wish to accommodate that decision makers are expected utility maximizers when facing pure risk. Monotonicity, a basic tenet of rationality, ensures that a robust dynamic utility operator is consistent with the posited atemporal utility: two consumption streams that are equivalent, at each point of time, in terms of atemporal utility have the same dynamic utility. Here, it is an especially noteworthy property because of the dependence of the robust dynamic utility operator on variability, through $\sigma_s(V)$.

Indeed, as is well-known, decision criteria that depend on the valuation variability are in general not monotone. To illustrate this important issue in the simplest possible way, consider the collection $\Delta(X)$ of lotteries p defined on a prize space X. Given a utility function $u : X \to \mathbb{R}$, with interval image $\operatorname{Im} u = (a, b)$, define the expected utility $\mathbb{E}_u : \Delta(X) \to \mathbb{R}$ and the standard deviation utility

[2] This is evident in the representations of not only α-maxmin expected utility (Ghirardato et al., 2004) and smooth ambiguity model (Klibanoff et al., 2005), but also other frameworks such as variational preferences (Maccheroni et al., 2006) and vector expected utility (Siniscalchi, 2009).

$\sigma_u : \Delta(X) \to \mathbb{R}$ functionals by[3]

$$\mathbb{E}_u(p) = \sum_{x \in \operatorname{supp} p} u(x)\, p(x) \quad \text{and} \quad \sigma_u(p) = \sqrt{\sum_{x \in \operatorname{supp} p} \left(u(x) - \mathbb{E}_u(p)\right)^2 p(x)} \tag{6}$$

Here $\sigma_u(p)$ indexes the *valuation* variability. Given a function $\lambda : (a, b) \to \mathbb{R}$, define a mean-variance criterion $M_u : \Delta(X) \to \mathbb{R}$ in utils by

$$M_u(p) = \mathbb{E}_u(p) - \frac{\lambda(\mathbb{E}_u(p))}{2} \sigma_u^2(p) \tag{7}$$

Unlike expected utility, this mean-variance criterion takes into account the variability of the valuation, not just its expectation.[4] Setting aside the normative appeal of the Independence Axiom, a basic observation in favor of expected utility, when compared to the criterion (7), is that the latter violates monotonicity, in the form of first-order stochastic dominance, arguably a basic tenet of rationality. Yet, monotonicity is restored when only lotteries p that have a small enough valuation variability $\sigma_u^2(p)$ are compared. The mean-variance criterion (7) might, therefore, have an "asymptotic" scope. For our purposes, this suggests that a continuous-time setting might be appropriate to study decision criteria that account for the valuation variability and still preserve monotonicity.

Yet, a folk result says that there exists a function $v : X \to \mathbb{R}$, ordinally equivalent to u, such that

$$M_u(p) = \mathbb{E}_v(p) + o\left(\sigma_u^2(p)\right)$$

for each lottery p.[5] So, the relevance of variability vanishes faster than its magnitude. For instance, for a "Brownianish" lottery b that pays the two prizes $\pm 1/\sqrt{n}$ with equal probability, we have $\mathbb{E}_v(b) = 0$ and $\sigma_u^2(b) = 1/n$, so that

$$M_u(b) = \mathbb{E}_v(b) + o\left(\frac{1}{n}\right) \approx \mathbb{E}_v(b)$$

Here the mean-variance criterion (7) is well approximated by a standard expected utility criterion.

This folk result casts some serious doubts on the asymptotic status of the mean-variance criterion (7). For our analysis, it indicates that in continuous time the role of the valuation variability might turn out to be fragile. Indeed, the ordinal transformation underlying this folk result is, essentially, that used by Duffie and Epstein (1992), p. 1409, to eliminate a variability term from their stochastic differential utility. Such transformations, we show, are no longer possible for our robust dynamic utility operator defined via the forward recursion (1) and consistent with exponential discounting and a cardinal atemporal utility function. For such operators the role of variability is thus not eliminable. It is, therefore, a main result of this paper to identify a general class of robust dynamic utility operators that are monotone and, yet, irreducibly depend on the valuation variability. Monotonicity (and dynamic consistency) is shown to follow from our engine room result, the equivalence between forward recursions like (1) and quadratic backward stochastic differential equations.

[3] As usual, $\operatorname{supp} p$ denotes the support $\{x \in X : p(x) > 0\}$ of lottery p.

[4] This criterion is, of course, relatively rudimentary compared to the utility defined in (1), however, they share a conceptual heart: non-separability. While $M_u(p)$, in general, violates the classic independence axiom (for risk), the utility defined in (1) will, in general, violate Savage's axiom P2, the so-called Sure Thing Principle (as ambiguity sensitive preferences would).

[5] See de Finetti (1952) pp. 16-18 (trans. pp. 270-272).

Specifications Two special cases of the robust aggregator f, in (5), are heuristically derived within our setting in Section 5. It is via these derivations that we obtain robust dynamic utility specifications corresponding to two different ambiguity aversion preference frameworks that allow for variable ambiguity attitude. Each derivation proceeds along the line of argument sketched earlier to explain the form the aggregator takes in the case of discounted expected utility and relies on our engine room result in the crucial final step.

One case is when the robust aggregator has the form

$$f(s, \mathbf{c}_s, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c})) = u(\mathbf{c}_s) - \beta V_s(\mathbf{c}) - \frac{1}{2}\gamma_s \sigma_s(V(\mathbf{c})) + \frac{1}{2}\kappa_s(2\alpha_s - 1)|\sigma_s(V(\mathbf{c}))| \tag{8}$$

where $\kappa_s \geq 0$ accounts for ambiguity aversion a la Gilboa and Schmeidler (1989). When $\alpha_t = 1$, we have a specification a la Chen and Epstein (2002); in the case κ_s is time invariant and γ_s is null, it reduces to their leading example of κ-ignorance. Going beyond Chen and Epstein (2002), we allow for $\alpha_s \in (0, 1)$ and derive a continuous-time version of α-maxmin expected utility that accommodates more flexibility in ambiguity attitude.

The second form of robust aggregator that we provide a heuristic foundation for is,

$$f(s, \mathbf{c}_s, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c})) = u(\mathbf{c}_s) - \beta V_s(\mathbf{c}) - \frac{1}{2}\upsilon\varsigma_s^2\sigma_s^2(V(\mathbf{c})) \tag{9}$$

Our heuristic development shows ς_s and $\upsilon \geq 0$ account, respectively, for model ambiguity perception and attitude corresponding to the smooth ambiguity model of Klibanoff et al. (2005, 2009). More specifically, we show, the aggregator (9) corresponds to the robust mean-variance criterion that arises via a quadratic approximation of the smooth ambiguity model, as in Maccheroni et al. (2013). Skiadas (2013) (and, also, Hansen and Sargent, 2011) pointed out a difficulty in taking the smooth ambiguity model to environments where information arrives at infinitesimally short time intervals, over which only small incremental risks are resolved (at a time): concern for ambiguity vanishes as these incremental risks tend to the infinitesimal. To address this difficulty, our heuristic development applies a suggestion of Hansen and Sargent (2011) and Hansen and Miao (2018). These papers suggest that the parametric specification of ambiguity aversion be adapted taking into account the frequency of information arrival in the decision environment; thereby, taking a limiting path different to the one considered by Skiadas. We provide a decision theoretic justification and a behavioral interpretation for the suggested parametric specification. Such a parametric adaptation, we argue formally in Section 5.3.3, ensures that ambiguity aversion, behaviorally defined, remains stable and consistent as the smooth ambiguity model (with this parametric specification) is applied, along the limit path, to a sequence of decision environments where time-increments and corresponding increments in risk decrease in size along the sequence, tending to the infinitesimal. Given (9), the recursion (1) becomes

$$V_t(\mathbf{c}) = \mathrm{E}_t\left[\int_t^T \left(u(\mathbf{c}_s) - \beta V_s(\mathbf{c}) - \frac{1}{2}\upsilon\varsigma_s^2\sigma_s^2(V(\mathbf{c}))\right) \mathrm{d}s\right] \tag{10}$$

Notice, the variability term in (9) appears in quadratic form. So, to establish (10) we could not rely on existing results and needed to apply our result proving equivalence between forward recursions and quadratic BSDEs. Finally note, in (10), if the term ς_t were unitary, we would get a robust recursion a la Hansen and Sargent (2001, 2008).

Hansen and Miao (2018, 2022) develop an extension of a version of the smooth ambiguity framework (that allows for non-unitary intertemporal elasticity of substitution) to a Brownian setting. We show, in Section 5.3.2, if the preferences they consider were required to treat time in the standard way (i.e., satisfy exponential discounting) then their extension can be accommodated in the heuristic development we apply. In this case, it would lead to an aggregator that is a special case of the general class (5) of aggregators we consider and therefore the resulting utility recursion is assured of the properties we prove this class (5) satisfies generally, in particular, existence and monotonicity. Section 5.3.2 also clarifies the relationship between this aggregator and the one described by (9).

Dynamic decision problems In a companion paper, Beissner et al. (2023), we study decision problems that feature robust dynamic utility operators as objective functions. Because of their dynamic consistency, we can establish a form of the classic Bellman's optimality principle for dynamic utility operators. To illustrate our analysis of decision problems, we apply the specification (10) to pricing aggregate equity and obtain analytic expressions for equity and risk free returns.

2 Dynamic utility

2.1 Streams

Our decision makers have to evaluate uncertain streams of a material *consequence* (or *outcome*) $c \in C$, say a consumption good, over a *time interval* $[0, T]$ that depend on exogenous *contingencies* $\omega \in \Omega$.[6] Throughout we make the following structural assumption.

A.0 The consequence space C is a a convex subset of a topological vector space, endowed with its Borel σ-algebra.

Uncertainty is governed by a probability measure $\mathbb{Q}$ on Ω, interpreted as the true *generative mechanism* or *model*. Decision makers have a *subjective* (*predictive*) *probability* $\mathbb{P}$ on Ω that quantifies their beliefs about the relative likelihood of the exogenous contingencies. It is their best guess of the true probability $\mathbb{Q}$.

Though $\mathbb{Q}$ is important for interpretation, it is $\mathbb{P}$ the protagonist of our analysis. Formally, we make the following assumption.

A.1 The quartet

$$(\Omega, \mathbb{F}, \mathcal{F}, \mathbb{P})$$

is the filtered probability space of our analysis, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space and $\mathbb{F} = (\mathcal{F}_t)$ is the natural filtration generated by a Brownian motion $B = (B_t)$.

The filtration $\mathbb{F}$ is the information structure available to the decision makers. Throughout the paper, processes are always predictable with respect to this filtration. In particular, this is the case for streams of consequences, as next we assume.

A.2 Stochastic streams of consequences $\mathbf{c} = (\mathbf{c}_t) \in \mathbf{C}$ are C-valued predictable processes.

At each node (t, ω), decision makers thus know the value $\mathbf{c}_t(\omega)$ of a stochastic stream $\mathbf{c}$. In view of A.2, in what follows $\mathbf{C}$ denotes the collection of all C-valued predictable processes $\mathbf{c}$.

[6] Points of time t are always understood to belong to $[0, T]$. For instance, we may just write "for all t" rather than "for all $t \in [0, T]$". Similarly, contingencies ω are always understood to belong to Ω. Background material is in Appendix A, where we also define the formal notions used in the text, like for example the spaces L^0 and L^∞.

2.2 Preferences

Our decision makers have a preference over stochastic streams.

Definition 1 *A* preference $\succcurlyeq$ *on* $\mathbf{C}$ *is a complete and transitive binary relation which is solvable, i.e., each stochastic stream* $\mathbf{c} \in \mathbf{C}$ *has a consequence* $c \in C$ *such that* $\mathbf{c} \sim \mathbf{1}_c$.[7]

This preference subsumes a preference over consequences: with a standard abuse of notation, for any pair of consequences $c, c' \in C$ we write

$$c \succcurlyeq c'$$

when we have $\mathbf{1}_c \succcurlyeq \mathbf{1}_{c'}$ for the corresponding constant streams.

We endow the collection of all preferences on $\mathbf{C}$ with the σ-algebra generated by the sets $\{\succeq : \mathbf{c} \succeq \mathbf{c}'\}$ of the preferences on $\mathbf{C}$ that, given any two streams, rank one above the other. This permits us to introduce preference processes.

Definition 2 *A* preference process $\precsim$ *is a predictable process that associates to each node* (t, ω) *a preference* $\precsim_{t,\omega}$ *on* $\mathbf{C}$.

Predictability ensures that, at each node (t, ω), decision makers know their own preference $\precsim_{t,\omega}$ over stochastic streams. In particular, given any two stochastic streams $\mathbf{c}$ and $\mathbf{c}'$ we write

$$\mathbf{c} \precsim \mathbf{c}'$$

when, for almost all nodes (t, ω), we have $\mathbf{c} \precsim_{t,\omega} \mathbf{c}'$.[8]

Decision makers' evaluation of material consequences is invariant across nodes, a standard assumption that we adopt next to ease the analysis.

A.3 There exists a preference $\dot{\precsim}$ on consequences such that, at each node (t, ω),

$$c \precsim_{t,\omega} c' \Longleftrightarrow c \mathrel{\dot{\precsim}} c'$$

for all consequences $c, c' \in C$.

In view of this assumption, given any two stochastic streams $\mathbf{c}$ and $\mathbf{c}'$ we write

$$\mathbf{c} \mathrel{\dot{\precsim}} \mathbf{c}'$$

when, for almost all nodes (t, ω), we have $\mathbf{c}_t(\omega) \mathrel{\dot{\precsim}} \mathbf{c}'_t(\omega)$; in words, when, at almost all nodes, stream $\mathbf{c}$ has a more preferred consequence than stream $\mathbf{c}'$. This notation comes in handy to introduce (subjective) monotonicity, a basic separability assumption that requires a preference process to rank higher stochastic streams that, at almost all nodes, deliver a more preferred consequence.

Definition 3 *A preference process* $\precsim$ *is* monotone *if, for every* $\mathbf{c}, \mathbf{c}' \in \mathbf{C}$,

$$\mathbf{c} \mathrel{\dot{\precsim}} \mathbf{c}' \Longrightarrow \mathbf{c} \precsim \mathbf{c}'$$

[7] Here $\mathbf{1}_c \in \mathbf{C}$, short for $c1_{[0,T]\times\Omega}$, denotes the deterministic stream constant to c.

[8] Almost all (t, ω) stands for $\mathrm{d}t \otimes \mathbb{P}$-a.e.. Binary relations, in particular equalities and inequalities, between processes hold $\mathrm{d}t \otimes \mathbb{P}$-a.e.

To give a dynamic version of this notion, let ${}_t\mathbf{c} \in L^0_{[t,T]}$ be the *continuation* at time t of a stochastic stream $\mathbf{c} \in \mathbf{C}$. Given any two stochastic streams $\mathbf{c}$ and $\mathbf{c}'$, we write

$${}_t\mathbf{c} \succsim {}_t\mathbf{c}'$$

when, for almost all nodes $(s,\omega) \in [t,T] \times \Omega$, we have $\mathbf{c} \succsim_{s,\omega} \mathbf{c}'$. The relation ${}_t\mathbf{c} \,\dot{\succsim}\, {}_t\mathbf{c}'$ is similarly defined.

Definition 4 *A preference process $\succsim$ is* dynamically monotone *if, for every $\mathbf{c}, \mathbf{c}' \in \mathbf{C}$, at each t we have*

$${}_t\mathbf{c} \,\dot{\succsim}\, {}_t\mathbf{c}' \Longrightarrow {}_t\mathbf{c} \succsim {}_t\mathbf{c}'$$

Monotonicity is the special case when only $t = 0$ is considered. Dynamic consistency is another important property of preference processes. To introduce it, denote by $\mathbf{c}_{[0,t]} = \{\mathbf{c}_s : s \in [0,t]\} \in L^0_{[0,t]}$ the initial history of a stochastic stream $\mathbf{c}$. Given any two stochastic streams $\mathbf{c}$ and $\mathbf{c}'$, we write

$$\mathbf{c}_{[0,t]} \succsim \mathbf{c}'_{[0,t]}$$

when, for almost all nodes $(s,\omega) \in [0,t] \times \Omega$, we have $\mathbf{c} \succsim_{s,\omega} \mathbf{c}'$. The relation $\mathbf{c}_{[0,t]} \,\dot{\succsim}\, \mathbf{c}'_{[0,t]}$ is similarly defined. Finally, we write

$$\mathbf{c} \succsim_t \mathbf{c}' \tag{11}$$

when $\mathbf{c} \succsim_{t,\omega} \mathbf{c}'$ for almost all ω.

Definition 5 *A preference process $\succsim$ is* dynamically consistent *if, for every $\mathbf{c}, \mathbf{c}' \in \mathbf{C}$, at each t we have*

$$\mathbf{c}_{[0,t]} \,\dot{\sim}\, \mathbf{c}'_{[0,t]} \quad \textit{and} \quad \mathbf{c} \succsim_t \mathbf{c}' \Longrightarrow \mathbf{c}_{[0,t]} \succsim \mathbf{c}'_{[0,t]}$$

2.3 Utility

Next we turn to the numerical representation of preference processes.

Definition 6 *A map $V : \mathbf{C} \to L^0$ is a* dynamic utility operator *for a preference process $\succsim$ if:*

(i) for every $\mathbf{c}, \mathbf{c}' \in \mathbf{C}$,

$$\mathbf{c} \succsim_{t,\omega} \mathbf{c}' \Longleftrightarrow V_{t,\omega}(\mathbf{c}) \geq V_{t,\omega}(\mathbf{c}') \tag{12}$$

for almost all nodes (t,ω);

(ii) $V_T(\mathbf{c}) = 0$ for all $\mathbf{c} \in \mathbf{C}$.

By (i), a dynamic utility operator V is a decision criterion that quantifies, at almost each node, the decision makers' evaluation of stochastic streams. By (ii), stochastic streams have no value for them at the final point of time T.

Next we assume that our dynamic utility operators rank deterministic streams via standard discounting. Time is thus treated, per se, in a conventional way as we are primarily interested in issues that are peculiar to uncertainty.[9]

[9] A similar assumption can be found in the discrete-time analysis of Klibanoff et al. (2009).

A.4 There exists a unique $\beta > 0$ and a, non-constant and cardinal,[10] continuous bounded utility function $u : C \to \mathbb{R}$ such that, for all deterministic streams $\mathbf{c} \in \mathbf{C}$,

$$V_{t,\omega}(\mathbf{c}) = \int_t^T e^{-\beta(s-t)} u(\mathbf{c}_s)\, \mathrm{d}s \tag{13}$$

The utility function u is a cardinal representation for the invariant preference $\dot{\succsim}$ over consequences, i.e.,

$$c \,\dot{\succsim}\, c' \Longleftrightarrow u(c) \geq u(c')$$

for all consequences $c, c' \in C$. There might be different reasons why the utility u is cardinal and the subjective discount factor β is unique, for instance a ranking over timed risk lotteries,[11] a need to give meaning to utility differences, or even a physiological underpinning. Axiomatically, a combination of the Savage (1954) axioms, with points of time playing the role of states, and of a stationarity axiom a la Koopmans (1960) is enough to justify the discounting criterion (13).[12] Be that as it may, we do not model explicitly these standard reasons and axioms, leaving them in the background. As a result, a cardinal u and a unique β are extra primitives of our analysis, along with the preference process $\succsim$.

Lemma 1 *A dynamic utility operator V for a preference process $\succsim$ is bounded and cardinal.*[13]

The operator V thus inherits the boundedness and cardinality of u. In particular, we can write it as $V : \mathbf{C} \to L^\infty$.

Assumption A.4 implies that, for all deterministic streams $\mathbf{c}$, we have:[14]

$$\mathrm{d}V_t(\mathbf{c}) = -\left[u(c_t) - \beta V_t(\mathbf{c})\right] \mathrm{d}t \tag{14}$$

To ease the analysis, in what follows we consider a stochastic version of this property by assuming that the utility process is a semimartingale of a classical form.

A.5 The utility process $V(\mathbf{c})$ is a bounded nice Ito process

$$\mathrm{d}V_t(\mathbf{c}) = \mu_t(V(\mathbf{c}))\, \mathrm{d}t + \sigma_t(V(\mathbf{c}))\, \mathrm{d}B_t \tag{15}$$

for all streams $\mathbf{c} \in \mathbf{C}$.

As well-known, the coefficients of the utility process are uniquely pinned down via the relations:[15]

$$\mathrm{E}_t\left[\mathrm{d}V_t(\mathbf{c})\right] = \mu_t(V(\mathbf{c}))\, h + o(h) \quad ; \quad \mathrm{E}_t\left[(\mathrm{d}V_t(\mathbf{c}))^2\right] = \sigma_t^2(V(\mathbf{c}))\, h + o(h) \tag{16}$$

When $\mathbf{c}$ is deterministic, also $V(\mathbf{c})$ is deterministic and so $\sigma_t(V(\mathbf{c})) = 0$ because $V_t(\mathbf{c})$ is of bounded variation in t.[16] Thus, to ensure the consistency of (15) with (14) we assume throughout that $\mu_t(V(\mathbf{c})) = -u(c_t) + \beta V_t(\mathbf{c})$ if $\sigma_t(V(\mathbf{c})) = 0$. In particular, this is why in Definition 7, below, we assume that $\eta(\cdot, \cdot, 0) = 0$.

[10] That is, unique up to a positive affine transformation $au + b$, where $a > 0$ and $b \in \mathbb{R}$.

[11] Since A.0 just requires the consequence space to be convex, a deterministic consequence stream may consist of (uncontingent) timed risk lotteries of the consumption good.

[12] See Kopylov (2010).

[13] That is, unique up to transformations $aV + b$ with $a : [0, T] \to (0, \infty)$ and $b : [0, T] \to \mathbb{R}$. It is a deterministic notion of cardinality, a dividend of A.4.

[14] Indeed, (13) solves the inhomogeneous linear differential equation (14). See, e.g., Corollary IV.2.1 of Hartman (1982).

[15] See (66) in appendix. Throughout the paper, to ease notation we denote the conditional expectation by $\mathrm{E}_t[\cdot]$ in place of $\mathrm{E}^{\mathbb{P}}[\cdot|\mathcal{F}_t]$.

[16] By (13), if $u \geq 0$ then $V_t(\mathbf{c})$ is decreasing in t on $[0, T]$. In general, by considering $u = u^+ - u^-$ we can then express $V_t(\mathbf{c})$ as the difference of two decreasing functions on $[0, T]$.

2.4 Properties

A dynamic utility operator V is easily seen to inherit the dynamic properties of the underlying preference process $\succsim$. Specifically:

(i) $\succsim$ is dynamically monotone if and only if so is V, i.e., for every $\mathbf{c}, \mathbf{c}' \in \mathbf{C}$, at each t we have:[17]

$$u \circ {}_t\mathbf{c} \geq u \circ {}_t\mathbf{c}' \Longrightarrow {}_tV\left(\mathbf{c}\right) \geq {}_tV\left(\mathbf{c}'\right)$$

(ii) $\succsim$ is dynamically consistent if and only if so is V, i.e., for every $\mathbf{c}, \mathbf{c}' \in \mathbf{C}$, at each t we have:[18]

$$u \circ \mathbf{c}_{[0,t]} = u \circ \mathbf{c}'_{[0,t]} \quad \text{and} \quad V_t\left(\mathbf{c}\right) \geq V_t\left(\mathbf{c}'\right) \Longrightarrow V_{[0,t]}\left(\mathbf{c}\right) \geq V_{[0,t]}\left(\mathbf{c}'\right) \tag{17}$$

The dynamic monotonicity of V implies that, at each point of time t,

$$u \circ {}_t\mathbf{c} = u \circ {}_t\mathbf{c}' \Longrightarrow {}_tV\left(\mathbf{c}\right) = {}_tV\left(\mathbf{c}'\right) \tag{18}$$

Only current and future consequences thus matter, a form of consequentialism: continuation values of a monotone dynamic utility operator depend only on continuation streams.[19] With an abuse of notation, we can write ${}_tV\left(\mathbf{c}\right) = {}_tV\left({}_t\mathbf{c}\right)$. A basic instance of a dynamically monotone dynamic utility operator is the standard discounting dynamic utility operator $U : \mathbf{C} \to L^\infty$ defined, at each point of time t, by

$$U_t\left(\mathbf{c}\right) = \mathrm{E}_t\left[\int_t^T e^{-\beta(s-t)} u(\mathbf{c}_s)\,\mathrm{d}s\right] \tag{19}$$

The dynamic consistency of V implies the recursive relation

$$u \circ \mathbf{c}_{[0,t]} = u \circ \mathbf{c}'_{[0,t]} \quad \text{and} \quad V_t\left(\mathbf{c}\right) = V_t\left(\mathbf{c}'\right) \Longrightarrow V_{[0,t]}\left(\mathbf{c}\right) = V_{[0,t]}\left(\mathbf{c}'\right)$$

Since $V_T\left(\mathbf{c}\right) = V_T\left(\mathbf{c}'\right) = 0$, this relation actually implies monotonicity. A strict version of consistency is useful: the operator V is *strictly dynamically consistent* when, besides (17), at each $t > 0$ we also have:[20]

$$u \circ \mathbf{c}_{[0,t]} = u \circ \mathbf{c}'_{[0,t]} \quad \text{and} \quad V_t\left(\mathbf{c}\right) > V_t\left(\mathbf{c}'\right) \Longrightarrow V_0\left(\mathbf{c}\right) > V_0\left(\mathbf{c}'\right) \tag{20}$$

For example, from (19) we have

$$U_0\left(\mathbf{c}\right) = \mathrm{E}\left[\int_0^t e^{-\beta s} u(\mathbf{c}_s)\,\mathrm{d}s + e^{-\beta t} U_t\left(\mathbf{c}\right)\right]$$

and so the standard discounting operator is strictly dynamically consistent.

[17] The composition $u \circ {}_t\mathbf{c} \in \mathbf{L}^0_{[t,T]}$ is the process $\left(u \circ \mathbf{c}\right)_s(\omega) = u\left(\mathbf{c}_s\left(\omega\right)\right)$ for all $(s,\omega) \in [t,T] \times \Omega$. Moreover, $u \circ {}_t\mathbf{c} \geq u \circ {}_t\mathbf{c}'$ means that, for almost all $(s,\omega) \in [t,T] \times \Omega$, $u\left(\mathbf{c}_s\left(\omega\right)\right) \geq u\left(\mathbf{c}'_s\left(\omega\right)\right)$. A similar remark applies to ${}_tV\left(\mathbf{c}\right) \geq {}_tV\left(\mathbf{c}'\right)$.

[18] Here $u \circ \mathbf{c}_{[0,t]} \geq u \circ \mathbf{c}'_{[0,t]}$ means that, for almost all $(s,\omega) \in [0,t] \times \Omega$, $u\left(\mathbf{c}_s\left(\omega\right)\right) \geq u\left(\mathbf{c}'_s\left(\omega\right)\right)$. A similar remark applies to $V_{[0,t]}\left(\mathbf{c}\right) \geq V_{[0,t]}\left(\mathbf{c}'\right)$.

[19] Cf. Klibanoff et al. (2009) p. 938.

[20] Cf. Duffie and Epstein (1992) p. 373.

3 Recursive dynamic utilities

In this paper we focus on recursive dynamic utility operators, whose definition rests on the following key notion.[21]

Definition 7 *A random map* $f : [0,T] \times \Omega \times C \times \mathbb{R}^2 \to \mathbb{R}$ *is an* aggregator *if it has the form*

$$f(t, c, y, z) = u(c) - \beta y + \eta(t, y, z)$$

where $\eta : [0,T] \times \Omega \times \mathbb{R}^2 \to \mathbb{R}$ *is a variability term, with* $\eta(\cdot, \cdot, 0) = 0$.[22]

Throughout we assume a few standard regularity conditions. Note that the growth condition (21) is linear in y and quadratic in z.

A.6 (i) The function $\eta(t, \cdot, \cdot) : \mathbb{R}^2 \to \mathbb{R}$ is continuous at each t;

(ii) there exists $k > 0$ such that

$$|\eta(t, y, z) - \eta(t, 0, 0)| \leq k\left(1 + |y| + z^2\right) \tag{21}$$

for all $(t, y, z) \in [0,T] \times C \times \mathbb{R}^2$;

(iii) the process $\eta(\cdot, y, z) : [0,T] \times \Omega \to \mathbb{R}$ is predictable for all $(y, z) \in \mathbb{R}^2$.

We can now introduce recursive dynamic utility operators.

Definition 8 *A dynamic utility operator* $V : \mathbf{C} \to L^\infty$ *is* recursive *if, at each point of time* t*, it solves the forward recursion*

$$V_t(\mathbf{c}) = \mathrm{E}_t\left[\int_t^T f(s, \mathbf{c}_s, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c})))\,\mathrm{d}s\right] \qquad \forall \mathbf{c} \in \mathbf{C} \tag{22}$$

When $\eta = 0$, the aggregator f has the standard form $u(c) - \beta y$ and this recursion takes a traditional form that, as well-known, has the specification (19) as its unique solution.[23] Otherwise, the recursion (22) features the diffusion of process V, interpreted as a volatility index for this valuation process.[24] When η is symmetric in z, recursion (22) becomes

$$V_t(\mathbf{c}) = \mathrm{E}_t\left[\int_t^T f(s, \mathbf{c}_s, V_s(\mathbf{c}), \sqrt{\frac{\mathrm{d}}{\mathrm{d}s}\langle V(\mathbf{c})\rangle_s})\,\mathrm{d}s\right]$$

[21] As remarked in the Introduction, aggregators that depend on z have been introduced by Lazrak and Quenez (2003), under a set of Lipschitzian conditions that, however, are not germane to our exercise. Schroder and Skiadas (2003) have studied a consumption-portfolio problem using concave aggregators of this kind as drivers of BSDEs that uniquely define utility processes – e.g., they satisfy condition (25) below.

[22] To ease notation, the dependence of η, so of f, on ω is omitted (equalities and inequalities that involve η and f hold $\mathbb{P}$-a.e.). In reading A.6 below it is important to keep in mind this omission.

[23] Up to a recursive risk aggregator, it is also the specification of stochastic differential utility in Chen and Epstein (2002) p. 1409 (to which richer specifications can be reduced via ordinal transformations, as described in Duffie and Epstein, 1992, p. 365).

[24] See Laeven and Stadje (2014) p. 1120 for a related interpretation.

where the quadratic variation of V takes the place of its diffusion $\sigma(V(\mathbf{c}))$.[25]

A solution $V : \mathbf{C} \to L^\infty$ of recursion (22) is a recursive dynamic utility operator for the underlying preference process defined via (12). The study of these operators thus amounts to the study of the solutions of recursion (22). The next routine result further clarifies the recursive nature of these solutions.

Proposition 2 *A map $V : \mathbf{C} \to L^\infty$ solves (22) if and only if*

$$V_t(\mathbf{c}) = \mathrm{E}_t\left[\int_t^\tau e^{-\beta(s-t)}\left(u(\mathbf{c}_s) + \eta(s, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c})))\right)\mathrm{d}s + e^{-\beta(\tau-t)}V_\tau(\mathbf{c})\right] \tag{23}$$

for all $0 \le t \le \tau \le T$.

For $\tau = T$, we have $V_T(\mathbf{c}) = 0$ and so formula (23) takes the following noteworthy form where the traditional specification (19), which corresponds to $\eta = 0$, is adjusted by a variability term:

$$V_t(\mathbf{c}) = \underbrace{U_t(\mathbf{c})}_{\text{traditional term}} + \underbrace{\mathrm{E}_t\left[\int_t^T e^{-\beta(s-t)}\eta(s, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c})))\,\mathrm{d}s\right]}_{\text{variability term}} \tag{24}$$

This form also suggests that recursive dynamic utility operators are dynamically consistent. To prove this important property we focus on maximal solutions of recursion (22).[26]

Proposition 3 *A unique maximal solution $V : \mathbf{C} \to L^\infty$ of recursion (22) exists and is dynamically both monotone and consistent.*

Monotonicity is an especially noteworthy property because of the dependence of recursive dynamic utility operator on variability. Indeed, decision criteria featuring a variance term, in utils or money, are in general not monotone.

Note that $V \le U$ under the condition

$$\eta(\cdot,\cdot,z) \le \eta(\cdot,\cdot,0) = 0 \qquad \forall z \in \mathbb{R}$$

This is a form of ambiguity aversion in the sense of Ghirardato and Marinacci (2002) when we interpret U, so the case $\eta = 0$, as a decision criterion under risk.[27] For instance, this condition holds when η is quasi-concave and symmetric in z. It is actually equivalent to require $\eta \le 0$, a sign condition on variability that thus characterizes an ambiguity averse criterion V. In view of this, we will refer to a dynamic utility operator and to its corresponding aggregator as *robust* whenever they incorporate a nonzero negative variability term η.

Maximal solutions are, of course, the unique solutions when recursion (22) admits a unique solution, a notoriously elusive issue. Next we report a full-fledged uniqueness result that requires, however, a substantial Lipschitzian strengthening of the growth condition (21). Yet, as a dividend of this strengthening we also get strict dynamic consistency.

[25] See (63) in Appendix A.2. Symmetry of η means that $\eta(\cdot,\cdot,z) = \eta(\cdot,\cdot,-z)$ for all $z \in \mathbb{R}$.

[26] Proposition 3 relies on an equivalence between forward recursions like (22) and quadratic backward stochastic differential equations established in the appendix (Theorem 15).

[27] Recall that U has been introduced in (19).

Proposition 4 *A unique solution $V : \mathbf{C} \to L^\infty$ of recursion (22) exists and is strictly dynamically consistent, provided (21) is strengthened as:*

$$|\eta(t, y_1, z_1) - \eta(t, y_2, z_2)| \leq k \left(|y_1 - y_2| + |z_1 - z_2| \left(1 + |z_1| + |z_2|\right)\right) \tag{25}$$

for all $t \in [0, T]$ and all $y_1, y_2, z_1, z_2 \in \mathbb{R}$.

4 Specifications

Next we introduce a convenient class of aggregators.

Definition 9 *An aggregator f is* separable *if its variability term has the form*

$$\eta(t, y, z) = -\upsilon(y)\,\varsigma_t^2 \varphi(z) \tag{26}$$

where $\varsigma = (\varsigma_t) \in L_+^0$ is a positive process and $\upsilon, \varphi : \mathbb{R} \to \mathbb{R}$ are real-valued functions, with $\varphi(z) = 0$ if and only if $z = 0$.

Separable aggregators correspond, as argued in the next section, to dynamic utilities founded on the ambiguity aversion frameworks more used in the economics and finance literature. The quadratic aggregator (9), i.e.,

$$\eta(t, y, z) = -\frac{\upsilon(y)}{2}\varsigma_t^2 z^2 \tag{27}$$

is a main example of a separable aggregator, with $\varphi(z) = z^2/2$. It may be justified on the basis of the smooth ambiguity model. Another leading instance of a separable aggregator is the absolute value example, $\varphi(z) = |z|$, where the basis is the multiple priors approach to ambiguity aversion. We discuss foundations later, in Section 5. Here, in the present section, we present some general properties of separable aggregators that facilitate their applicability.

We assume that the separable aggregator (26) satisfies our maintained assumption A.6, as well as the following specific regularity condition.

A.7 The process ς is bounded and the functions υ and φ are bounded and continuous.

For a separable aggregator f, recursion (23) becomes

$$V_t(\mathbf{c}) = \mathrm{E}_t\left[\int_t^\tau e^{-\beta(s-t)}\left(u(\mathbf{c}_s) - \upsilon(V_s(\mathbf{c}))\,\varsigma_s^2\varphi(\sigma_s(V(\mathbf{c})))\right)\mathrm{d}s + e^{-\beta(\tau-t)}V_\tau(\mathbf{c})\right] \tag{28}$$

for each $0 \leq t \leq \tau \leq T$. In particular, recursion (24) becomes

$$V_t(\mathbf{c}) = U_t(\mathbf{c}) - \mathrm{E}_t\left[\int_t^T e^{-\beta(s-t)}\upsilon(V_s(\mathbf{c}))\,\varsigma_s^2\varphi(\sigma_s(V(\mathbf{c})))\,\mathrm{d}s\right] \tag{29}$$

We interpret the elements of a separable aggregator as follows:

I.1 The process ς indexes the degree of (perceived) model ambiguity over time: the higher ς is, the higher such ambiguity is, with $\varsigma = 0$ when it is absent (so, we are back to risk).

I.2 The function υ indexes attitudes toward model ambiguity: the (pointwise) higher υ is, the higher is the aversion to such uncertainty, with neutrality when $\upsilon = 0$.

I.3 The function φ describes the impact of the volatility of uncertain streams' valuations, with no impact when volatility is absent.

Note that interpretations I.1 and I.2 are about an information and (what is in part) a taste trait, respectively. Process ς describes the evolution of the degree of model ambiguity in the specific decision problem at hand. It enters multiplicatively and quadratically – as heuristically justified in Section 5 – the separable aggregator and makes it stochastic.

Under either (perceived) absence of model ambiguity or a neutral attitude toward it, the aggregator f takes the standard risk form $u(c) - \beta y$ and recursion (28) reduces to the standard specification (19), so $V = U$. This common reduction under "absence" and "neutrality" is an important feature of the separable specification (26) that underlies the following interpretation.

I.4 The no-variability risk aggregator $f(t, c, y) = u(c) - \beta y$ represents decision makers who either perceives no model ambiguity or have a neutral attitude toward it.

If υ is a constant coefficient (of aversion to model ambiguity), we can write recursion (29) as

$$V_t(\mathbf{c}) = U_t(\mathbf{c}) - \upsilon \, \Sigma_t(V(\mathbf{c})) \tag{30}$$

where the process $\Sigma : L^\infty \to L^0$ given by

$$\Sigma_t(V(\mathbf{c})) = \mathrm{E}_t\left[\int_t^T e^{-\beta(s-t)} \varsigma_s^2 \varphi(\sigma_s(V(\mathbf{c})))\, \mathrm{d}s\right]$$

can be viewed as a variability index for the dynamic utility operator V, which thus gets a mean-variance flavor.[28]

A dynamic utility operator V that solves recursion (28) admits a basic comparative statics exercise: by keeping all else equal, except model ambiguity attitude and perception, we have

$$\upsilon \leq \tilde{\upsilon} \quad \text{or} \quad \varsigma \leq \tilde{\varsigma} \Longrightarrow V(\mathbf{c}) \geq \tilde{V}(\mathbf{c}) \tag{31}$$

A lower ambiguity aversion or perception thus results in a higher utility of a given stream.[29]

In addition to this comparative property, a dynamic utility operator that solves recursion (28) has noteworthy uniqueness properties when υ is constant, as the next version of Proposition 4 shows.

Proposition 5 *A dynamic utility operator V that solves the forward recursion (28), with constant υ, is unique if there exists $k > 0$ such that*

$$|\varphi(z_1) - \varphi(z_2)| \leq k\,|z_1 - z_2|\,(1 + |z_1| + |z_2|)$$

for all $z_1, z_2 \in \mathbb{R}$.

[28] Index $\Sigma(V(\mathbf{c}))$ inherits from $\sigma(V(\mathbf{c}))$ the following properties: $\Sigma((V+k)(\mathbf{c})) = \Sigma(V(\mathbf{c}))$ and $\Sigma(kV(\mathbf{c})) = k^2\Sigma(V(\mathbf{c}))$ for all scalars k.

[29] This comparative remark is a simple consequence of Theorem 8, due to Lepeltier and San Martín (1998), reported in appendix.

For instance, in this case the operator V is unique for both $\varphi(z) = |z|$ and $\varphi(z) = z^2$.

Besides the quadratic (27), an absolute value example of a separable aggregator features

$$\eta(t, y, z) = -\kappa |z| \tag{32}$$

where ς is unitary (so trivial), $\upsilon = \kappa \geq 0$ and $\varphi(z) = |z|$. It corresponds to a κ-ignorance specification studied by Chen and Epstein (2002) p. 1417 as their leading example, who interpret $\kappa |z|$ as modeling ambiguity aversion within a multiple priors approach. In terms of model ambiguity, here $\kappa = 0$ amounts to both "absence" and "neutrality". Interpretation I.4 thus applies to aggregator (32) as well.[30] This aggregator is symmetric in z, so recursion (30) takes here the form

$$V_t(\mathbf{c}) = U_t(\mathbf{c}) - \kappa \mathrm{E}_t \left[\int_t^T e^{-\beta(s-t)} \sqrt{\frac{\mathrm{d}}{\mathrm{d}s} \langle V(\mathbf{c}) \rangle_s} \, \mathrm{d}s \right]$$

It reduces to the traditional specification (19) when $\kappa = 0$.

5 Heuristic foundation

In this section we provide an heuristic derivation of some principal specifications of the separable aggregator introduced in the previous section. Thereby we obtain robust dynamic utility formulations corresponding to the bases of the smooth ambiguity and α-maxmin expected utility, two ambiguity aversion frameworks which incorporate variable ambiguity attitude.[31]

5.1 Model ambiguity

Ambiguity is embodied in the uncertainty over models perceived by the decision makers: they are unsure which would be the right probabilistic model to apply to evaluate prospects and keep in consideration a set of alternative probabilistic models. The subjective (predictive) probability $\mathbb{P}$, under which (B_t) is a Brownian motion, is decision makers' best guess of the correct predictive probability $\mathbb{Q}$. However, a (countable) family of equivalent alternatives is considered, describing the decision makers' perceived model ambiguity. More precisely, they may posit a set

$$\left\{ \mathbb{Q}^\xi : \xi \in \Xi \right\}$$

of probability measures equivalent to $\mathbb{P}$, each interpreted as an alternative possibly true model.[32] As well-known, we can regard Ξ as a collection of time-varying parameters $\xi \in L^0$, with $\int_0^T \xi_s^2 \, \mathrm{d}s < \infty$, that parametrize the models $\mathbb{Q}^\xi$ via the relation

$$\mathrm{E}_t \left[\frac{\mathrm{d}\mathbb{Q}^\xi}{\mathrm{d}\mathbb{P}} \right] = \exp \left(-\frac{1}{2} \int_0^t \xi_s^2 \, \mathrm{d}s + \int_0^t \xi_s \, \mathrm{d}B_s \right) \tag{33}$$

In particular, a null ξ corresponds to $\mathbb{P}$.

[30] Indeed, the multiple priors model does not distinguish between perception and aversion to ambiguity (see Ghirardato and Marinacci, 2002).

[31] Appendix 5 contains some omitted arguments and computations for this section.

[32] This posited set of models, as an additional primitive in the study of decision problems under uncertainty, is proposed in Cerreia-Vioglio et al. (2013) and discussed in Marinacci (2015).

Under each model ξ the utility process has the Ito form

$$\mathrm{d}V_t(\mathbf{c}) = \mu_t^{\xi}(V(\mathbf{c}))\,\mathrm{d}t + \sigma_t(V(\mathbf{c}))\,\mathrm{d}B_t^{\xi} \tag{34}$$

where $\mu_t^{\xi}(V(\mathbf{c})) = \mu_t(V(\mathbf{c})) + \xi_t\sigma_t(V(\mathbf{c}))$.[33] Thus,

$$\mathrm{E}_t^{\xi}[\mathrm{d}V_t(\mathbf{c})] = [\mu(V(\mathbf{c})) + \xi\sigma(V(\mathbf{c}))]\,h + o(h) \quad ; \quad \mathrm{E}_t^{\xi}\left[(\mathrm{d}V_t(\mathbf{c}))^2\right] = \sigma_t^2(V(\mathbf{c}))\,h + o(h) \tag{35}$$

For instance, when drift and diffusion are time invariant the value process $V(\mathbf{c})$ is under $\mathbb{P}$ a Brownian motion with drift $\mu(V(\mathbf{c}))$ and diffusion $\sigma(V(\mathbf{c}))$, while it is under $\mathbb{Q}^{\xi}$ a Brownian motion with adjusted drift $\mu(V(\mathbf{c}))+\xi\sigma(V(\mathbf{c}))$ and same diffusion. Hence, model uncertainty is reduced to drift uncertainty.

5.2 The master operator equation

Next, we derive an operator equation that will be key for introducing ambiguity aversion into the value process.

A conditional *certainty utility equivalent* operator $\mathcal{M}_t : L_T^0 \to L_t^0$, with the property $\mathcal{M}_t(V_t(\mathbf{c})) = V_t(\mathbf{c})$, permits us to define the following stochastic operator equation:[34] at each t,

$$V_t(\mathbf{c}) = u(\mathbf{c}_t)\,h + \frac{1}{1+\beta h}\mathcal{M}_t(V_{t+h}(\mathbf{c})) \tag{36}$$

The future utility $V_{t+h}(\mathbf{c})$ of the stochastic stream $\mathbf{c}$ is a random variable from the perspective of time t, so needs to be converted into a conditional certainty utility equivalent $\mathcal{M}_t(V_{t+h}(\mathbf{c}))$ before being discounted via $(1+\beta h)^{-1}$. A basic example of $\mathcal{M}_t$ is the conditional expectation $\mathcal{M}_t(V_{t+h}(\mathbf{c})) = \mathrm{E}_t[V_{t+h}(\mathbf{c})]$.

We have, from (36),

$$\frac{\mathcal{M}_t(V_{t+h}(\mathbf{c})) - \mathcal{M}_t(V_t(\mathbf{c}))}{h} = -u(\mathbf{c}_t) + \beta V_t(\mathbf{c}) + o(1) \tag{37}$$

Note that $V_t(\mathbf{c})$ appears on both sides of this equality. Define a differential operator $\dot{\mathcal{M}} : L^0 \to L^0$ by

$$\dot{\mathcal{M}}_t(V(\mathbf{c})) = \lim_{h\to 0}\frac{\mathcal{M}_t(V_{t+h}(\mathbf{c})) - \mathcal{M}_t(V_t(\mathbf{c}))}{h} \tag{38}$$

if the limit exists (say, in probability). With this, (37) implies the *master* (*operator*) *equation*: at each t,

$$\dot{\mathcal{M}}_t(V(\mathbf{c})) = -u(\mathbf{c}_t) + \beta V_t(\mathbf{c}) \tag{39}$$

The unknown is $V(\mathbf{c})$. For a deterministic stream $\mathbf{c}$, the master equation reduces to (14). When $\mathcal{M}_t$ is constant additive, the difference quotient in (38) takes the simpler form

$$\dot{\mathcal{M}}_t(V(\mathbf{c})) = \lim_{h\to 0}\frac{\mathcal{M}_t(\mathrm{d}V_t(\mathbf{c}))}{h} \tag{40}$$

[33] See Theorem 1 of Girsanov (1960).

[34] For $h = 1$ we have a standard discrete time stochastic operator equation (see Marinacci and Montrucchio, 2010, for the analysis of a general version of this equation).

where $\mathrm{d}V_t(\mathbf{c}) = V_{t+h}(\mathbf{c}) - V_t(\mathbf{c})$. For instance, in the basic case $\mathcal{M}_t(V_{t+h}(\mathbf{c})) = \mathrm{E}_t[V_{t+h}(\mathbf{c})]$, we have $\dot{\mathcal{M}}_t(V(\mathbf{c})) = \mu_t(V(\mathbf{c}))$ because, in view of (16), it holds $\mathrm{E}_t[\mathrm{d}V(\mathbf{c})] = \mu_t(V(\mathbf{c}))h + o(h)$. The master equation thus becomes

$$\mu_t(V(\mathbf{c})) = -u(\mathbf{c}_t) + \beta V_t(\mathbf{c})$$

Its solution is an Ito value process that satisfies the basic BSDE

$$\mathrm{d}V_t(\mathbf{c}) = -[u(\mathbf{c}_t) - \beta V_t(\mathbf{c})]\,\mathrm{d}t + \sigma_t(V(\mathbf{c}))\,\mathrm{d}B_t$$

and so it is the standard discounting operator $U_t(\mathbf{c})$ defined by (19).

5.3 Two specifications

Decision makers concerned about model ambiguity may care about how the conditional expectation of continuation value, given a model ξ, varies as this model varies in Ξ. Bearing in mind our Ito assumption A.5 on $V(\mathbf{c})$, we assume that the model-dependent certainty utility equivalent has a basic conditional expectation form:

$$\mathrm{E}_t^\xi[V_{t+h}(\mathbf{c})] = \mathrm{E}_t^\xi[V_t(\mathbf{c}) + \mathrm{d}V_t(\mathbf{c})] = V_t(\mathbf{c}) + \mu_t^\xi(V(\mathbf{c}))h + o(h) \tag{41}$$

The decision makers do not know the value of parameter ξ and so face model ambiguity, a second layer of uncertainty represented by the map

$$\xi \mapsto \mathrm{E}_t^\xi[V_{t+h}(\mathbf{c})]$$

Then the question boils down to how the model-dependent certainty utility equivalents are aggregated across the models ξ in Ξ. Different ambiguity aversion models suggest alternative answers and thereby yield alternative constructs of $\mathcal{M}_t$. Note, the assumption that the model-dependent certainty utility equivalent $\mathrm{E}_t^\xi[V_{t+h}(\mathbf{c})]$ has a basic conditional expectation form matches a feature, normative in spirit, common to most models of ambiguity aversion: given a probability distribution, decision makers are expected utility maximizers.

5.3.1 Maxmin analysis

Assume that the decision makers aggregate the model-dependent certainty utility equivalents via an α-maxmin certainty utility equivalent operator

$$\mathcal{M}_t(V_{t+h}(\mathbf{c})) = \alpha_t \min_{\xi\in\Xi} \mathrm{E}_t^\xi[V_{t+h}(\mathbf{c})] + (1-\alpha_t)\max_{\xi\in\Xi} \mathrm{E}_t^\xi[V_{t+h}(\mathbf{c})]$$

This operator is constant additive. In particular, the $[0,1]$-valued process $\alpha = (\alpha_t) \in L^0$ captures model ambiguity attitudes a la Gilboa and Schmeidler (1989) and Ghirardato et al. (2004).

Given two processes $\underline{\xi}, \bar{\xi} \in L^\infty$ with $\underline{\xi} \le \bar{\xi}$, let $\Xi \subseteq L^\infty$ be a finite collection of processes $\underline{\xi} \le \xi \le \bar{\xi}$ that contains both $\underline{\xi}$ and $\bar{\xi}$. This collection Ξ is rectangular (cf. Chen and Epstein, 2002, p. 1411) and some simple algebra shows that:

$$\mathcal{M}_t(\mathrm{d}V_t(\mathbf{c})) = \mu_t(V(\mathbf{c}))h + \frac{1}{2}\gamma_t\sigma_t(V(\mathbf{c}))h + \frac{1}{2}\kappa_t(2\alpha_t - 1)|\sigma_t(V(\mathbf{c}))|h + o(h) \tag{42}$$

where $\gamma_t = \underline{\xi}_t + \bar{\xi}_t$ and $\kappa_t = \underline{\xi}_t - \bar{\xi}_t$. In turn, this implies

$$\dot{\mathcal{M}}_t\left(V\left(\mathbf{c}\right)\right) = \mu_t\left(V\left(\mathbf{c}\right)\right) + \frac{1}{2}\gamma_t\sigma_t\left(V\left(\mathbf{c}\right)\right) + \frac{1}{2}\kappa_t\left(2\alpha_t - 1\right)\left|\sigma_t\left(V\left(\mathbf{c}\right)\right)\right|$$

and the master equation becomes

$$\mu_t\left(V\left(\mathbf{c}\right)\right) + \frac{1}{2}\gamma_t\sigma_t\left(V\left(\mathbf{c}\right)\right) + \frac{1}{2}\kappa_t\left(V\left(\mathbf{c}\right)\right)\left(2\alpha_t - 1\right)\left|\sigma_t\left(V\left(\mathbf{c}\right)\right)\right| = -u\left(\mathbf{c}_t\right) + \beta V_t\left(\mathbf{c}\right)$$

Its solution is an Ito value process that satisfies the BSDE

$$\mathrm{d}V_t\left(\mathbf{c}\right) = -\left[u\left(\mathbf{c}_t\right) - \beta V_t\left(\mathbf{c}\right) + \frac{1}{2}\gamma_t\sigma_t\left(V\left(\mathbf{c}\right)\right) + \frac{1}{2}\kappa_t\left(V\left(\mathbf{c}\right)\right)\left(2\alpha_t - 1\right)\left|\sigma_t\right|\right]\mathrm{d}t + \sigma_t\left(V\left(\mathbf{c}\right)\right)\mathrm{d}B_t$$

and so, in view of (29), the forward recursion:[35]

$$V_t\left(\mathbf{c}\right) = U_t\left(\mathbf{c}\right) + \frac{1}{2}\mathrm{E}_t\left[\int_t^T e^{-\beta(s-t)}\gamma_t\sigma_t\left(V\left(\mathbf{c}\right)\right) + \kappa_t\left(2\alpha_t - 1\right)\left|\sigma_t\left(V\left(\mathbf{c}\right)\right)\right|\mathrm{d}s\right]$$

Summing up, the maxmin aggregator is separable and has variability term

$$\eta\left(t, y, z\right) = \frac{1}{2}\gamma_t z + \frac{1}{2}\kappa_t\left(2\alpha_t - 1\right)\left|z\right| \tag{43}$$

As $\kappa_t \leq 0$, this function is concave in z if $\alpha_t \geq 1/2$ and convex otherwise. When $\alpha_t = 1$, it has the form (60), below, with γ_t and κ_t functions of $\underline{\xi}_t$ and $\bar{\xi}_t$; in particular, it is the variability term of the aggregator in the BSDE (2.17) of Chen and Epstein (2002) and, when $\kappa = \bar{\xi}_t = -\underline{\xi}_t$ for each t, it reduces to that of the κ-ignorance aggregator (32), their leading example.[36]

5.3.2 Robust Bayesian analysis

Consider a belief process

$$\pi = \left(\pi_t\right) \subseteq \Delta\left(\Xi\right)$$

that describes the evolution of decision makers' beliefs about the unknown parameter ξ (to ease notation, we omit the dependence of π on $\mathbf{c}$). For instance, a Bayesian analysis starts with a joint probability $\Pr : 2^{\Xi} \times \mathcal{F} \to [0,1]$ given by $\Pr\left(\xi, F\right) = \pi_0\left(\xi\right)\mathbb{Q}^{\xi}\left(F\right)$ that induces, at each time t, the posterior $\pi_t\left(\xi\right) = \pi_0\left(\xi\right)\mathbb{Q}_t^{\xi} / \sum_{\xi' \in \Xi}\pi_0\left(\xi'\right)\mathbb{Q}_t^{\xi'}$ of π_0 which accounts for the decision makers' learning about the unknown parameter ξ.[37]

The belief process permits to evaluate the map $\mathcal{M}_t\left(V_{t+h}\left(\mathbf{c}\right)\right)$ via a *smooth ambiguity* certainty utility equivalent:[38]

$$\mathcal{M}_t\left(V_{t+h}\left(\mathbf{c}\right), h\right) = \phi_h^{-1}\left(\sum_{\xi}\phi_h\left(\mathrm{E}_t^{\xi}\left[V_{t+h}\left(\mathbf{c}\right)\right]\right)\pi_t\left(\xi\right)\right) \tag{44}$$

[35] Due to the equivalence established in Theorem 15 of the appendix. This is the case also for the other forward recursions of this section.

[36] Beissner et al. (2018) extend the analysis of Chen and Epstein (2002) to the α-maxmin criterion axiomatized by Ghirardato et al. (2004).

[37] Here $\mathbb{Q}_t^{\xi}$ denotes the conditional probability of $\mathbb{Q}^{\xi}$ with respect to $\mathcal{F}_t$.

[38] Cf. Klibanoff et al. (2009) p. 943. A caveat: here $\mathcal{M}_t : L^0 \times [0,\infty) \to L^0$ because the conditional certainty utility equivalent may depend on h. For simplicity, so far we neglected this possible dependence on h, but what we did so far can be extended to this case.

Here ϕ_h is a real-valued map defined on an interval of the real line that captures model ambiguity attitudes a la smooth ambiguity model of Klibanoff et al. (2005). Inspired by Hansen and Sargent (2011) and Hansen and Miao (2018), we consider specifications of ϕ_h that depend on the infinitesimal time-increment h.[39]

Cumulants The expected and the standard deviation parameter processes $\vartheta = (\vartheta_t) \in L^0$ and $\varsigma = (\varsigma_t) \in L^2$ are given by

$$\vartheta_t = \sum_{\xi} \xi_t \, \pi_t (\xi) \quad \text{and} \quad \varsigma_t = \sqrt{\sum_{\xi} \xi_t^2 \pi_t (\xi) - \vartheta_t^2}$$

In what follows we assume that $\vartheta_t = 0$ for each t, so that $\mathbb{P} = \mathbb{Q}^{\vartheta}$ and

$$\varsigma_t^2 = \sum_{\xi} \xi_t^2 \pi_t(\xi)$$

The process ς accounts for the decision makers' perception of model ambiguity, a subjective feature that depends on the belief process. In particular, we have $\varsigma_t = 0$ for each t if and only if $\pi_t (\xi : \xi_t = 0) = 1$ for each t. That is, process ς is null when decision makers do not perceive any model ambiguity in that their beliefs are concentrated on $\mathbb{P}$, so they consider their subjective beliefs to be correct. All this is in accordance with interpretation I.4.

Note that ϑ_t and ς_t^2 are the first two cumulants of ξ_t under π_t.[40] If we denote by $\chi_{i,t}$ the cumulant of order i, we thus have

$$\chi_{1,t} = \vartheta_t = 0 \quad \text{and} \quad \chi_{2,t} = \varsigma_t^2 \tag{45}$$

an observation that will be useful momentarily.

Mean-variance analysis To proceed further, we consider for the increment $\mathrm{d}V_t (\mathbf{c})$ the *robust mean-variance* criterion

$$\mathcal{M}_t (\mathrm{d}V_t (\mathbf{c}), h) = \sum_{\xi} \mathrm{E}_t^{\xi} [\mathrm{d}V_t (\mathbf{c})] \, \pi_t (\xi) - \frac{\upsilon}{2h} \mathrm{Var}_{\pi} \left(\mathrm{E}_t^{\xi} [\mathrm{d}V_t (\mathbf{c})] \right) \tag{46}$$

that arises via a quadratic approximation of (44), as in Maccheroni et al. (2013). Indeed, (35) implies that, for each ξ, we have $\mathrm{E}_t^{\xi} [\mathrm{d}V_t (\mathbf{c})] \to 0$ as $h \to 0$. Implicit here is a negative exponential specification $\phi_h (y) = -e^{-\frac{y}{\kappa h}}$, with $\upsilon = 1/\kappa = \lim_{h \to 0} \lambda^{\phi_h} h$, that makes the certainty utility equivalent (44) constant additive. With this, in view of (40) we have the master equation

$$-\frac{\upsilon}{2} \varsigma_t^2 \sigma_t^2 (V (\mathbf{c})) + \mu_t (V (\mathbf{c})) = -u (\mathbf{c}_t) + \beta V_t (\mathbf{c}) \tag{47}$$

Its solution is an Ito value process that satisfies the BSDE

$$\mathrm{d}V_t (\mathbf{c}) = -[u (\mathbf{c}_t) - \beta V_t (\mathbf{c}) - \frac{\upsilon}{2} \varsigma_t^2 \sigma_t^2 (V (\mathbf{c}))] \, \mathrm{d}t + \sigma_t (V (\mathbf{c})) \, \mathrm{d}B_t$$

[39] See Hansen and Sargent (2011), Hansen and Miao (2018), and the discussion in Section 5.3.3, below, on the justification of this h-dependence assumption. It is key in this diffusion setting, otherwise only the drift term would be $o(h)$, as remarked by Skiadas (2013).

[40] On cumulants, see for instance Billingsley (1995) pp. 142-145.

and so, in view of (29), the forward recursion:

$$V_t(\mathbf{c}) = U_t(\mathbf{c}) - \frac{\upsilon}{2}\mathrm{E}_t\left[\int_t^T e^{-\beta(s-t)}\varsigma_s^2\sigma_s^2(V(\mathbf{c}))\,\mathrm{d}s\right] \tag{48}$$

This is the form that takes here the decomposition (30).

Summing up, the robust mean-variance aggregator is separable and has the variability term (27) with constant υ, i.e.,

$$\eta(t,y,z) = -\frac{\upsilon}{2}\varsigma_t^2 z^2 \tag{49}$$

This function is concave in z and ≤ 0 provided $\upsilon \geq 0$. The robust dynamic utility V defined by the recursion (48) is thus a criterion based on a quadratic approximation of the smooth ambiguity criterion. It features a standard discounted expected utility recursion adjusted, at each point in time, by a product of a measure υ of ambiguity aversion and a measure $\varsigma_s^2\sigma_s^2$ of total variance of continuation utility taking model ambiguity into account.

Negative exponential analysis Consider the negative exponential specification of ϕ_h implicitly assumed in the mean-variance approximation. Using (35), from the certainty utility equivalent (44) we get

$$\mathcal{M}_t(\mathrm{d}V_t(\mathbf{c}), h) = \mu_t(V(\mathbf{c}))\,h - \frac{h}{\upsilon}\log\sum_\xi e^{-\upsilon\sigma_t(V(\mathbf{c}))\xi_t}\pi_t(\xi) + o(h) \tag{50}$$

In our setting, this can be seen as the route taken by Hansen and Miao (2018, 2022).[41] It leads to the master equation

$$-\frac{1}{\upsilon}\log\sum_\xi e^{-\upsilon\sigma_t(V(\mathbf{c}))\xi_t}\pi_t(\xi) + \mu_t(V(\mathbf{c})) = -u(\mathbf{c}_t) + \beta V_t(\mathbf{c})$$

Its solution is an Ito value process that satisfies the BSDE

$$\mathrm{d}V_t(\mathbf{c}) = -[u(\mathbf{c}_t) - \beta V_t(\mathbf{c}) - \frac{1}{\upsilon}\log\sum_\xi e^{-\upsilon\sigma_t(V(\mathbf{c}))\xi_t}\pi_t(\xi)]\,\mathrm{d}t + \sigma_t\mathrm{d}B_t$$

and so, in view of (24), the forward recursion is given by

$$V_t(\mathbf{c}) = U_t(\mathbf{c}) - \frac{1}{\upsilon}\mathrm{E}_t\left[\int_t^T e^{-\beta(s-t)}\log\sum_\xi e^{-\upsilon\sigma_t(V(\mathbf{c}))\xi_t}\pi_t(\xi)\,\mathrm{d}s\right]. \tag{51}$$

In sum, the negative exponential aggregator, while a special case of the general class we consider, is not separable and has the variability term

$$\eta(t,y,z) = -\frac{1}{\upsilon}\log\sum_\xi e^{-\upsilon z\xi_t}\pi_t(\xi) \tag{52}$$

This function is concave in z. It is also ≤ 0 since, by the Jensen inequality, $\eta(t,\cdot,z) \leq z\vartheta_t = 0$. The robust dynamic utility V defined by the recursion (51) is thus a criterion based on the smooth ambiguity model.

[41] The works of Hansen and Miao consider a filtering setting, allow for non-unitary intertemporal elasticities of substitution, and use a version of assumption A.5 for $\mathrm{d}V_t(\mathbf{c})/V_t(\mathbf{c})$ rather than for $\mathrm{d}V_t(\mathbf{c})$.

To see the relations between (49) and (52), note that $\log \sum_{\xi} e^{-\upsilon z \xi_t} \pi_t(\xi)$ is the cumulant generating function of ξ_t under π_t. In view of (45), we can write (52) as:

$$\eta(t, y, z) = -\frac{1}{\upsilon} \sum_{i=1}^{\infty} \frac{\chi_{i,t}}{i!} (-\upsilon z)^i = \underbrace{-\frac{\upsilon}{2} \varsigma_t^2 z^2}_{\text{mean-variance term}} - \frac{1}{\upsilon} \sum_{i=3}^{\infty} \frac{\chi_{i,t}}{i!} (-\upsilon z)^i$$

for values of υz that belong to some neighborhood of 0. For these values, the mean-variance variability term (49) is thus a truncated version of the negative exponential variability term (52). The two terms coincide in the Gaussian case where $\chi_{i,t} = 0$ for $i \geq 3$.[42]

To conclude, (52) is the variability term in the smooth ambiguity case with a negative exponential ϕ_h. It has a quadratic truncation (49) that may improve its tractability as it features a separable aggregator.

5.3.3 Discussion: ambiguity attitude and h-dependence

In the heuristic derivation of the smooth ambiguity forward recursions (48) and (51), we let ϕ_h depend on infinitesimal time intervals h. Here, we clarify the behavioral content of this parametrization. In particular, we show, contrary to first impressions, the embodied ambiguity aversion neither explodes nor implodes as we pass to the infinitesimal limit. To proceed, consider first a static smooth ambiguity model described by the parameters (π, ϕ, u), where π is a probability measure on elements of Ξ, a collection of parameters identifying probabilistic models under consideration, while ϕ and u parametrize ambiguity and risk attitude, respectively. Since u is cardinal, we can take consequences c_* and c^* such that $u(c_*) = 0 < u(c^*)$. We assume that ϕ is increasing and, to model ambiguity aversion, concave.

Let E be a collection of models in Ξ. The decision makers' prior probability that the correct model belongs to E is $\pi(E) = m$. Denote by b a bet on models that pays c^* on E and c_* off it, and by $\bar{b}$ a dual bet that does the opposite. The smooth model evaluates these bets as:

$$V(b) = m\phi(u(c^*)) + (1-m)\phi(u(c_*)) = m\phi(u(c^*)) + (1-m)\phi(0)$$

and

$$V(\bar{b}) = (1-m)\phi(u(c^*)) + m\phi(0)$$

Consider next a lottery ℓ^p which pays c^* with known probability p and c_* otherwise. The smooth model evaluates it as

$$V(\ell^p) = \phi(pu(c^*))$$

We have $V(b) \leq V(\ell^m)$ and $V(\bar{b}) \leq V(\ell^{1-m})$. Define $q, \bar{q} \in [0,1]$ to be such that

$$V(\ell^q) = V(b) \quad \text{and} \quad V(\ell^{\bar{q}}) = V(\bar{b})$$

We then have:[43]

$$m \in [q, 1-\bar{q}] \tag{53}$$

Hence, the smooth ambiguity decision maker behaves as if the prior belief that the correct model is in E matches an interval of known (i.e., lottery) probabilities, thereby revealing the decision maker's

[42] Formally, this case is not included here as Ξ is finite, but the spirit of this section is heuristic.

[43] With ϕ increasing, $\phi(qu(c^*)) = V(\ell^q) = V(b) \leq V(\ell^m) = \phi(mu(c^*))$ implies $q \leq m$. Similarly, $m \leq 1-\bar{q}$.

doubts about the accuracy of the prior belief. This is a critical property of the smooth ambiguity model, an implication of ambiguity aversion as modeled via a concave ϕ. It is the reason why we apply the moniker "robust Bayesian".

For instance, consider a negative exponential $\phi(y) = \lambda\left(1 - e^{-\frac{y}{\lambda}}\right)$, where a higher $\lambda > 0$ parametrizes a lower degree of ambiguity aversion. We have:

$$q = -\frac{\lambda}{u(c^*)}\log\left(1 - m(1 - e^{-\frac{u(c^*)}{\lambda}})\right) \quad \text{and} \quad \bar{q} = -\frac{\lambda}{u(c^*)}\log\left(1 - (1-m)(1 - e^{-\frac{u(c^*)}{\lambda}})\right) \tag{54}$$

It is easy to check that $q, \bar{q} \in [0,1]$ when $m \in [0,1]$. A lower λ, so a higher ambiguity aversion, corresponds to a larger matching interval (53), thus to the (revealed perception of a) less accurate prior belief. In particular, this interval becomes maximal, $[0,1]$, as λ goes to 0 because both q and $\bar{q}$ go to 0. In contrast, it shrinks to the singleton $\{m\}$ as λ goes to $+\infty$ because q then goes to m and $\bar{q}$ to $1-m$. Appropriately, at $m = 0$ and $m = 1$, the two cases that represent certain knowledge, the corresponding matching interval is degenerate, on 0 and 1, respectively.

The framework of the smooth ambiguity model does not include time. Its recursive version in Klibanoff et al. (2009) is in discrete time, with information and outcomes arriving at *fixed* time increments. Here we are extending the recursive smooth model to continuous time by varying, shrinking, the size of the intervals over which outcomes and information accrue. Over such short time intervals only small incremental risks are resolved (at a time): at the limit, infinitesimal time intervals allow for only infinitesimal payoff changes. To account for this peculiar feature of small time intervals, we scale payoffs changes as we go across environments with different time increments. To ease matters, assume monetary consequences and let the time-scaled consequences in our probability matching exercise be $c_* + h$ and c_*, with $h > 0$. Consider the negative exponential $\phi_h(y) = \kappa h(1 - e^{-\frac{y}{\kappa h}})$, which depends on the time increment h. In view of (54),[44] as h goes to 0 we have:

$$q_h \to -\frac{\kappa}{u'(c_*)}\log\left(1 - m(1 - e^{-\frac{u'(c_*)}{\kappa}})\right) < m \tag{55}$$

and

$$\bar{q}_h \to -\frac{\kappa}{u'(c_*)}\log\left(1 - (1-m)(1 - e^{-\frac{u'(c_*)}{\kappa}})\right) < 1 - m \tag{56}$$

The limit matching interval is thus non-trivial. So, the doubt about the accuracy of the prior belief remains as we pass to the continuous time limit: the key behavioral property ambiguity aversion in the smooth model continues to hold in the same way at the limit. This is what is ensured by the functional dependence on h in the parametrization of ambiguity attitude.

A higher $1/\kappa$ is easily seen to correspond to a larger matching interval: indeed, the quantitative role of κ in (55) and (56) is identical to that of λ in (54). This suggests that κ may be interpreted as an index of ambiguity aversion for the h-dependent specification of the negative exponential. To clarify this interpretation, recall that, as previously observed, over small time intervals decision makers are typically able to choose only among alternatives with a small payoff variability. The study of their ambiguity attitudes in this peculiar context has then to be conducted within a time-scaled, "miniature", decision environment; we should not invoke results derived within a standard static setting in which payoffs' variability is unconstrained.

To outline a such miniature environment, consider monetary consequences and define the variability of a simple (i.e., finitely many-valued) act $f : S \to \mathbb{R}$ by $\omega_f = \max_{s,s' \in S} |f(s) - f(s')|$. Given a time

[44] We assume that u is differentiable, with $u' > 0$, and consider the first-order approximation $u(c_* + h) = u'(c_*)h + o(h)$.

scale $h > 0$, the collection $\mathcal{H}$ of all simple acts f such that $\omega_f \leq h$ is then the decision makers' time-scaled decision environment. Their preferences on $\mathcal{H}$ are represented by a smooth ambiguity criterion $V_h : \mathcal{H} \to \mathbb{R}$ given by

$$V_h(f) = \phi_h^{-1}\left(\sum_{\xi} \phi_h\left(\sum_{s} u(f(s))\,\mathbb{Q}^{\xi}(s)\right)\pi(\xi)\right) \tag{57}$$

with a scale-invariant, continuous and strictly increasing u. A constant act $\mathbf{1}_c$ has no variability and so belongs to $\mathcal{H}$, with $V(\mathbf{1}_c) = u(c)$. We say that $V_{1,h}$ is *more ambiguity averse* than $V_{2,h}$ if, for all $c \in C$ and all $f \in \mathcal{H}$,

$$V_{1,h}(f) > u(c) \Longrightarrow V_{2,h}(f) > u(c) \quad \text{and} \quad V_{1,h}(f) \geq u(c) \Longrightarrow V_{2,h}(f) \geq u(c) \tag{58}$$

If the acts in $\mathcal{H}$ have certainty equivalents, this condition amounts to require that $V_{2,h} \geq V_{1,h}$ on $\mathcal{H}$.[45]

Let us turn to the negative exponential specification $\phi_h(y) = -e^{-\frac{y}{\kappa h}}$ previously used. In this case, criterion (57) becomes

$$V_h(f) = -\kappa h \log \sum_{\xi} e^{-\frac{\sum_s u(f(s))\mathbb{Q}^{\xi}(s)}{\kappa h}}\pi(\xi) \tag{59}$$

Within a time-scaled decision environment, a higher ambiguity aversion amounts to a higher coefficient $1/\kappa$, as next we show.

Proposition 6 *Criterion $V_{1,h}$ is more ambiguity averse than criterion $V_{2,h}$ if and only if*

$$\frac{1}{\kappa_1} \geq \frac{1}{\kappa_2}$$

provided $\pi_{1,h} = \pi_{2,h}$.

In time-scaled decision environments, the coefficient $1/\kappa$ may be formally interpreted as an index of (smooth) ambiguity aversion, not $1/\kappa h$. Our earlier observation that a larger $1/\kappa$ also corresponds to a larger matching interval – even as h goes to 0 – is consistent with, and enlightened by, this interpretation. Thus, we summarize our discussion as follows. Our analysis posited a time-scaled parametrization of ambiguity aversion, $\phi_h(y) = -e^{-\frac{y}{\kappa h}}$, and a sequence of time-scaled decision environments along which we considered the corresponding limit, $\lim_{h\to 0} \lambda^{\phi_h} h = (1/\kappa h)\, h = 1/\kappa$. The time-scaled parametrization ensures that the behavioral content of smooth ambiguity aversion remains consistent across h environments, and that the ambiguity averse behavior implied by the scaled ϕ_h remains stable as we pass to the continuous time limit. In other words, it is the parametrization that depends on h, not the implied ambiguity attitude. This justifies viewing $\upsilon \geq 0$ in the robust dynamic utility specification in (48) as an index of ambiguity aversion based on the smooth ambiguity model.[46]

[45] See Ghirardato and Marinacci (2002), where this comparative notion of ambiguity aversion is presented in terms of preferences. For brevity, here we state it directly in terms of the decision criteria that represent them.

[46] More generally, suppose $\phi_h(y)$ is such that $\lim_{h\to 0} \lambda^{\phi_h} h$ exists and is equal to, say, $\upsilon(y)$ (e.g., $\phi_h(y) = y^{1-\frac{\alpha}{h}}$ implies $\upsilon(y) = \alpha/y$). Then, analogous versions properties we demonstrate for the exponential case, on probability matching intervals and the stability and identification of ambiguity aversion in time-scaled environments, hold for the more general case.

6 Concluding remarks

We conclude with a few technical remarks.

1. The general version of the absolute value specification studied at the end of Section 4 is (8), i.e.,

$$\eta(t, y, z) = -\gamma_t z - \kappa_t |z| \tag{60}$$

   where $\gamma = (\gamma_t) \in L^\infty$ and $\kappa = (\kappa_t) \in L^\infty_+$ are two bounded processes. This aggregator is, however, no longer a separable aggregator (26) because of the term $\gamma_t z$. It can be seen, indeed, as a special case of the different specification

$$\eta(t, y, z) = -\gamma_t z - \kappa_t \varphi(z) \tag{61}$$

   Yet, by the Girsanov Theorem the term $\gamma_t z$ can be "absorbed" via an equivalent probability measure $\mathbb{P}^\gamma$ defined by $\mathrm{d}\mathbb{P}^\gamma/\mathrm{d}\mathbb{P} = \mathcal{E}_T(-\gamma)$ that permits to consider just $\eta(t, y, z) = -\kappa_t \varphi(z)$, which is a special case of (26) up to a change of probability measure. For this reason, we do not expatiate on specification (61). Note that it includes as a special case the rectangular specification of Chen and Epstein (2002), which is thus included in the analysis of our paper.In this regard, observe that the aggregator (61) satisfies the Lipschitz condition (25) when φ is Lipschitz, so in this important case Proposition 4 applies to it. We showed in Section 5 that a basis for (60) is the variant of α-maxmin expected utility formulation, a generalization of the maxmin criterion which allows for some flexibility of ambiguity attitude.

2. When $\upsilon\varsigma_t^2$ is non-random and time invariant, so information independent, recursion (48) reduces to that studied by Duffie and Epstein (1992) and equation (47) is then the counterpart here of their eq. (17) p. 362. As is well-known, in their case a suitable ordinal transformation permits one to eliminate the variability term and to go back to the standard recursion $V_t(\mathbf{c}) = \mathrm{E}_t[\int_t^T u(\mathbf{c}_s) - \beta V_s(\mathbf{c})\, \mathrm{d}s]$. This is no longer true in our case when $\upsilon\varsigma_t^2$ is random and time dependent (all the more when υ is non-constant).

A Background material

A.1 Basic notions

Filtered probability space Throughout we fix a time interval $[0, T]$, with generic element t, and a filtered probability space $(\Omega, \mathbb{F}, \mathcal{F}, \mathbb{P})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, and $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is a filtration that satisfies the usual conditions, i.e., it is complete (i.e., $\mathcal{F}_0$ contains all P-null sets) and right-continuous. Throughout, "adapted" is always understood to be with respect to this filtration and we denote the conditional expectation by $\mathrm{E}_t[\cdot]$ in place of $\mathrm{E}^{\mathbb{P}}[\cdot|\mathcal{F}_t]$.

To ease matters, throughout we assume that $\mathbb{F}$ is the natural Brownian filtration. Yet, we expect that most of our analysis would continue to hold if $\mathbb{F}$ were only required to have the predictable representation property (indeed, this is what assumed in Lazrak and Quenez, 2003).

Spaces We denote by $\mathcal{P}$ the predictable sigma algebra on $[0,T]\times\Omega$ generated by adapted and continuous processes. We denote by $L_t^0 = L^0(\Omega, \mathcal{F}_t, \mathbb{P})$, for each $t\in[0,T]$, the collection of $\mathcal{F}_t$-measurable real-valued functions, while $L^0 = L^0(\Omega\times[0,T], \mathcal{P}, \mathbb{F}, \mathrm{d}t\otimes\mathbb{P})$ denotes the collection of the real-valued processes that are predictable (so, adapted).

On the spaces L_t^0 and L^0 we consider the usual norms $\|X_t\|^2 = \mathrm{E}[X_t^2]$ and $\|X\|^2 = \mathrm{E}[\int_0^T X_t^2 \mathrm{d}t]$, respectively. We denote by L_t^2 and L^2 the spaces of functions and processes that have such norms finite. The space L^2 is an Hilbert space when endowed with the inner product $\langle X, Y\rangle = \mathrm{E}[\int_0^T X_t Y_t \mathrm{d}t]$.

The spaces L_t^∞ and L^∞ are similarly defined with respect to the supnorms $\|X_t\|_\infty = \mathrm{essup}_{\omega\in\Omega}|X_t(\omega)|$ and $\|X\|_\infty = \mathrm{essup}_{(\omega,t)\in\Omega\times[0,T]}|X_t(\omega)|$. Processes that belong to L^∞ are called *bounded*. Clearly,

$$L_t^\infty \subseteq L_t^2 \subseteq L_t^0 \quad \text{and} \quad L^\infty \subseteq L^2 \subseteq L^0$$

Given any $\tau\in[0,T)$, denote $\mathcal{P}_{[\tau,T]} = \mathcal{P}\cap([\tau,T]\times\Omega)$ and

$$L^2_{[\tau,T]} = \left\{X : [\tau,T]\times\Omega\to\mathbb{R} : X \text{ is } \mathcal{P}_{[\tau,T]}\text{-measurable and } \mathrm{E}\left[\int_\tau^T X_s^2\,\mathrm{d}s\right] < \infty\right\}$$

In particular, $L^2_{[0,T]} = L^2$.

The following table summarizes the spaces of real-valued processes that we will consider:

L^0	Predictable processes
L^2	Predictable processes X with $\mathrm{E}[\int_0^T X_t^2\mathrm{d}t] < \infty$
L^∞	Bounded predictable processes

Besides real-valued processes, in the paper we also consider processes $X = (X_t)$ taking values in a topological space S endowed with its Borel sigma algebra $\mathcal{B}_S$. In this case, each X_t is a measurable map with $X_t^{-1}(B)\in\mathcal{F}_t$ for all $B\in\mathcal{B}_S$.

Unless otherwise stated, equalities and inequalities between random variables defined on $(\Omega,\mathcal{F})$ hold $\mathbb{P}$-a.e., while those between processes defined on $(\Omega\times[0,T],\mathcal{P})$ hold $\mathrm{d}t\otimes\mathbb{P}$-a.e.. Note that t may appear sometimes as an argument and sometimes as an index of a function, say $f_t(\cdot)$ and $f(t,\cdot)$. Context should clarify, so this inconsistent notation should cause no confusion.

We close with some terminology on real-valued functions of a single variable. In particular, a such function $f : \mathrm{dom}\, f\subseteq\mathbb{R}^n\to\mathbb{R}$ is *Lipschitz* if there exists $k>0$ such that $|f(x)-f(y)|\le k\|x-y\|$ for all $x,y\in\mathrm{dom}\, f$, while it has *linear growth* if there exists $k>0$ such that $|f(x)|\le k(1+\|x\|)$ for all $x\in\mathrm{dom}\, f$.

A.2 Ito processes

A stochastic process $X\in L^0$ is *continuous* if the map $t\longmapsto X_t$ is $\mathbb{P}$-a.e. continuous.

Definition 10 *Given a $\tau\in[0,T]$, a continuous process $X=(X_t)_{t\in[\tau,T]}$ is a* Ito process *on $[\tau,T]$ if*

$$X_t = X_\tau + \int_\tau^t a_s\,\mathrm{d}s + \int_\tau^t b_s\,\mathrm{d}B_s \qquad \forall t\in[\tau,T] \tag{62}$$

where X_τ is $\mathcal{F}_\tau$-measurable, and $a,b\in L^0$ are such that $\int_0^T |a_s|\,\mathrm{d}s<\infty$ and $\int_0^T b_s^2\mathrm{d}s<\infty$.

The decomposition (62) is unique.[47] The *stochastic differential* of a Ito process X is, heuristically, written as

$$\mathrm{d}X_t = a_t\,\mathrm{d}t + b_t\,\mathrm{d}B_t \qquad \forall t \in [0,T]$$

The quadratic variation of a Ito process X on $[0,T]$ is given by $\langle X\rangle_t = \int_0^t b_s^2\,\mathrm{d}s$ for all $t \in [0,T]$. The square of b_t is thus the integrand of the quadratic variation of the process X, with

$$\sqrt{\frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle_t} = |b_t| \tag{63}$$

In this sense, the process b can be interpreted in terms of "variability" of the process. So, b is called *volatility* of X and is denoted by $\sigma(X)$. In this notation, (63) becomes

$$\sqrt{\frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle_t} = |\sigma_t(X)| \tag{64}$$

An Ito process is *nice* if it is bounded and $b \in L^2$. In particular, $\mathrm{E}\left[\int_0^T b_s^2\,\mathrm{d}s\right] < \infty$ and so $\langle X\rangle$ is integrable.

Let $X = (X_t)_{t\in[0,T]}$ be an Ito process on $[0,T]$. Its restriction $(X_t)_{t\in[\tau,T]}$ on each time interval $[\tau,T]$ is, for each $\tau \in [0,T]$, also an Ito process. For each $0 \le \tau < t \le T$, we thus have

$$\mathrm{E}_\tau[X_t] = \mathrm{E}_\tau\left[X_\tau + \int_\tau^t a_s\,\mathrm{d}s + \int_\tau^t b_s\,\mathrm{d}B_s\right] = X_\tau + \int_\tau^t a_s\,\mathrm{d}s + \mathrm{E}_\tau\left[\left[\int_\tau^t b_s\,\mathrm{d}B_s\right]\right] = X_\tau + \int_\tau^t a_s\,\mathrm{d}s \tag{65}$$

because $\mathrm{E}_\tau\left[\int_\tau^t b_s\,\mathrm{d}B_s\right] = 0$ (see, e.g., eq. 5.9 p. 106 in Le Gall, 2016). Define $\varphi(t) = \int_\tau^t a_s\mathrm{d}s$ on a small enough interval $[\tau, \tau+\varepsilon)$. If a is continuous in s, we have $\varphi'(s) = a_s$ for each $s \in [\tau, t]$, and so $\varphi(t) = \varphi(\tau) + \varphi'(\tau)(t-\tau) + o((t-\tau))$, i.e., the following approximation of the last equation

$$\mathrm{E}_\tau[X_t] = X_\tau + a_\tau(t-\tau) + o((t-\tau)) \tag{66}$$

A.3 SDE

Given $\mathcal{P}\otimes\mathcal{B}(\mathbb{R})$-measurable coefficients $\alpha, \beta : [0,T]\times\Omega\times\mathbb{R} \to \mathbb{R}$ and an initial time $\tau \in [0,T]$, we call

$$X_t = \xi_\tau + \int_\tau^t \alpha(s, X_s)\,\mathrm{d}s + \int_\tau^t \beta(s, X_s)\,\mathrm{d}B_s \qquad \forall t \in [\tau, T] \tag{67}$$

a *stochastic differential equation* (SDE) with initial (random) condition $\xi_\tau \in L^0_\tau$. The (Markovian) coefficients α and β are called *drift* and *diffusion*. Heuristically, we write

$$\mathrm{d}X_t = \alpha(t, X_t)\,\mathrm{d}t + \beta(t, X_t)\,\mathrm{d}B_t \qquad X_\tau = \xi_\tau$$

Note that the coefficients α and β are random. When $\tau = 0$ the initial condition ξ_0 is a scalar.

Definition 11 *A (strong)* solution *of the SDE (67) is a continuous and adapted process* $(X_t)_{t\in[\tau,T]}$ *that satisfies it.*

[47] Of course, $\mathrm{d}t\otimes\mathbb{P}$-a.e. (see, e.g., Gihman and Skorohod, 1979, Theorem III.2). Note that the processes a and b are defined on $[0,T]$ (so, if defined on $[\tau,T]$, they are required to be extendable on $[0,T]$ as predictable processes).

A solution X is unique if, given any other solution X', we have $\mathbb{P}\left(X_t = X'_t, \tau \leq t \leq T\right) = 1$. Unique existence is ensured by the *Ito conditions*: for all $(t, \omega) \in [0, T] \times \Omega$, the functions $\alpha\left(t, \omega, \cdot\right), \beta\left(t, \omega, \cdot\right) : \mathbb{R} \to \mathbb{R}$ are Lipschitz and have linear growth (see, e.g., Fleming and Rishel, 1975, p. 120, and Nisio, 2015, p. 16).

Given $\tau, t \in [0, T]$, we denote by $\mathrm{SDE}([\tau, t], \alpha, \beta, x_\tau)$ a stochastic differential equation on a time interval $[\tau, t]$ with initial condition $x_\tau \in L^0_\tau$ and drift and diffusion coefficients $\alpha, \beta : [0, T] \times \Omega \times \mathbb{R} \to \mathbb{R}$.

A.4 BSDE

Given a $\tau \in [0, T]$, we call

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s)\, \mathrm{d}s - \int_t^T Z_s\, \mathrm{d}B_s \qquad \forall t \in [\tau, T] \tag{68}$$

a *backward stochastic differential equation* (BSDE) on $[\tau, T]$ with driver $g : [0, T] \times \Omega \times \mathbb{R}^2 \to \mathbb{R}$ and final condition $\xi \in L^0_T$. We often use the notation $\mathrm{BSDE}\left(g, \xi, [\tau, T]\right)$ to refer to this backward stochastic differential equation.

Throughout we make the following assumptions.

(B.i) $g : [0, T] \times \Omega \times \mathbb{R}^2 \to \mathbb{R}$ is a $\mathcal{P} \otimes \mathcal{B}\left(\mathbb{R}^2\right)$-measurable function.

(B.ii) The final condition is bounded, i.e., $\xi \in L^\infty_T$.

Definition 12 *A (bounded)* solution *of the BSDE (68) is a pair $(Y, Z) \in L^\infty \times L^2$, with Y continuous, that satisfies it.*

A solution is *maximal* if, given any other solution (Y', Z') of (68), we have $Y' \leq Y$. Clearly, a maximal solution is unique.

Since $Y_\tau = \xi + \int_\tau^T g(s, Y_s, Z_s)\mathrm{d}s - \int_\tau^T Z_s \mathrm{d}B_s$, for each $t \in [\tau, T]$ we have:

$$\begin{aligned} Y_t &= Y_\tau - \int_\tau^T g(s, Y_s, Z_s)\, \mathrm{d}s + \int_\tau^T Z_s\, \mathrm{d}B_s + \int_t^T g(s, Y_s, Z_s)\, \mathrm{d}s - \int_t^T Z_s\, \mathrm{d}B_s \\ &= Y_\tau - \int_\tau^t g(s, Y_s, Z_s)\, \mathrm{d}s + \int_\tau^t Z_s\, \mathrm{d}B_s \end{aligned} \tag{69}$$

Thus, often one writes a BSDE as

$$\mathrm{d}Y_t = -g(t, Y_t, Z_t)\, \mathrm{d}t + Z_t\, \mathrm{d}B_t \qquad Y_T = \xi$$

A BSDE is *Lipschitz* if it satisfies the following standard condition.

(L) There exists a constant $c > 0$ such that

$$\left|g\left(t, y_1, z_1\right) - g\left(t, y_2, z_2\right)\right| \leq c\left(|y_1 - y_2| + |z_1 - z_2|\right)$$

for all $t \in [0, T]$ and all $y_1, y_2, z_1, z_2 \in \mathbb{R}$.

Unique existence for a Lipschitz BSDE has been proved by Pardoux and Peng (1990).

A BSDE (68) is *quadratic* provided its driver g is *quadratic*, that is, (Q.i) for each $(t,\omega) \in [0,T] \times \Omega$, the function $g(t,\omega,\cdot,\cdot)$ is continuous;

(Q.ii) there exist $A, B, C \geq 0$ such that, for each $(t,y,z) \in [0,T] \times \mathbb{R}^2$, we have

$$|g(t,y,z)| \leq A + B|y| + Cz^2$$

Existence and uniqueness of a solution for quadratic BSDE has been studied by Lepeltier and San Martin (1998), Kobylanski (2000), Briand and Hu (2006, 2008) and Tevzadze (2008). For our purposes, two results of Lepeltier and San Martin (1998) are important. We begin with a special case of their Theorem 1.

Theorem 7 (Lepeltier-San Martin) *A quadratic BSDE on $[0,T]$ has a maximal solution.*

Next we report a comparison result, a special case of Corollary 2 of Lepeltier and San Martin (1998).

Theorem 8 (Lepeltier-San Martin) *Consider two quadratic BSDEs on $[0,T]$ that feature quadratic drivers g and $\tilde{g}$ and final conditions ξ and $\tilde{\xi}$, respectively. If Y and $\tilde{Y}$ are maximal solutions, then*

$$\xi \geq \tilde{\xi} \quad \textit{and} \quad g \geq \tilde{g} \Longrightarrow Y \geq \tilde{Y}$$

Next we consider a unique existence result, due to Kobylanski (2000). See also Touzi (2013) p. 168 and Briand and Elie (2013).

Theorem 9 (Kobylanski) *A quadratic BSDE has a unique solution if the following condition holds:*

(K) there exists a constant $c > 0$ such that

$$|g(t,y_1,z_1) - g(t,y_2,z_2)| \leq c|y_1 - y_2| + c|z_1 - z_2|(1 + |z_1| + |z_2|)$$

for all $y_1, y_2, z_1, z_2 \in \mathbb{R}$.

Under condition (K) it is possible to establish a sharp relation, due to Tevzadze (2008), between solutions and final conditions.

Theorem 10 (Tevzadze) *Consider two quadratic BSDEs on $[0,T]$ that feature a common quadratic driver g satisfying condition (K) and final conditions ξ and $\tilde{\xi}$, respectively. If Y and $\tilde{Y}$ are their unique solutions, then there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ and a bounded process $\alpha \in L^0$ such that*

$$Y_t - \tilde{Y}_t = \mathrm{E}_t^{\mathbb{Q}}\left[\mathrm{e}^{\int_t^T \alpha_s \mathrm{d}s}(\xi - \tilde{\xi})\right] \qquad \forall t \in [0,T] \tag{70}$$

Proof In the notation of Tevzadze (2008), in our setup we have $d = n = 1$, $\alpha_t = \gamma_t = g_t = 0$ for each $t \in [0,T]$, $M = B$ (and so $\sigma = 1$) and $K_t = t$. The hypotheses of Theorem 2 of Tevzadze (2008) are satisfied (cf. the remark on p. 514 of Tevzadze, 2008) by quadratic BSDEs featuring a bounded final condition and a quadratic driver satisfying condition (K). Equality (70) appears at the end of the proof of his Theorem 2, with $\alpha_t = (g(t,Y_t,Z_t) - g(t,\tilde{Y}_t,Z_t))/(Y_t - \tilde{Y}_t)$. By condition (K), for each $t \in [0,T]$ we have $|\alpha_t| \leq c$, so process α is bounded. ■

A.5 Conditioning and all that

Let $X = (X_t) \in L^0_+$ be a positive process with $\mathrm{E}[X_T] = 1$ and $X_T > 0$. Define a probability measure $\mathbb{P}^X$ on $(\Omega, \mathcal{F})$ by $\mathrm{d}\mathbb{P}^X/\mathrm{d}\mathbb{P} = X_T$. For each t we have[48]

$$\frac{\mathrm{d}\mathbb{P}^X_t}{\mathrm{d}\mathbb{P}_t} = \mathrm{E}_t[X_T]$$

Since $X_T > 0$, the probabilities $\mathbb{P}^X$ and $\mathbb{P}$ are equivalent. The map $\mathbb{P} \longmapsto \mathbb{P}^X$ is called a *change of measure.*

Let $Z \in L^0$ with $\mathrm{E}^X[|Z_t|] < \infty$ for each t. By the *conditional Bayes' rule* (see, e.g., Nisio, 2015, p. 18), for $0 \le t \le s \le T$ we have

$$\mathrm{E}^X_t[Z_s] = \frac{\mathrm{E}_t[Z_s \mathrm{E}_s[X_T]]}{\mathrm{E}_t[X_T]}$$

where we write $\mathrm{E}^X_t[\cdot]$ in place of $\mathrm{E}^{\mathbb{P}^X}[\cdot|\mathcal{F}_t]$. So, process (Z_t) is a martingale with respect to $\mathbb{P}^X$ if and only if process $(Z_t \mathrm{E}_t[X_T])$ is a martingale with respect to $\mathbb{P}$.

Assume that, in addition, the process X is a martingale (with respect to $\mathbb{P}$). For each t we then have

$$\mathrm{E}_t\left[\frac{\mathrm{d}\mathbb{P}^X}{\mathrm{d}\mathbb{P}}\right] = X_t \tag{71}$$

and the conditional Bayes' rule takes the form (cf. Karatzas and Shreve, 1991, Lemma 5.3):

$$\mathrm{E}^X_t[Z_s] = \frac{1}{X_t}\mathrm{E}_t[Z_s X_s] \tag{72}$$

We have the following conditional version of the Fubini Theorem (cf. Schilling, 2017, p. 354).

Proposition 11 *Let $\mathcal{G}$ be a sigma algebra included in $\mathcal{F}$. Given any $0 \le t \le \tau \le T$, if $\mathrm{E}[\int_t^\tau |X_s|\,\mathrm{d}s] = \int_t^\tau \mathrm{E}[|X_s|]\,\mathrm{d}s < \infty$ then:*

$$\mathrm{E}\left[\int_t^\tau X_t\,\mathrm{d}s \mid \mathcal{G}\right] = \int_t^\tau \mathrm{E}[X_t|\,\mathcal{G}]\,\mathrm{d}s$$

In particular, if $\mathrm{E}[\int_0^T |X_s|]\,\mathrm{d}s = \int_0^T \mathrm{E}[|X_s|]\,\mathrm{d}s < \infty$, then $\mathrm{E}_t\left[\int_t^T X_s\,\mathrm{d}s\right] = \int_t^T \mathrm{E}_t[X_s]\,\mathrm{d}s$ for all $0 \le t \le T$. In turn, along with (72) this implies:

$$\mathrm{E}^X_t\left[\int_t^T Z_s \mathrm{d}s\right] = \int_t^T \mathrm{E}^X_t[Z_s]\,\mathrm{d}s = \int_t^T \frac{1}{X_t}\mathrm{E}_t[X_s Z_s]\,\mathrm{d}s = \frac{1}{X_t}\mathrm{E}_t\left[\int_t^T X_s Z_s \mathrm{d}s\right] \tag{73}$$

Example 1 By (73), the standard intertemporal utility function

$$V^X_t(\mathbf{c}) = \mathrm{E}^X_t\left[\int_t^T e^{-\beta(s-t)} u(c_s)\,\mathrm{d}s\right] \qquad \forall \mathbf{c} \in L^0_+$$

with respect to the probability $\mathbb{P}^X$ can be written in terms of $\mathbb{P}$ as

$$V^X_t(\mathbf{c}) = \frac{1}{X_t}\mathrm{E}_t\left[\int_t^T X_s e^{-\beta(s-t)} u(c_s)\mathrm{d}s\right] \qquad \forall \mathbf{c} \in L^0_+ \tag{74}$$

▲

[48] Here $\mathbb{P}^Y_t$ and $\mathbb{P}_t$ denote, respectively, the conditional probabilities of $\mathbb{P}^Y$ and $\mathbb{P}$ with respect to $\mathcal{F}_t$.

Given a process $X \in L^0$ with $\int_0^T X_s^2\,\mathrm{d}s < \infty$, define a continuous positive supermartingale $\mathcal{E}(X) \in L_+^0$ by

$$\mathcal{E}_t(X) = \exp\left(-\frac{1}{2}\int_0^t X_s^2\,\mathrm{d}s + \int_0^t X_s \mathrm{d}B_s\right) \qquad \forall t \in [0,T]$$

This supermartingale is a martingale, called *exponential martingale*, if and only if $\mathrm{E}[\mathcal{E}_T(X)] = 1$.[49] By Ito's Lemma, we have $\mathcal{E}_t(X) = 1 + \int_0^t \mathcal{E}_s(X) X_s \mathrm{d}B_s$ (cf. McKean, 1969, p. 33).

The *Novikov condition* $\mathrm{E}[e^{\frac{1}{2}\int_0^T X_s^2\,\mathrm{d}s}] < \infty$ ensures that $\mathcal{E}(X)$ is a martingale with $\mathrm{E}[\mathcal{E}_t(X)] = 1$ for each $t \in [0,T]$.

When process $\mathcal{E}(X)$ is a martingale, the change of measure $\mathbb{P} \longmapsto \mathbb{P}^{\mathcal{E}(X)}$ is called a *Girsanov transformation.* By (71), we have

$$\mathrm{E}_t\left[\frac{\mathrm{d}\mathbb{P}^{\mathcal{E}(X)}}{\mathrm{d}\mathbb{P}}\right] = \exp\left(-\frac{1}{2}\int_0^t X_s^2\,\mathrm{d}s + \int_0^t X_s \mathrm{d}B_s\right) \tag{75}$$

A version of the Girsanov Theorem ensures that, for $X \in L^0$ with $\int_0^T X_s^2\,\mathrm{d}s < \infty$, if $\mathrm{E}[\mathcal{E}_T(X)] = 1$ then the process B^X defined by

$$B_t^X = B_t - \int_0^t X_s \mathrm{d}s \qquad \forall t \in [0,T] \tag{76}$$

is a Brownian motion under $\mathbb{P}^{\mathcal{E}(X)}$. In differential form,

$$\mathrm{d}B_t^X = \mathrm{d}B_t - X_t \mathrm{d}t \tag{77}$$

Example 2 In terms of BSDEs, since $\int_t^T Z_s \mathrm{d}B_s^X = -\int_t^T X_s Z_s \mathrm{d}s + \int_t^T Z_s \mathrm{d}B_s$, we have

$$\begin{aligned} Y_t &= \xi + \int_t^T g(s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}B_s \\ &= \xi + \int_t^T [g(s, Y_s, Z_s) - X_s Z_s]\,\mathrm{d}s - \int_t^T Z_s \mathrm{d}B_s \end{aligned}$$

A pair $(Y, Z) \in L_Y^2 \times L_Y^2$ that solves this BSDE satisfies the forward recursion

$$Y_t = \mathrm{E}_t^{\mathcal{E}(X)}\left[\xi + \int_t^T g(s, Y_s, Z_s)\mathrm{d}s\right] = \mathrm{E}_t\left[\xi + \int_t^T [g(s, Y_s, Z_s) - X_s Z_s]\,\mathrm{d}s\right]$$

▲

B Proofs and related analysis

B.1 Preamble on quadratic BSDEs

Proposition 12 *Let g be a quadratic driver. The following conditions are equivalent for a pair $(Y, Z) \in L_{[\tau,T]}^\infty(\mathbb{R}) \times L_{[\tau,T]}^2(\mathbb{R})$:*

(i) (Y, Z) is a solution of $\mathrm{BSDE}(g, \xi, [\tau, T])$*;*

[49] See, e.g., Section 6.2 of Liptser and Shiryaev (2001).

(ii) Y is a nice Ito process such that

$$Y_t = \xi + \int_t^T g(s, Y_s, \sigma_s(Y))\mathrm{d}s - \int_t^T \sigma_s(Y)\,\mathrm{d}B_s \qquad \forall t \in [\tau, T] \tag{78}$$

and $Z = \sigma(Y)$.

Proof (i) implies (ii). If (Y, Z) solves the BSDE$\,(g, \xi, [\tau, T])$, then $Y, Z \in L^2_{[\tau,T]}(\mathbb{R})$, $(Y_t)_{t\in[\tau,T]}$ is continuous and $\|Y\|_\infty < \infty$. Moreover, by (69) we have $Y_t = Y_\tau - \int_\tau^t g(s, Y_s, Z_s)\,\mathrm{d}s + \int_\tau^t Z_s \mathrm{d}B_s$ for all $t \in [\tau, T]$. Since g is a quadratic driver,

$$\begin{aligned}\int_\tau^T |g(s, Y_s, Z_s)|\,\mathrm{d}s \;&\leq\; \int_\tau^T A\mathrm{d}s + B\int_\tau^T |Y_s|\,\mathrm{d}s + C\int_\tau^T |Z_s|^2\,\mathrm{d}s \\ &\leq\; (A + B\|Y\|_\infty)(T-\tau) + C\int_\tau^T |Z_s|^2\,\mathrm{d}s\end{aligned}$$

and $\int_\tau^T |Z_s|^2\,\mathrm{d}s < \infty$ because $Z \in L^2_{[\tau,T]}(\mathbb{R})$. Thus, Y is an Ito process, with $\sigma(Y) = Z$, that satisfies (78).

(ii) implies (i). If Y is a nice Ito process that satisfies (78) and $Z = \sigma(Y)$, the pair (Y, Z) solves (68). Moreover, since Y is nice, we have $Y, Z \in L^2_{[\tau,T]}(\mathbb{R})$ and $\|Y\|_\infty < \infty$, as desired. ■

Since Z is uniquely determined by Y, we can say that a nice Ito process $(Y_t)_{t\in[\tau,T]}$ is a solution of a quadratic BSDE$\,(g, \xi, [\tau, T])$ if it is such that

$$Y_t = \xi + \int_t^T g(s, Y_s, \sigma_s(Y))\mathrm{d}s - \int_t^T \sigma_s(Y)\,\mathrm{d}B_s \qquad \forall t \in [\tau, T] \tag{79}$$

With an abuse of notation, we then write $Y \in \mathrm{BSDE}\,(g, \xi, [\tau, T])$. In particular, Y is maximal if $Y \geq Y'$ for all $Y' \in \mathrm{BSDE}\,(g, \xi, [\tau, T])$.

The next result is a simple extension of Theorem 8 of Lepeltier and San Martin (1998) which is important for our purposes.

Proposition 13 *If, for all $(t, \omega, y, z) \in [\tau, T] \times \Omega \times \mathbb{R}^2$,*

$$g(t, \omega, y, z) \geq h(t, \omega, y, z) \qquad \text{and} \qquad \xi(\omega) \geq \zeta(\omega)$$

then

$$(Y_t^{g,\xi})_{[\tau,T]} \geq (Y_t^{h,\zeta})_{[\tau,T]}$$

where $Y^{g,\xi}$ and $Y^{h,\zeta}$ are the maximal solutions of quadratic BSDE$\,(g, \xi, [0, T])$ *and quadratic* BSDE$\,(h, \zeta, [0, T])$*, respectively.*[50]

The idea is that the restriction $(Y_t^{g,\xi})_{t\in[\tau,T]}$ to $[\tau, T]$ of the maximal solution $(Y_t^{g,\xi})_{t\in[0,T]}$ of the quadratic BSDE$\,(g, \xi, [0, T])$ only depends on the driver's restriction $g_{|[\tau,T]\times\Omega\times\mathbb{R}^2}$. We prove it by showing that such a restriction is the maximal solution of the quadratic BSDE$\,(g, \xi, [\tau, T])$ – which, if exists, is unique and clearly depends only on $g_{|[\tau,T]\times\Omega\times\mathbb{R}^2}$.

[50] In terms of notation, the context should clarify when the letter h denotes an increment or a driver.

Lemma 14 *If* $(Y_t^{g,\xi})_{t\in[0,T]}$ *is the maximal solution of quadratic* BSDE$(g,\xi,[0,T])$*, then:*

(i) $(Y_t^{g,\xi})_{t\in[\tau,T]}$ *is the maximal solution of the quadratic* BSDE$(g,\xi,[\tau,T])$*;*

(ii) $(Y_t^{g,\xi})_{t\in[0,\tau]}$ *is the maximal solution of the quadratic* BSDE$(g,Y_\tau^{g,\xi},[0,\tau])$*.*

To prove (ii) note that a process $Y\in L^0$ solves, for each $t\in[0,T]$, the recursion

$$Y_t = \mathrm{E}_t\left[\xi + \int_t^T g(s,Y_s,Z_s)\,\mathrm{d}s\right]$$

if and only if, for each $0\le t\le \tau\le T$,

$$Y_t = \mathrm{E}_t\left[\int_t^\tau g(s,Y_s,Z_s)\,\mathrm{d}s + Y_\tau\right] \tag{80}$$

Indeed,

$$\begin{aligned}
Y_t &= \mathrm{E}_t\left[\xi+\int_t^T g(s,Y_s,Z_s)\,\mathrm{d}s\right] = \mathrm{E}_t\left[\mathrm{E}_\tau\left[\xi+\int_t^T g(s,Y_s,Z_s)\,\mathrm{d}s\right]\right]\\
&= \mathrm{E}_t\left[\int_t^\tau g(s,Y_s,Z_s)\,\mathrm{d}s\right] + \mathrm{E}_\tau\left[\xi+\int_\tau^T g(s,Y_s,Z_s)\,\mathrm{d}s\right] = \mathrm{E}_t\left[\int_t^\tau g(s,Y_s,Z_s)\,\mathrm{d}s + Y_\tau\right]
\end{aligned}$$

Proof (i) By Proposition 12, $(Y_t^{g,\xi})_{t\in[\tau,T]}$ is a nice Ito process $(Y_t)_{t\in[\tau,T]}$ such that

$$Y_t = \xi + \int_t^T g(s,Y_s,\sigma_s(Y))\mathrm{d}s - \int_t^T \sigma_s(Y)\,\mathrm{d}B_s \qquad \forall t\in[\tau,T]$$

that is, $(Y_t^{g,\xi})_{t\in[\tau,T]}\in$ BSDE$(g,\xi,[\tau,T])$. Now let $(V_t)_{t\in[\tau,T]}\in$ BSDE$(g,\xi,[\tau,T])$. Consider the maximal solution $(Z_t)_{t\in[0,\tau]}$ of BSDE$(g,V_\tau,[0,\tau])$. Since $Z_\tau=V_\tau$, we can define $(X_t)_{t\in[0,T]}$ by

$$X_t = \begin{cases} Z_t & \forall t\in[0,\tau] \\ V_t & \forall t\in[\tau,T]\end{cases}$$

If $t\in[0,\tau]$, then

$$\begin{aligned}
X_t &= Z_t = V_\tau + \int_t^\tau g(s,Z_s,\sigma_s(Z))\mathrm{d}s - \int_t^\tau \sigma_s(Z)\,\mathrm{d}B_s\\
&= V_\tau + \int_t^\tau g(s,X_s,\sigma_s(X))\mathrm{d}s - \int_t^\tau \sigma_s(X)\,\mathrm{d}B_s\\
&= \left(\xi + \int_\tau^T g(s,V_s,\sigma_s(V))\mathrm{d}s - \int_\tau^T \sigma_s(V)\,\mathrm{d}B_s\right) + \int_t^\tau g(s,X_s,\sigma_s(X))\mathrm{d}s - \int_t^\tau \sigma_s(X)\,\mathrm{d}B_s\\
&= \xi + \int_\tau^T g(s,X_s,\sigma_s(X))\mathrm{d}s - \int_\tau^T \sigma_s(X)\,\mathrm{d}B_s + \int_t^\tau g(s,X_s,\sigma_s(X))\mathrm{d}s - \int_t^\tau \sigma_s(X)\,\mathrm{d}B_s\\
&= \xi + \int_t^T g(s,X_s,\sigma_s(X))\mathrm{d}s - \int_t^T \sigma_s(X)\,\mathrm{d}B_s
\end{aligned}$$

and the same equality holds for $t\in(\tau,T]$ because $X_t=V_t$ and $(V_t)_{t\in[\tau,T]}\in$ BSDE$(g,\xi,[\tau,T])$. Therefore $X\in$ BSDE$(g,\xi,[0,T])$ and, by the maximality of $Y^{g,\xi}$,

$$Y^{g,\xi}\ge X$$

Hence, $(Y_t^{g,\xi})_{t\in[\tau,T]} \geq (X_t)_{t\in[\tau,T]} = (V_t)_{t\in[\tau,T]}$, as desired.

(ii) In view of (80), $(Y_t^{g,\xi})_{t\in[0,\tau]}$ is a nice Ito process that solves the BSDE$(g, Y_\tau^{g,\xi}, [0,\tau])$. Consider the maximal solution $(Z_t)_{t\in[0,\tau]}$ of BSDE$(g, Y_\tau^{g,\xi}, [0,\tau])$. Since $Z_\tau = Y_\tau^{g,\xi}$, we can define $(X_t)_{t\in[0,T]}$ by

$$X_t = \begin{cases} Z_t & \forall t \in [0,\tau] \\ Y_t^{g,\xi} & \forall t \in [\tau,T] \end{cases}$$

If $t \in [0,\tau]$, then

$$\begin{aligned} X_t &= Z_t = Y_\tau^{g,\xi} + \int_t^\tau g(s, Z_s, \sigma_s(Z))\mathrm{d}s - \int_t^\tau \sigma_s(Z)\,\mathrm{d}B_s \\ &= Y_\tau^{g,\xi} + \int_t^\tau g(s, X_s, \sigma_s(X))\mathrm{d}s - \int_t^\tau \sigma_s(X)\,\mathrm{d}B_s \\ &= \left(\xi + \int_\tau^T g(s, Y_s^{g,\xi}, \sigma_s(Y^{g,\xi}))\mathrm{d}s - \int_\tau^T \sigma_s(Y^{g,\xi})\,\mathrm{d}B_s\right) + \int_t^\tau g(s, X_s, \sigma_s(X))\mathrm{d}s - \int_t^\tau \sigma_s(X)\,\mathrm{d}B_s \\ &= \xi + \int_\tau^T g(s, X_s, \sigma_s(X))\mathrm{d}s - \int_\tau^T \sigma_s(X)\,\mathrm{d}B_s + \int_t^\tau g(s, X_s, \sigma_s(X))\mathrm{d}s - \int_t^\tau \sigma_s(X)\,\mathrm{d}B_s \\ &= \xi + \int_t^T g(s, X_s, \sigma_s(X))\mathrm{d}s - \int_t^T \sigma_s(X)\,\mathrm{d}B_s \end{aligned}$$

and the same equality holds for $t \in (\tau, T]$ because $X_t = Y_t^{g,\xi}$ and $(Y_t^{g,\xi})_{t\in[\tau,T]} \in \mathrm{BSDE}(g,\xi,[\tau,T])$. Therefore $X \in \mathrm{BSDE}(g,\xi,[0,T])$ and, by the maximality of $Y^{g,\xi}$,

$$Y^{g,\xi} \geq X$$

Hence, $(Y_t^{g,\xi})_{t\in[0,\tau]} \geq (X_t)_{t\in[0,\tau]} = (Z_t)_{t\in[0,\tau]}$, as desired. ■

Proof of Proposition 13 Assume that, for all $(t,\omega,y,z) \in [\tau,T] \times \Omega \times \mathbb{R}^2$,

$$g(t,\omega,y,z) \geq h(t,\omega,y,z) \qquad \text{and} \qquad \xi(\omega) \geq \zeta(\omega)$$

Define, for all $(t,\omega,y,z) \in [0,T] \times \Omega \times \mathbb{R}^2$,

$$\bar{g}(t,\omega,y,z) = 1_{[\tau,T]}(t)\, g(t,\omega,y,z) \qquad \text{and} \qquad \bar{h}(t,\omega,y,z) = 1_{[\tau,T]}(t)\, g(t,\omega,y,z)$$

Now $\bar{g}$ and $\bar{h}$ are quadratic drivers because g and h are. For all $(t,\omega,y,z) \in [0,T] \times \Omega \times \mathbb{R}^2$, we have

$$\bar{g}(t,\omega,y,z) \geq \bar{h}(t,\omega,y,z) \qquad \text{and} \qquad \xi(\omega) \geq \zeta(\omega)$$

By Theorem 8, $(Y_t^{\bar{g},\xi})_{t\in[0,T]} \geq (Y_t^{\bar{h},\zeta})_{t\in[0,T]}$. By Lemma 14, $(Y_t^{g,\xi})_{t\in[\tau,T]} = (Y_t^{\bar{g},\xi})_{t\in[\tau,T]}$ and $(Y_t^{h,\zeta})_{t\in[\tau,T]} = (Y_t^{\bar{h},\zeta})_{t\in[\tau,T]}$, as desired. ■

B.2 Equivalence of quadratic BDSEs and recursions

The next result shows the relevance in our setting of quadratic BSDEs.[51]

Theorem 15 *Let g be a quadratic driver. The following conditions are equivalent for a pair (Y,Z).*

[51] Since in this section we consider only BSDEs on $[0,T]$, we write BSDE(g,ξ) instead of BSDE$(g,\xi,[0,T])$.

(i) (Y, Z) *is a solution of* BSDE (g, ξ)*;*

(ii) Y *is a nice Ito process such that*

$$Y_t = \mathrm{E}_t \left[\xi + \int_t^T g(s, Y_s, \sigma_s(Y))\, \mathrm{d}s \right] \qquad \forall t \in [0, T]$$

and $Z = \sigma(Y)$*.*

We prove this result through a few lemmas.

Lemma 16 *If* (Y, Z) *is a solution of* BSDE (g, ξ), *then* Y *is a nice Ito process such that*

$$Y_t = \mathrm{E}_t \left[\xi + \int_t^T g(s, Y_s, Z_s)\, \mathrm{d}s \right]$$

for all $t \in [0, T]$*.*

Proof For each $t \in [0, T]$,

$$\begin{aligned}
Y_t &= \xi + \int_t^T g(s, Y_s, Z_s)\, \mathrm{d}s - \int_t^T Z_s \mathrm{d}B_s \text{ (but } Y_t \text{ is } \mathcal{F}_t \text{ measurable)} \\
&= \mathrm{E}_t \left[\xi + \int_t^T g(s, Y_s, Z_s)\, \mathrm{d}s \right] - \mathrm{E}_t \left[\int_t^T Z_s \mathrm{d}B_s \right] \\
&= \mathrm{E}_t \left[\xi + \int_t^T g(s, Y_s, Z_s)\, \mathrm{d}s \right] - \left(\mathrm{E}_t \left[\int_t^T Z_s \mathrm{d}B_s \right] + \mathrm{E}_t \left[\int_0^t Z_s \mathrm{d}B_s \right] - \mathrm{E}_t \left[\int_0^t Z_s \mathrm{d}B_s \right] \right) \\
&= \mathrm{E}_t \left[\xi + \int_t^T g(s, Y_s, Z_s)\, \mathrm{d}s \right] - \left(\mathrm{E}_t \left[\int_0^T Z_s \mathrm{d}B_s \right] - \mathrm{E}_t \left[\int_0^t Z_s \mathrm{d}B_s \right] \right) \text{ (but } \int_0^t Z_s \mathrm{d}B_s \text{ is a martingale)} \\
&= \mathrm{E}_t \left[\xi + \int_t^T g(s, Y_s, Z_s)\, \mathrm{d}s \right] - \left(\int_0^t Z_s \mathrm{d}B_s - \int_0^t Z_s \mathrm{d}B_s \right) \\
&= \mathrm{E}_t \left[\xi + \int_t^T g(s, Y_s, Z_s)\, \mathrm{d}s \right]
\end{aligned}$$

as desired. ■

In view of this lemma, by Proposition 12 it follows that (i) implies (ii). Next we prove the converse implication.

Lemma 17 *Let* g *be a quadratic driver. If* Y *is a nice Ito process such that*

$$Y_t = \mathrm{E}_t \left[\xi + \int_t^T g(s, Y_s, \sigma_s(Y))\, \mathrm{d}s \right]$$

for all $t \in [0, T]$, *then* $Y \in$ BSDE (g, ξ)*.*

Proof Set $\sigma = \sigma(Y)$. Note that $\xi \in L_T^\infty$. Moreover,

$$\int_0^T |g(s, Y_s, \sigma_s)|\, \mathrm{d}s \le (A + B \|Y\|_\infty) T + C \int_0^T |\sigma_s|^2\, \mathrm{d}s$$

and $\int_0^T |\sigma_s|^2 \mathrm{d}s < \infty$ because $\sigma \in L^2$. *A fortiori*, $\int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s \in L_T^0$. Also observe that (again, because $\sigma \in L^2$)

$$\mathrm{E}\left[\int_0^T |g(s, Y_s, \sigma_s)| \mathrm{d}s\right] \leq (A + B\|Y\|_\infty) T + C\mathrm{E}\left[\int_0^T |\sigma_s|^2 \mathrm{d}s\right] < \infty$$

Then, $\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s$ is summable. In turn, this implies that

$$M_t = \mathrm{E}_t\left[\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s\right]$$

is a martingale. By the Martingale Representation Theorem (see, e.g., Le Gall, 2016, p. 127 and Theorem 5.7 of Lipster and Shiryaev, 2001), there exists a unique predictable process $Z = (Z_t)_{t\in[0,T]}$ such that $\mathbb{P}\left(\int_0^T Z_s^2 \mathrm{d}s < \infty\right) = 1$ and

$$M_t = M_0 + \int_0^t Z_s \mathrm{d}B_s$$

That is,

$$\mathrm{E}_t\left[\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s\right] = \mathrm{E}\left[\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s\right] + \int_0^t Z_s \mathrm{d}B_s$$

which for $t = T$ amounts to

$$\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s = \mathrm{E}\left[\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s\right] + \int_0^T Z_s \mathrm{d}B_s$$

By replacing the first equation into the second, we have

$$\begin{aligned}
\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s &= \mathrm{E}\left[\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s\right] + \int_0^t Z_s \mathrm{d}B_s + \int_t^T Z_s \mathrm{d}B_s \\
&= \mathrm{E}_t\left[\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s\right] + \int_t^T Z_s \mathrm{d}B_s
\end{aligned}$$

Hence

$$\begin{aligned}
\xi &= \mathrm{E}_t\left[\xi + \int_t^T g(s, Y_s, \sigma_s) \mathrm{d}s\right] + \int_0^t g(s, Y_s, \sigma_s) \mathrm{d}s + \int_t^T Z_s \mathrm{d}B_s - \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s \\
&= Y_t - \int_t^T g(s, Y_s, \sigma_s) \mathrm{d}s + \int_t^T Z_s \mathrm{d}B_s
\end{aligned}$$

But then

$$\begin{aligned}
Y_t &= \xi + \int_t^T g(s, Y_s, \sigma_s) \mathrm{d}s - \int_t^T Z_s \mathrm{d}B_s \\
&= \xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s - \int_0^t g(s, Y_s, \sigma_s) \mathrm{d}s - \int_0^T Z_s \mathrm{d}B_s + \int_0^t Z_s \mathrm{d}B_s \\
&= \mathrm{E}\left[\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s\right] + \int_0^T Z_s \mathrm{d}B_s - \int_0^t g(s, Y_s, \sigma_s) \mathrm{d}s - \int_0^T Z_s \mathrm{d}B_s + \int_0^t Z_s \mathrm{d}B_s \\
&= \mathrm{E}\left[\xi + \int_0^T g(s, Y_s, \sigma_s) \mathrm{d}s\right] - \int_0^t g(s, Y_s, \sigma_s) \mathrm{d}s + \int_0^t Z_s \mathrm{d}B_s
\end{aligned}$$

Since Y is an Ito process and its decomposition is unique, then $Z = \sigma(Y)$. The first line of the previous batch then shows that

$$Y_t = \xi + \int_t^T g(s, Y_s, \sigma_s(Y))\, \mathrm{d}s - \int_t^T \sigma_s(Y)\, \mathrm{d}B_s$$

as desired (cf. Proposition 12). ■

Next we state an important consequence of Theorem 15, a version relevant here of a classic integral equality (see, e.g., Dynkin, 2002, p. 61). In reading it, note that, given any $\beta \in \mathbb{R}$, $g(s, y, z)$ is a quadratic driver if and only if $g(s, y, z) - \beta y$ is.

Corollary 18 *Let h be a quadratic driver and $\xi \in L^\infty(\Omega, \mathcal{F}_T, \mathbb{P})$. The following conditions are equivalent for a nice Ito process Y with $\sigma(Y) = Z$:*

(i) $Y_t = \mathrm{E}_t\left[\xi + \int_t^T (h(s, Y_s, Z_s) - \beta Y_s)\, \mathrm{d}s\right]$ *for all* $t \in [0, T]$,

(ii) $Y_t = \mathrm{E}_t\left[e^{-\beta(T-t)}\xi + \int_t^T e^{-\beta(s-t)} h(s, Y_s, Z_s)\, \mathrm{d}s\right]$ *for all* $t \in [0, T]$.

Proof Set $\mathrm{d}X_t = -\beta X_t \mathrm{d}t$, with solution $X_t = e^{-\beta t}$ and notice that:

- X_t is differentiable with continuous derivative;
- there exist two constants $\alpha, \gamma > 0$ such that $\alpha < X_t = e^{-\beta t} < \gamma$ for all $t \in [0, T]$.

(i) implies (ii). Let Y be a nice Ito process with $\sigma(Y) = Z$ that satisfies

$$Y_t = \mathrm{E}_t\left[\xi + \int_t^T (h(s, Y_s, Z_s) - \beta Y_s)\, \mathrm{d}s\right] \qquad \forall t \in [0, T]$$

The function $h(s, y, z) - \beta y$ is a quadratic driver. Then, in view of Theorem 15, Y is an Ito process that satisfies the BSDE

$$\mathrm{d}Y_t = -(h(t, Y_t, Z_t) - \beta Y_t)\mathrm{d}t + Z_t \mathrm{d}B_t \qquad Y_T = \xi$$

Since X has finite variation, then

$$\begin{aligned}
X_t Y_t &= X_0 Y_0 + \int_0^t X_s \mathrm{d}Y_s + \int_0^t Y_s \mathrm{d}X_s \\
&= X_0 Y_0 - \int_0^t X_s(h(s, Y_s, Z_s) - \beta Y_s)\mathrm{d}s + \int_0^t X_s Z_s \mathrm{d}B_s - \int_0^t Y_s \beta X_s \mathrm{d}s \\
&= X_0 Y_0 - \int_0^t X_s h(s, Y_s, Z_s)\, \mathrm{d}s + \beta \int_0^t X_s Y_s \mathrm{d}s + \int_0^t X_s Z_s \mathrm{d}B_s - \beta \int_0^t X_s Y_s \mathrm{d}s \\
&= X_0 Y_0 - \int_0^t X_s h(s, Y_s, Z_s)\, \mathrm{d}s + \int_0^t X_s Z_s \mathrm{d}B_s
\end{aligned}$$

Hence,

$$\begin{aligned}
X_t Y_t &= X_0 Y_0 - \int_0^t X_s h(s, Y_s, Z_s)\, \mathrm{d}s + \int_0^t X_s Z_s \mathrm{d}B_s \\
X_T Y_T &= X_0 Y_0 - \int_0^T X_s h(s, Y_s, Z_s)\, \mathrm{d}s + \int_0^T X_s Z_s \mathrm{d}B_s
\end{aligned}$$

and so

$$X_tY_t - X_TY_T = \int_t^T X_s h(s, Y_s, Z_s)\,\mathrm{d}s - \int_t^T X_s Z_s \mathrm{d}B_s$$

Since X_tY_t is $\mathcal{F}_t$ measurable, in turn this implies

$$X_tY_t = \mathrm{E}_t[X_tY_t] = \mathrm{E}_t\left[X_TY_T + \int_t^T X_s h(s, Y_s, Z_s)\,\mathrm{d}s\right]$$

because $\mathrm{E}_t\left[\int_t^T X_s Z_s \mathrm{d}B_s\right] = \mathrm{E}_t\left[\int_t^T e^{-\beta s} Z_s \mathrm{d}B_s\right] = 0$ since $\left(e^{-\beta t}Z_t\right)_{t\in[0,T]} \in L^2$ (as Z does). So, for each $t \in [0, T]$ we have (notice that X_t is strictly positive):

$$Y_t = \frac{X_tY_t}{X_t} = \mathrm{E}_t\left[\frac{X_T}{X_t}Y_T + \int_t^T \frac{X_s}{X_t} h(s, Y_s, Z_s)\,\mathrm{d}s\right] = \mathrm{E}_t\left[e^{-\beta(T-t)}\xi + \int_t^T e^{-\beta(s-t)} h(s, Y_s, Z_s)\,\mathrm{d}s\right]$$

as desired.

(ii) implies (i). If $Y_t = \mathrm{E}_t\left[e^{-\beta(T-t)}\xi + \int_t^T e^{-\beta(s-t)} h(s, Y_s, Z_s)\,\mathrm{d}s\right]$ for all $t \in [0, T]$, then (recall that $X_t = e^{-\beta t}$)

$$\begin{aligned} Y_t &= \mathrm{E}_t\left[e^{-\beta(T-t)}\xi + \int_t^T e^{-\beta(s-t)} h(s, Y_s, Z_s)\,\mathrm{d}s\right] = \mathrm{E}_t\left[\frac{e^{-\beta T}}{e^{-\beta t}}Y_T + \int_t^T \frac{e^{-\beta s}}{e^{-\beta t}} h(s, Y_s, Z_s)\,\mathrm{d}s\right] \\ &= \mathrm{E}_t\left[\frac{X_T}{X_t}Y_T + \int_t^T \frac{X_s}{X_t} h(s, Y_s, Z_s)\,\mathrm{d}s\right] \end{aligned}$$

So,

$$X_tY_t = \mathrm{E}_t\left[X_TY_T + \int_t^T X_s h(s, Y_s, Z_s)\,\mathrm{d}s\right] = \mathrm{E}_t\left[e^{-\beta T}\xi + \int_t^T e^{-\beta s} h\left(s, \frac{X_sY_s}{e^{-\beta s}}, \frac{X_sZ_s}{e^{-\beta s}}\right)\mathrm{d}s\right] \tag{81}$$

But, Y is a nice Ito process,

$$\mathrm{d}Y_t = F_t\mathrm{d}t + Z_t\mathrm{d}B_t \qquad Y_T = \xi$$

Since X has finite variation, then

$$\begin{aligned} X_tY_t &= X_0Y_0 + \int_0^t X_s\mathrm{d}Y_s + \int_0^t Y_s\mathrm{d}X_s \\ &= X_0Y_0 + \int_0^t X_sF_s\mathrm{d}s + \int_0^t X_sZ_s\mathrm{d}B_s + \int_0^t (-Y_s\beta X_s)\,\mathrm{d}s \\ &= X_0Y_0 + \int_0^t (X_sF_s - Y_s\beta X_s)\,\mathrm{d}s + \int_0^t X_sZ_s\mathrm{d}B_s = X_0Y_0 + \int_0^t X_s(F_s - \beta Y_s)\,\mathrm{d}s + \int_0^t X_sZ_s\mathrm{d}B_s \end{aligned}$$

and

$$X_tY_t = X_0Y_0 + \int_0^t X_s(F_s - \beta Y_s)\,\mathrm{d}s + \int_0^t X_sZ_s\mathrm{d}B_s \qquad X_TY_T = e^{-\beta T}\xi \tag{82}$$

Moreover, $XY = (X_tY_t)_{t\in[0,T]}$ is a continuous stochastic process and

$$\begin{aligned} \int_0^T |X_s(F_s - \beta Y_s)|\,\mathrm{d}s &\leq \int_0^T \gamma\,|(F_s - \beta Y_s)|\,\mathrm{d}s \leq \gamma\int_0^T |F_s|\,\mathrm{d}s + \gamma\beta\int_0^T |Y_s|\,\mathrm{d}s < \infty \\ &\leq \gamma\int_0^T |F_s|\,\mathrm{d}s + \gamma\beta\,\|Y\|_\infty T < \infty \end{aligned}$$

and $\int_0^T |X_s Z_s|^2 \, \mathrm{d}s \leq \gamma^2 \int_0^T |Z_s|^2 \, \mathrm{d}s < \infty$. So XY is an Ito process. It is also nice, in fact, $XY, \sigma(XY) = XZ \in L^2$ and $\|XY\|_\infty < \infty$.[52]

Now the terminal condition $e^{-\beta T}\xi$ of (81) is bounded and the driver

$$g(s,\omega,u,v) = e^{-\beta s} h\left(s, \omega, \frac{u}{e^{-\beta s}}, \frac{v}{e^{-\beta s}}\right)$$

is quadratic

$$|g(s,\omega,u,v)| \leq \gamma \left| h\left(s,\omega,\frac{u}{e^{-\beta s}},\frac{v}{e^{-\beta s}}\right)\right| \leq \gamma\left(A + \frac{B}{e^{-\beta s}}|u| + \frac{C}{e^{-\beta s}}|v|^2\right) \leq \gamma\left(A + \frac{B}{\alpha}|u| + \frac{C}{\alpha}|v|^2\right)$$

so that (81) can be rewritten as

$$X_t Y_t = \mathrm{E}_t\left[e^{-\beta T}\xi + \int_t^T e^{-\beta s} h\left(s, \frac{X_s Y_s}{e^{-\beta s}}, \frac{\sigma_s(XY)}{e^{-\beta s}}\right) \mathrm{d}s\right] = \mathrm{E}_t\left[e^{-\beta T}\xi + \int_t^T g\left(s, X_s Y_s, \sigma_s(XY)\right) \mathrm{d}s\right]$$

Then, in view of Theorem 15, XY is an Ito process that satisfies the BSDE

$$\mathrm{d}(X_t Y_t) = -e^{-\beta s} h\left(s, \frac{X_s Y_s}{e^{-\beta s}}, \frac{\sigma_s(XY)}{e^{-\beta s}}\right) \mathrm{d}t + X_t Z_t \mathrm{d}B_t \qquad X_T Y_T = e^{-\beta T}\xi \tag{83}$$

By the uniqueness of the Ito decomposition of XY, we then have

$$\begin{aligned}
-e^{-\beta s} h(s, Y_s, Z_s) &= -e^{-\beta s} h\left(s, \frac{X_s Y_s}{e^{-\beta s}}, \frac{\sigma_s(XY)}{e^{-\beta s}}\right) = X_s(F_s - \beta Y_s) \\
-h(s, Y_s, Z_s) &= F_s - \beta Y_s \\
F_s &= -(h(s, Y_s, Z_s) - \beta Y_s)
\end{aligned}$$

Finally,

$$\mathrm{d}Y_t = -(h(t, Y_t, Z_t) - \beta Y_t)\,\mathrm{d}t + Z_t \mathrm{d}B_t \qquad Y_T = \xi$$

But this means that (Y, Z) is a solution of BSDE $(h(s,y,z) - \beta y, \xi)$. Another application of Theorem 15 concludes the proof. ■

B.3 Parametric BSDEs

We say that a $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^3)$-measurable function $f : [0,T] \times \Omega \times \mathbb{R}^3 \to \mathbb{R}$ is a *parametric quadratic driver* if

(i) $f(t,\omega,\theta,\cdot,\cdot) : \mathbb{R}^2 \to \mathbb{R}$ is continuous for all $(t,\omega,\theta) \in [0,T] \times \Omega \times \mathbb{R}$,

(ii) $|f(t,\omega,\theta,y,z)| \leq A + B|y| + C|z|^2$ for some $A, B, C \in \mathbb{R}$ and all $(t,\omega,\theta,y,z) \in [0,T] \times \Omega \times \mathbb{R}^3$.

A parametric quadratic driver f is *monotone* if it is increasing in θ.

Lemma 19 *If f is a parametric quadratic driver then, for each process $\boldsymbol{\theta} \in L^0$, the function $f_{\boldsymbol{\theta}} : [0,T] \times \Omega \times \mathbb{R}^2 \to \mathbb{R}$ defined by*

$$f_{\boldsymbol{\theta}}(t,\omega,y,z) = f(t,\omega,\boldsymbol{\theta}_t(\omega), y, z)$$

is a quadratic driver.

[52] Note that if $W \in L^2$, so does XW.

Proof Let $\boldsymbol{\theta} \in L^0$. The $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^2)$-measurability of $f_{\boldsymbol{\theta}}$ is a long routine exercise. For all $(t,\omega) \in [0,T]\times\Omega$, $f_{\boldsymbol{\theta}}(t,\omega,\cdot,\cdot) = f(t,\omega,\boldsymbol{\theta}_t(\omega),\cdot,\cdot) : \mathbb{R}^2 \to \mathbb{R}$ is continuous. Moreover, for all $(t,\omega,y,z) \in [0,T]\times\Omega\times\mathbb{R}^2$ we have

$$|f_{\boldsymbol{\theta}}(t,\omega,y,z)| = |f(t,\omega,\boldsymbol{\theta}_t(\omega),y,z)| \leq A + B|y| + C|z|^2$$

as desired. ■

Given $\tau \in [0,T]$, denote by ${}_\tau\boldsymbol{\theta} = (\boldsymbol{\theta}_t)_{t\in[\tau,T]}$ the restriction of process $\boldsymbol{\theta}$ on $[\tau,T]$. The restriction of $f_{\boldsymbol{\theta}}$ on $[\tau,T]\times\Omega\times\mathbb{R}^2$ is the map $f_{{}_\tau\boldsymbol{\theta}} : [\tau,T]\times\Omega\times\mathbb{R}^2$ given, for each $(t,\omega,y,z) \in [\tau,T]\times\Omega\times\mathbb{R}^2$, by

$$f_{{}_\tau\boldsymbol{\theta}}(t,\omega,y,z) = f_{\boldsymbol{\theta}}(t,\omega,y,z) = f(t,\omega,\boldsymbol{\theta}_t(\omega),y,z) = f(t,\omega,({}_\tau\boldsymbol{\theta})_t(\omega),y,z)$$

It only depends on ${}_\tau\theta$, so the notation $f_{{}_\tau\boldsymbol{\theta}}$.

Analogously, denote by $\boldsymbol{\theta}^\tau = (\boldsymbol{\theta}_t)_{t\in[0,\tau]}$ the restriction of process $\boldsymbol{\theta}$ on $[0,\tau]$. The restriction of $f_{\boldsymbol{\theta}}$ on $[0,\tau]\times\Omega\times\mathbb{R}^2$ is the map $f_{\boldsymbol{\theta}^\tau} : [0,\tau]\times\Omega\times\mathbb{R}^2$ given, for each $(t,\omega,y,z) \in [\tau,T]\times\Omega\times\mathbb{R}^2$, by

$$f_{\boldsymbol{\theta}^\tau}(t,\omega,y,z) = f_{\boldsymbol{\theta}}(t,\omega,y,z) = f(t,\omega,\boldsymbol{\theta}_t(\omega),y,z) = f(t,\omega,(\boldsymbol{\theta}^\tau)_t(\omega),y,z)$$

It only depends on θ^τ, so the notation $f_{\boldsymbol{\theta}^\tau}$.

Let f be a parametric quadratic driver and $v : \mathbb{R} \to \mathbb{R}$ a bounded and Borel measurable function. Given a subset $\mathcal{L}$ of L^0, we denote by $V : \mathcal{L} \to L^0$ the operator that associates to each $\boldsymbol{\theta} \in \mathcal{L}$ the maximal solution $V(\boldsymbol{\theta})$, within the family of nice Ito processes, of the forward recursion

$$V_t(\boldsymbol{\theta})(\omega) = \mathrm{E}_t\left[v(\boldsymbol{\theta}_T(\omega)) + \int_t^T f(s,\omega,\boldsymbol{\theta}_s(\omega),V_s(\boldsymbol{\theta})(\omega),\sigma_s(V(\boldsymbol{\theta}))(\omega))\,\mathrm{d}s\right](\omega) \quad \forall (t,\omega) \in [0,T]\times\Omega$$

In view of Lemma 19, the following result follows from Theorem 15.

Proposition 20 *If f is a parametric quadratic driver and $v : \mathbb{R} \to \mathbb{R}$ is a bounded and Borel measurable function, then*

$$V(\boldsymbol{\theta}) = Y^{f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T)} \qquad \forall \boldsymbol{\theta} \in \mathcal{L}$$

where $Y^{f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T)}$ is the maximal solution of the quadratic $\mathrm{BSDE}(f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T),[0,T])$.

The next result builds upon Lemma 14-(i).

Lemma 21 *Let f be a parametric quadratic driver, $v : \mathbb{R} \to \mathbb{R}$ a bounded and Borel measurable function and $\boldsymbol{\theta} \in \mathcal{L}$. If $\tau \in [0,T)$, then $V_\tau(\boldsymbol{\theta})$ is the initial value $Y_\tau^{f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T)}$ of the maximal solution of the* $\mathrm{BSDE}(f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T),[\tau,T])$.

Since this value only depends on ${}_\tau\boldsymbol{\theta}$, with an abuse notation we write $Y_\tau^{f_{{}_\tau\boldsymbol{\theta}},v(\boldsymbol{\theta}_T)} = Y_\tau^{f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T)}$ and $V_\tau({}_\tau\boldsymbol{\theta}) = V_\tau(\boldsymbol{\theta})$.

Proof By Lemma 14, if $(Y_t^{f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T)})_{t\in[0,T]}$ is the maximal solution of $\mathrm{BSDE}(f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T),[0,T])$, then $(Y_t^{f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T)})_{t\in[\tau,T]}$ is the maximal solution of $\mathrm{BSDE}(f_{{}_\tau\boldsymbol{\theta}},v(\boldsymbol{\theta}_T),[\tau,T])$, which only depends on the restriction $f_{{}_\tau\boldsymbol{\theta}}$ of $f_{\boldsymbol{\theta}}$ on $[\tau,T]\times\Omega\times\mathbb{R}^2$ that, in turn, depends only on ${}_\tau\boldsymbol{\theta}$. ■

Unlike the last result, next we assume monotonicity.

Proposition 22 *Let f be a monotone quadratic driver, $v : \mathbb{R} \to \mathbb{R}$ a bounded and Borel measurable function and $\tau \in [0, T)$. For each $\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}} \in \mathcal{L}$ we have*

$$_\tau\boldsymbol{\theta} \geq {}_\tau\tilde{\boldsymbol{\theta}} \quad \text{and} \quad v(\boldsymbol{\theta}_T) \geq v(\tilde{\boldsymbol{\theta}}_T) \Longrightarrow (V_t(\boldsymbol{\theta}))_{t\in[\tau,T]} \geq (V_t(\tilde{\boldsymbol{\theta}}))_{t\in[\tau,T]}$$

Proof Since f is increasing in θ, for any given t, ω, y, z, then

$$_\tau\boldsymbol{\theta} \geq {}_\tau\tilde{\boldsymbol{\theta}} \iff (\boldsymbol{\theta}_t(\omega))_{t\in[\tau,T]} \geq (\tilde{\boldsymbol{\theta}}_t(\omega))_{t\in[\tau,T]} \implies f_{\boldsymbol{\theta}} \geq f_{\tilde{\boldsymbol{\theta}}}$$

Since $v(\boldsymbol{\theta}_T) \geq v(\tilde{\boldsymbol{\theta}}_T)$, by Proposition 13 we then have $\left(Y_t^{f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T)}\right)_{t\in[\tau,T]} \geq \left(Y_t^{f_{\boldsymbol{\theta}'},v\left(\tilde{\boldsymbol{\theta}}_T\right)}\right)_{t\in[\tau,T]}$, as desired. ■

The next result deals with time consistency.

Proposition 23 *Let f be a quadratic driver, $v : \mathbb{R} \to \mathbb{R}$ a bounded and Borel measurable function and $\tau \in [0, T)$. For each $\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}} \in \mathcal{L}$ we have*

$$V_\tau(\boldsymbol{\theta}) \geq V_\tau(\tilde{\boldsymbol{\theta}}) \Longrightarrow (V_t(\boldsymbol{\theta}))_{t\in[0,\tau]} \geq (V_t(\tilde{\boldsymbol{\theta}}))_{t\in[0,\tau]}$$

provided $\boldsymbol{\theta}^\tau = \tilde{\boldsymbol{\theta}}^\tau$.

Proof By Proposition 20, $V(\boldsymbol{\theta}) = Y^{f_{\boldsymbol{\theta}},v(\boldsymbol{\theta}_T)}$ is the maximal solution of BSDE$(f_{\boldsymbol{\theta}}, v(\boldsymbol{\theta}_T), [0, T])$. By Lemma 14-(ii), $(V_t(\boldsymbol{\theta}))_{t\in[0,\tau]}$ is then the maximal solution of BSDE$(f_{\boldsymbol{\theta}}, V_\tau(\boldsymbol{\theta}), [0, \tau])$. Similarly, $(V_t(\tilde{\boldsymbol{\theta}}))_{t\in[0,\tau]}$ is the maximal solution of BSDE$\left(f_{\boldsymbol{\theta}'}, V_\tau(\tilde{\boldsymbol{\theta}}), [0, \tau]\right)$. We have that $\boldsymbol{\theta}^\tau = \tilde{\boldsymbol{\theta}}^\tau$ implies $f_{\boldsymbol{\theta}^\tau} = f_{\tilde{\boldsymbol{\theta}}^\tau}$. Since $V_\tau(\boldsymbol{\theta}) \geq V_\tau(\tilde{\boldsymbol{\theta}})$, by Theorem 8 we then have $(V_t(\boldsymbol{\theta}))_{t\in[0,\tau]} \geq (V_t(\tilde{\boldsymbol{\theta}}))_{t\in[0,\tau]}$. ■

Proposition 24 *Let f be a quadratic driver that satisfies condition (K), $v : \mathbb{R} \to \mathbb{R}$ a bounded and Borel measurable function and $t \in [0, T)$. For each $\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}} \in \mathcal{L}$ we have*

$$V_t(\boldsymbol{\theta}) = V_t(\tilde{\boldsymbol{\theta}}) \Longrightarrow V_0(\boldsymbol{\theta}) = V_0(\tilde{\boldsymbol{\theta}})$$

and

$$V_t(\boldsymbol{\theta}) > V_t(\tilde{\boldsymbol{\theta}}) \Longrightarrow V_0(\boldsymbol{\theta}) > V_0(\tilde{\boldsymbol{\theta}})$$

provided $\boldsymbol{\theta}^\tau = \tilde{\boldsymbol{\theta}}^\tau$.

Proof By Theorem 9, $V(\boldsymbol{\theta}) = (V_t(\boldsymbol{\theta}))$ is unique for each $\boldsymbol{\theta} \in \mathcal{L}$. Let $V_t(\boldsymbol{\theta}) > V_t(\tilde{\boldsymbol{\theta}})$, i.e., $V_t(\boldsymbol{\theta}) \geq V_t(\tilde{\boldsymbol{\theta}})$ and $V_t(\boldsymbol{\theta}) \neq V_t(\tilde{\boldsymbol{\theta}})$. By Proposition 23, $V_0(\boldsymbol{\theta}) \geq V_0(\tilde{\boldsymbol{\theta}})$. Suppose, *per contra*, that $V_0(\boldsymbol{\theta}) = V_0(\tilde{\boldsymbol{\theta}})$. By Lemma 14, $(V_s(\boldsymbol{\theta}))_{s\in[0,t]}$ solves the BSDE$(f_{\boldsymbol{\theta}^\tau}, V_t(\boldsymbol{\theta}), [0, t])$ and $(V_s(\tilde{\boldsymbol{\theta}}))_{s\in[0,t]}$ solves the BSDE$(f_{\tilde{\boldsymbol{\theta}}^\tau}, V_t(\tilde{\boldsymbol{\theta}}), [0, t])$. Since $\boldsymbol{\theta}^\tau = \tilde{\boldsymbol{\theta}}^\tau$, we have $f_{\boldsymbol{\theta}^\tau} = f_{\tilde{\boldsymbol{\theta}}^\tau}$. By (70), from $V_0(\boldsymbol{\theta}) = V_0(\tilde{\boldsymbol{\theta}})$ it then follows that $V_t(\boldsymbol{\theta}) = V_t(\tilde{\boldsymbol{\theta}})$ since we assumed that $V_t(\boldsymbol{\theta}) \geq V_t(\tilde{\boldsymbol{\theta}})$. This contradiction proves that $V_0(\boldsymbol{\theta}) > V_0(\tilde{\boldsymbol{\theta}})$. ■

B.4 Proofs

Proof of Lemma 1 Let V be a dynamic utility operator for $\succsim$. By A.4, Fix t. By A.4, there exists $A_t \in \mathcal{P}$ with $\mathbb{P}(A_t) = 1$ such that, for each $c \in C$, $V_{t,\omega}(\mathbf{1}_c) = u(c) \int_t^T e^{-\beta(s-t)} \, \mathrm{d}s = u(c)(1 - e^{-\beta(T-t)})/\beta$ at all $\omega \in A_t$. Since u is bounded, for each $c \in C$ it holds $|V_{t,\omega}(\mathbf{1}_c)| \leq \|u\|_\infty (1 - e^{-\beta(T-t)})/\beta \leq \|u\|_\infty / \beta$ at all $\omega \in A_t$.

Fix $\mathbf{c} \in \mathbf{C}$. Given $\omega \in A_t$, let $c \in C$ be such that $\mathbf{c} \sim_{t,\omega} \mathbf{1}_c$. Then, $|V_{t,\omega}(\mathbf{c})| = |V_{t,\omega}(\mathbf{1}_c)| \leq \|u\|_\infty / \beta$. Since $\mathbf{c}$ was arbitrarily chosen in $\mathbf{C}$, we conclude that $\|V_{t,\omega}\|_\infty \leq \|u\|_\infty / \beta$ at all $\omega \in A_t$. Hence, $\|V_t\|_\infty = \text{essup}_{\omega \in \Omega} |V_t(\omega)| \leq \|u\|_\infty / \beta$ and so $\|V\|_\infty = \text{essup}_{(\omega,t) \in \Omega \times [0,T]} |V_t(\omega)| \leq \|u\|_\infty / \beta$. Thus, V belongs to L^∞, i.e., it is bounded.

Let V and $\tilde{V}$ be dynamic utility operators for $\succsim$. Fix t. By A.4, there exists $A_t \in \mathcal{P}$ with $\mathbb{P}(A_t) = 1$ such that, for each $c \in C$, $V_{t,\omega}(\mathbf{1}_c) = u(c)(1 - e^{-\beta(T-t)})/\beta$ at all $\omega \in A_t$. Similarly, $\tilde{V}_{t,\omega}(\mathbf{1}_c) = \tilde{u}(c)(1 - e^{-\beta(T-t)})/\beta$ for all $\omega \in \tilde{A}_t$. By the cardinality of u, there exist $a > 0$ and $b \in \mathbb{R}$ such that $\tilde{u} = au + b$. In turn, this implies the existence of scalars $\bar{a}_t > 0$ and $\bar{b}_t \in \mathbb{R}$ such that $V_{t,\omega}(\mathbf{1}_c) = \bar{a}_t \tilde{V}_{t,\omega}(\mathbf{1}_c) + \bar{b}_t$ for all $\omega \in \bar{A}_t = A_t \cap \tilde{A}_t$. Clearly, $\mathbb{P}(\bar{A}_t) = 1$.

Fix $\mathbf{c} \in \mathbf{C}$. Given $\omega \in A_t$, let $c \in C$ be such that $\mathbf{c} \sim_{t,\omega} \mathbf{1}_c$. Then,

$$V_{t,\omega}(\mathbf{c}) = V_{t,\omega}(\mathbf{1}_c) = \bar{a}_t \tilde{V}_{t,\omega}(\mathbf{1}_c) + \bar{b}_t = \bar{a}_t \tilde{V}_{t,\omega}(\mathbf{c}) + \bar{b}_t$$

Since ω was arbitrarily chosen in $\bar{A}_t$, we thus have $V_{t,\omega}(\mathbf{c}) = \bar{a}_t \tilde{V}_{t,\omega}(\mathbf{c}) + \bar{b}_t$ for all $\omega \in \bar{A}_t$. Since $\mathbf{c}$ was arbitrarily chosen in $\mathbf{C}$, we conclude that $V_{t,\omega} = \bar{a}_t \tilde{V}_{t,\omega} + \bar{b}_t$, as desired. ■

Define $\varphi : \Omega \times [0,T] \times \operatorname{Im} u \times \mathbb{R}^2 \to \mathbb{R}$ by $\varphi(\omega, t, u(c), y, z) = f(\omega, t, c, y, z)$. The function φ is well defined by properties (i) and (ii) of Definition 7. Given $\mathbf{c} \in \mathbf{C}$, define $u \circ \mathbf{c} \in L^0_+$ by $(u \circ \mathbf{c})_t(\omega) = u(\mathbf{c}_t(\omega))$. Define a driver $\varphi_{u \circ \mathbf{c}} : \Omega \times [0,T] \times \mathbb{R}^2 \to \mathbb{R}$ by

$$\varphi_{u \circ \mathbf{c}}(\omega, t, y, z) = \varphi(\omega, t, u(\mathbf{c}_t(\omega)), y, z) = u(\mathbf{c}_t(\omega)) - \beta y + \eta(\omega, t, y, z)$$

Note that, in view of what established so far in the appendix, the results that we are going to prove actually hold for a more general class of, possibly non-separable, aggregators f. For brevity, we omit details.

Proof of Proposition 2 We continue to denote by h the quadratic driver in Corollary 18. By setting $h(t, \omega, y, z) = u(\mathbf{c}_t(\omega)) + \eta(t, \omega, y, z)$, from this corollary we have

$$V_t(\mathbf{c}) = \mathrm{E}_t \left[e^{-\beta(T-t)} v(\mathbf{c}_T) + \int_t^T e^{-\beta(s-t)} \left(u(\mathbf{c}_s) + \eta(t, V_s(\mathbf{c}), \sigma_s(V(\mathbf{c}))) \right) \mathrm{d}s \right]$$

So,

$$\begin{aligned}
V_t(\mathbf{c}) &= \mathrm{E}_t\left[e^{-\beta(T-t)}\upsilon(\mathbf{c}_T)+\int_t^T e^{-\beta(s-t)}\left(u(\mathbf{c}_s)+\eta(t,V_s(\mathbf{c}),\sigma_s(V(\mathbf{c})))\right)\mathrm{d}s\right] \\
&= \mathrm{E}_t\left[\int_t^\tau e^{-\beta(s-t)}\left(u(\mathbf{c}_s)+\eta(t,V_s(\mathbf{c}),\sigma_s(V(\mathbf{c})))\right)\mathrm{d}s\right. \\
&\quad \left.+\mathrm{E}_\tau\left[e^{-\beta(T-t)}\upsilon(\mathbf{c}_T)+\int_\tau^T e^{-\beta(s-t)}\left(u(\mathbf{c}_s)+\eta(t,V_s(\mathbf{c}),\sigma_s(V(\mathbf{c})))\right)\mathrm{d}s\right]\right] \\
&= \mathrm{E}_t\left[\int_t^\tau e^{-\beta(s-t)}\left(u(\mathbf{c}_s)-\eta(t,V_s(\mathbf{c}),\sigma_s(V(\mathbf{c})))\right)\mathrm{d}s\right. \\
&\quad \left.+e^{-\beta(\tau-t)}\mathrm{E}_\tau\left[e^{-\beta(T-\tau)}\upsilon(\mathbf{c}_T)+\int_\tau^T e^{-\beta(s-\tau)}\left(u(\mathbf{c}_s)+\eta(t,V_s(\mathbf{c}),\sigma_s(V(\mathbf{c})))\right)\mathrm{d}s\right]\right] \\
&= \mathrm{E}_t\left[\int_t^\tau e^{-\beta(s-t)}\left(u(\mathbf{c}_s)+\eta(t,V_s(\mathbf{c}),\sigma_s(V(\mathbf{c})))\right)\mathrm{d}s+e^{-\beta(\tau-t)}V_\tau(\mathbf{c})\right]
\end{aligned}$$

as desired. In particular, if $\tau = T$ we have $V_T(\mathbf{c}) = \upsilon_T(\mathbf{c}) = 0$ and so

$$V_t(\mathbf{c}) = \mathrm{E}_t\left[\int_t^T e^{-\beta(s-t)}\left(u(\mathbf{c}_s)+\eta(t,V_s(\mathbf{c}),\sigma_s(V(\mathbf{c})))\right)\mathrm{d}s\right]$$

as desired. ■

Proof of Proposition 3 By Lemma 19, $\varphi_{u\circ\mathbf{c}}$ is a quadratic driver. Existence of a maximal bounded solution then follows from Theorem 7. Dynamic monotonicity follows from Proposition 22 and dynamic consistency from Proposition 23. ■

Proof of Proposition 4 Since condition (25) implies (K), uniqueness follows from Theorem 9. Using the driver $\varphi_{u\circ\mathbf{c}}$, strict dynamic consistency follows from Proposition 24. ■

Proof of Proposition 5 We have

$$\begin{aligned}
|\eta(t,y_1,z_1)-\eta(t,y_2,z_2)| &= \left|-\varsigma_t^2\upsilon\varphi(z_1)+\varsigma_t^2\upsilon\varphi(z_2)\right| = |\upsilon|\left|\varsigma_t^2\right|\left|\varphi(z_1)-\varphi(z_2)\right| \\
&\leq k\,|\upsilon|\left\|\varsigma^2\right\|_\infty |z_1-z_2|\left(|z_1|+|z_2|\right)
\end{aligned}$$

So, condition (25) is satisfied. By Proposition 4, the result holds. ■

Proof of Proposition 6 We prove the "only if" as the converse is easily checked. Since u and $\phi_{i,h}$ are both continuous, all acts have certainty equivalents and is thus enough to prove that $V_2 \geq V_1$ implies $\kappa_2 \leq \kappa_1$. To this end, consider a bet on states b_ε that pays c on some event $E \subseteq S$ and $c+\varepsilon$ otherwise, with $0 < \varepsilon \leq h$ (so that $b_\varepsilon \in \mathcal{H}$). For $i = 1, 2$, criterion (59) is

$$V_{i,h}(b_\varepsilon) = -\kappa_2 h \log \sum_\xi e^{-\frac{q^\xi u(c+\varepsilon)+(1-q^\xi)u(c)}{\kappa_2 h}}\pi(\xi)$$

where $\mathbb{Q}^{\xi}(E) = q^{\xi} \in (0,1)$ is the probability that model $\mathbb{Q}^{\xi}$ assigns to the event E, and π is the common belief over models. Normalize $u(c) = 0$. It holds:

$$\begin{aligned} V_{2,h}(b_\varepsilon) \geq V_{1,h}(b_\varepsilon) &\iff -\kappa_2 h \log \sum_{\xi} e^{-\frac{q^{\xi} u(c+\varepsilon)}{\kappa_2 h}} \pi(\xi) \geq -\kappa_1 h \log \sum_{\xi} e^{-\frac{q^{\xi} u(c+\varepsilon)}{\kappa_1 h}} \pi(\xi) \\ &\iff \kappa_2 \log \sum_{\xi} e^{-\frac{q^{\xi} u(c+\varepsilon)}{\kappa_2 h}} \pi(\xi) \leq \kappa_1 \log \sum_{\xi} e^{-\frac{q^{\xi} u(c+\varepsilon)}{\kappa_1 h}} \pi(\xi) \iff \kappa_2 \leq \kappa_1 \end{aligned}$$

where the last equivalence holds because the function

$$(0,\infty) \ni \kappa \longmapsto \kappa \log \sum_{\xi} e^{-\frac{q^{\xi} u(c+\varepsilon)}{\kappa h}} \pi(\xi) \in (-\infty, 0)$$

is strictly increasing (note that $q^{\xi} u(c+\varepsilon)/\kappa h > 0$ since $u(c+\varepsilon) > u(c) = 0$). We conclude that $V_2 \geq V_1$ implies $\kappa_2 \leq \kappa_1$, as desired. ■

B.5 Section 5

To ease notation we write $V_{t+h} = V_{t+h}(\mathbf{c})$, $\mu_t = \mu_t(V(\mathbf{c}))$ and $\sigma_t = \sigma_t(V(\mathbf{c}))$.

Eq. (42) Given two processes $\underline{\xi}, \bar{\xi} \in L^0$ with $\underline{\xi} \leq \bar{\xi}$, let Ξ be the collection of all processes $\xi \in L^0$ such that $\underline{\xi} \leq \xi \leq \bar{\xi}$. We have

$$\begin{aligned} \min_{\xi \in \Xi} (\sigma_t \xi_t) &= \begin{cases} \sigma_t \min_{\xi \in \Xi} \xi_t & \text{if } \sigma_t \geq 0 \\ \sigma_t \max_{\xi \in \Xi} \xi_t & \text{if } \sigma_t < 0 \end{cases} = \begin{cases} \sigma_t \underline{\xi}_t & \text{if } \sigma_t \geq 0 \\ \sigma_t \bar{\xi}_t & \text{if } \sigma_t < 0 \end{cases} \\ &= \frac{1}{2} \underline{\xi}_t \sigma_t [1 + \operatorname{sgn} \sigma_t] - \frac{1}{2} \bar{\xi}_t \sigma_t [-1 + \operatorname{sgn} \sigma_t] = \frac{1}{2} \left[\left(\underline{\xi}_t + \bar{\xi}_t \right) \sigma_t + \left(\underline{\xi}_t - \bar{\xi}_t \right) \sigma_t \operatorname{sgn} \sigma_t \right] \\ &= \frac{1}{2} \left[\left(\underline{\xi}_t + \bar{\xi}_t \right) \sigma_t + \left(\underline{\xi}_t - \bar{\xi}_t \right) |\sigma_t| \right] \end{aligned}$$

as well as

$$\begin{aligned} \max_{\xi \in \Xi} (\sigma_t \xi_t) &= \begin{cases} \sigma_t \max_{\xi \in \Xi} \xi_t & \text{if } \sigma_t \geq 0 \\ \sigma_t \min_{\xi \in \Xi} \xi_t & \text{if } \sigma_t < 0 \end{cases} = \begin{cases} \sigma_t \bar{\xi}_t & \text{if } \sigma_t \geq 0 \\ \sigma_t \underline{\xi}_t & \text{if } \sigma_t < 0 \end{cases} \\ &= \frac{1}{2} \bar{\xi}_t \sigma_t [1 + \operatorname{sgn} \sigma_t] - \frac{1}{2} \underline{\xi}_t \sigma_t [-1 + \operatorname{sgn} \sigma_t] = \frac{1}{2} \left[\left(\underline{\xi}_t + \bar{\xi}_t \right) \sigma_t + \left(\bar{\xi}_t - \underline{\xi}_t \right) \sigma_t \operatorname{sgn} \sigma_t \right] \\ &= \frac{1}{2} \left[\left(\underline{\xi}_t + \bar{\xi}_t \right) \sigma_t + \left(\bar{\xi}_t - \underline{\xi}_t \right) |\sigma_t| \right] \end{aligned}$$

So,

$$\begin{aligned} & \alpha_t \min_{\xi \in \Xi} (\sigma_t \xi_t) + (1 - \alpha_t) \max_{\xi \in \Xi} (\sigma_t \xi_t) \\ =\ & \alpha_t \frac{1}{2} \left[\left(\underline{\xi}_t + \bar{\xi}_t \right) \sigma_t + \left(\underline{\xi}_t - \bar{\xi}_t \right) |\sigma_t| \right] + (1 - \alpha_t) \frac{1}{2} \left[\left(\underline{\xi}_t + \bar{\xi}_t \right) \sigma_t + \left(\bar{\xi}_t - \underline{\xi}_t \right) |\sigma_t| \right] \\ =\ & \frac{1}{2} \left(\underline{\xi}_t + \bar{\xi}_t \right) \sigma_t + \alpha_t \frac{1}{2} \left(\underline{\xi}_t - \bar{\xi}_t \right) |\sigma_t| + (1 - \alpha_t) \frac{1}{2} \left(\bar{\xi}_t - \underline{\xi}_t \right) |\sigma_t| \\ =\ & \frac{1}{2} \left(\underline{\xi}_t + \bar{\xi}_t \right) \sigma_t + \frac{|\sigma_t|}{2} \left(\alpha_t \left(\underline{\xi}_t - \bar{\xi}_t \right) + (1 - \alpha_t) \left(\bar{\xi}_t - \underline{\xi}_t \right) \right) \\ =\ & \frac{1}{2} \left(\underline{\xi}_t + \bar{\xi}_t \right) \sigma_t + \frac{|\sigma_t|}{2} \left(\underline{\xi}_t - \bar{\xi}_t \right) (2\alpha_t - 1) \end{aligned}$$

We conclude that

$$\begin{aligned}
\mathcal{M}_t\left(V_{t+h}\right) &= \alpha_t \min_{\xi\in\Xi} \mathrm{E}_t^{\xi}\left[V_{t+h}\right] + \left(1-\alpha_t\right) \max_{\xi\in\Xi} \mathrm{E}_t^{\xi}\left[V_{t+h}\right] \\
&= V_t + \mu_t h + \left(\alpha_t \min_{\xi\in\Xi} \sigma_t\xi_t + \left(1-\alpha_t\right) \max_{\xi\in\Xi} \sigma_t\xi_t\right) h + o\left(h\right) \\
&= V_t + \mu_t h + \left(\frac{1}{2}\left(\underline{\xi}_t + \bar{\xi}_t\right)\sigma_t + \frac{\left|\sigma_t\right|}{2}\left(\underline{\xi}_t - \bar{\xi}_t\right)\left(2\alpha_t - 1\right)\right) h + o\left(h\right)
\end{aligned}$$

This proves (42).

Eqs. (47) and (50) Recall from (41) that $\mathrm{E}_t^{\xi}\left[\mathrm{d}V_t\left(\mathbf{c}\right)\right] = \left(\mu_t + \xi_t\sigma_t\right) h + o\left(h\right)$. So,

$$\begin{aligned}
\mathcal{M}_t\left(\mathrm{d}V_t, h\right) &= \sum_{\xi} \mathrm{E}_t^{\xi}\left[\mathrm{d}V_t\right]\pi_t\left(\xi\right) - \frac{\upsilon}{2h}\left(\sum_{\xi} \mathrm{E}_t^{\xi}\left[\mathrm{d}V_t\right]^2\pi_t\left(\xi\right) - \left(\sum_{\xi} \mathrm{E}_t^{\xi}\left[\mathrm{d}V_t\right]\pi_t\left(\xi\right)\right)^2\right) \\
&= \mu_t h - \frac{\upsilon}{2h}\left(\sum_{\xi}\left(\xi_t\sigma_t h\right)^2\pi_t\left(\xi\right) - \left(\sigma_t h\sum_{\xi}\xi_t\pi_t\left(\xi\right)\right)^2\right) + o\left(h\right) \\
&= \mu_t h - \frac{\upsilon}{2}\sum_{\xi}\xi_t^2\sigma_t^2 h\pi_t\left(\xi\right) = \mu_t h - \frac{\upsilon}{2}\sigma_t^2 h\sum_{\xi}\xi_t^2\pi_t\left(\xi\right) + o\left(h\right) \\
&= \left[\mu_t - \frac{\upsilon\sigma_t^2}{2}\varsigma_t^2\right] h + o\left(h\right)
\end{aligned}$$

This proves (47). Moreover, we have

$$\begin{aligned}
\mathcal{M}_t\left(\mathrm{d}V_t, h\right) &= \phi_h^{-1}\left(\sum_{\xi}\phi_h\left(\mathrm{E}_t^{\xi}\left[\mathrm{d}V_t\right]\right)\pi_t\left(\xi\right)\right) = -\kappa h\log\sum_{\xi} e^{-\frac{\left(\mu_t+\xi_t\sigma_t\right)h+o(h)}{\kappa h}}\pi_t(\xi_t) \\
&= -\kappa h\log\sum_{\xi} e^{-\frac{\mu_t+\xi_t\sigma_t}{\kappa}+\frac{o(h)}{h}}\pi_t(\xi_t) = -\kappa h\log\sum_{\xi} e^{-\frac{\sigma_t}{\kappa}\xi_t}\pi_t(\xi_t) + h\mu_t + o\left(h\right)
\end{aligned}$$

This proves (50) with $\upsilon = 1/\kappa$.

Finally, observe that, as the process ς is bounded, by the Novikov condition the probability $\mathbb{Q}^{-\frac{\upsilon}{2}\varsigma_t^2} = \mathbb{P}^{\mathcal{E}\left(-\frac{\upsilon}{2}\varsigma_t^2\right)}$ is well defined and we have $\mathrm{E}_t^{-\frac{\upsilon}{2}\varsigma_t^2}\left[\mathrm{d}V_t\left(\mathbf{c}\right)\right] = \left(\mu_t - \left(\upsilon/2\right)\varsigma_t^2\sigma_t\right) h + o\left(h\right)$. In turn, this implies the master equation (47).

Section 5.3.3 Let $w > v = 0$ and ϕ be a strictly increasing and continuous function such that $\phi\left(0\right) = 0$. Define $\psi_w : \mathbb{R} \to \mathbb{R}$ by $\psi_w\left(t\right) = \phi\left(wt\right)$. The function ψ_w is strictly increasing, with inverse

$$\psi_w^{-1} = \frac{1}{w}\phi^{-1}$$

Fix any $m \in (0,1)$. We have

$$\begin{aligned}
\phi\left(qw + \left(1-q\right)v\right) = m\phi\left(w\right) + \left(1-m\right)\phi\left(v\right) \quad &\Longleftrightarrow \quad qw = \phi^{-1}\left(m\phi\left(w\right) + \left(1-m\right)\phi\left(v\right)\right) \\
&\Longleftrightarrow \quad q = \frac{1}{w}\phi^{-1}\left(m\phi\left(w\cdot 1\right) + \left(1-m\right)\phi\left(w\cdot 0\right)\right) \\
&\Longleftrightarrow \quad q = \psi_w^{-1}\left(m\psi_w\left(1\right) + \left(1-m\right)\psi_w\left(0\right)\right)
\end{aligned}$$

Define $\bar{\phi} : (0,1) \to \mathbb{R}$ by $\bar{\phi}(m) = \psi_w^{-1}(m\psi_w(1) + (1-m)\psi_w(0))$, so that $q = \bar{\phi}(m)$. With this, q is the quasi-arithmetic mean under ψ_w of the scalars 1 and 0 with weight m. In particular:

(i) $\bar{\phi}(m) \in (0,1)$;

(ii) if ϕ is (strictly) concave, so is ψ_w; thus, $\bar{\phi}(m)$ is (strictly) smaller than the arithmetic mean of the scalars 1 and 0 with weight m, that is,

$$\bar{\phi}(m)\,(<) \leq m1 + (1-m)\,0 = m \tag{84}$$

with equality holding when ψ_w is affine (i.e., when ϕ is affine).

In particular, (i) ensures that the matching probabilities q and $\bar{q}$ are, indeed, probabilities and so belong to $[0,1]$.

For each $\lambda > 0$, set $\phi_\lambda(z) = \lambda(1 - e^{-\frac{z}{\lambda}})$. We have:

$$\bar{\phi}_\lambda(m) = \frac{1}{w}\phi_\lambda^{-1}(m\phi_\lambda(w\cdot 1) + (1-m)\phi_\lambda(w\cdot 0)) = -\frac{\lambda}{w}\log\left(me^{-\frac{w}{\lambda}} + 1 - m\right) = -\frac{\lambda}{w}\log\left(1 - m\left(1 - e^{-\frac{w}{\lambda}}\right)\right)$$

Since ϕ_λ is strictly concave, by (84) we have $\bar{\phi}_\lambda(m) < m$. Moreover,

$$\begin{aligned}
\frac{\mathrm{d}\bar{\phi}_\lambda(m)}{\mathrm{d}\lambda} &= -\frac{1}{w}\left[\log\left(1 - m(1 - e^{-\frac{w}{\lambda}})\right) + \lambda\frac{-m\frac{1}{\lambda^2}e^{-\frac{w}{\lambda}}}{1 - m(1 - e^{-\frac{w}{\lambda}})}\right] \\
&= -\frac{1}{w}\log\underbrace{\left(1 - m(1 - e^{-\frac{w}{\lambda}})\right)}_{\in(0,1)} + \frac{\lambda}{w}\frac{m\frac{1}{\lambda^2}e^{-\frac{w}{\lambda}}}{\underbrace{1 - m(1 - e^{-\frac{w}{\lambda}})}_{>0}} > 0
\end{aligned}$$

So, $\bar{\phi}_\lambda(m)$ strictly increases in λ. Moreover,

$$\lim_{\lambda\to 0}\bar{\phi}_\lambda(m) = 0 \quad \text{and} \quad \lim_{\lambda\to+\infty}\bar{\phi}_\lambda(m) = m$$

Indeed,

$$\begin{aligned}
\lim_{\lambda\to+\infty}\bar{\phi}_\lambda(m) &= \lim_{\lambda\to+\infty} -\frac{\lambda}{w}\log\left(1 - m(1 - e^{-\frac{w}{\lambda}})\right) = \lim_{\lambda\to+\infty}\frac{\log\left(1 - m(1 - e^{-\frac{w}{\lambda}})\right)}{-\frac{w}{\lambda}} \\
&= \lim_{\lambda\to+\infty}\frac{\log\left(1 - \frac{1}{\lambda}m\left(w + o\left(\frac{1}{\lambda}\right)\right)\right)}{-\frac{w}{\lambda}} = \lim_{\lambda\to+\infty}\frac{\log\left(1 - m\frac{w}{\lambda} + o\left(\frac{1}{\lambda^2}\right)\right)}{-\frac{w}{\lambda}} = m
\end{aligned}$$

All this is easily seen to imply (54), (55) and (56).

References

[1] P. Beissner, Q. Lin and F Riedel, Dynamically consistent alpha-maxmin expected utility, mimeo, 2018.

[2] P. Beissner, F. Macheroni, M. Marinacci and S. Mukerji, Decision problems with dynamic robust utility, mimeo, 2021.

[3] P. Billingsley, *Probability and measure*, 3rd ed., Wiley, New York, 1995.

[4] P. Briand and Y. Hu, BSDE with quadratic growth and unbounded terminal value, *Probability Theory and Related Fields*, 136.4, 604-618, 2006

[5] P. Briand and Y. Hu, Quadratic BSDEs with convex generators and unbounded terminal conditions, *Probability Theory and Related Fields*, 141, 543-567, 2008.

[6] P. Briand and R. Elie, A simple constructive approach to quadratic BSDEs with or without delay, *Stochastic processes and their applications*, 123, 2921-2939, 2013.

[7] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci and L. Montrucchio, Ambiguity and robust statistics, *Journal of Economic Theory*, 148, 974-1049, 2013.

[8] Z. Chen and L. Epstein, Ambiguity, risk, and asset returns in continuous time, *Econometrica*, 70, 1403-1443, 2002.

[9] B. de Finetti, Sulla preferibilità, *Giornale degli Economisti*, 11, 685-709, 1952 (trans. in *Giornale degli Economisti,* 71, 257-279, 2012).

[10] D. Duffie and L. G. Epstein, Stochastic differential utility, *Econometrica*, 60, 353-394, 1992.

[11] E. B. Dynkin, *Diffusions, superdiffusions and partial differential equations*, AMS, Providence, 2002.

[12] N. El Karoui, S. Peng and M. C. Quenez, Backward stochastic differential equations in finance, *Mathematical Finance*, 7, 1-71, 1997.

[13] L. G. Epstein and S. E. Zin, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework, *Econometrica*, 57, 937-969, 1989.

[14] W. H. Fleming and R. W. Rishel, *Deterministic and stochastic optimal control*, Springer, New York, 1975.

[15] P. Ghirardato and M. Marinacci, Ambiguity made precise: a comparative foundation, *Journal of Economic Theory*, 102, 251-289, 2002.

[16] P. Ghirardato, F. Maccheroni and M. Marinacci, Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory*, 118, 133-173, 2004.

[17] I. Gihman and A. V. Skorohod, *The theory of stochastic processes III*, Springer, New York, 1979.

[18] I. V. Girsanov, On transforming a certain class of stochastic processes by absolutely continuous substitution of measures, *Theory of Probability and Its Applications*, 5, 285-301, 1960.

[19] I. Gilboa and D. Schmeidler, Maxmin expected utility with a non-unique prior, *Journal of Mathematical Economics*, 18, 141-153, 1989.

[20] L. P. Hansen and M. Marinacci, Ambiguity aversion and model misspecification: an economic perspective, *Statistical Science*, 31, 511-515, 2016.

[21] L. P. Hansen and J. Miao, Aversion to ambiguity and model misspecification in dynamic stochastic environments, *Proceedings of National Academy of Sciences*, 115, 9163-9168, 2018.

[22] L. P. Hansen and J. Miao, Asset pricing under smooth ambiguity in continuous time, mimeo, 2022.

[23] L. P. Hansen and T. J. Sargent, Robust control and model uncertainty, *American Economic Review*, 91, 60-66, 2001.

[24] L. P. Hansen and T. J. Sargent, *Robustness*, Princeton University Press, Princeton, 2008.

[25] L. P. Hansen and T. J. Sargent, Fragile beliefs and the price of uncertainty, *Quantitative Economics,* 1, 129-162, 2010.

[26] L. P. Hansen and T. J. Sargent, Robustness and ambiguity in continuous time, *Journal of Economic Theory*, 146, 1195-1223, 2011.

[27] L. P. Hansen and T. J. Sargent, Structured uncertainty and model misspecification, mimeo, 2018.

[28] K. Jurado, S. C. Ludvigson, and S. Ng , *Measuring uncertainty*, *American Economic Review*, 105.3, 1177-1216, 2015.

[29] I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus*, 2^{nd} ed., Springer, New York, 1991.

[30] P. Klibanoff, M. Marinacci and S. Mukerji, A smooth model of decision making under ambiguity, *Econometrica*, 73, 1849-1892, 2005.

[31] P. Klibanoff, M. Marinacci and S. Mukerji, Recursive smooth ambiguity preferences, *Journal of Economic Theory*, 144, 930-976, 2009.

[32] M. Kobylanski, Backward stochastic differential equations and partial differential equations with quadratic growth, *Annals of Probability*, 28.2, 558-602, 2000.

[33] T. C. Koopmans, Stationary utility and impatience, *Econometrica*, 28, 287-309, 1960.

[34] R. J. A. Laeven and M. Stadje, Robust portfolio choice and indifference valuation, *Mathematics of Operations Research*, 39, 1109-1141, 2014.

[35] A. Lazrak and M. C. Quenez, A generalized stochastic differential utility, *Mathematics of Operations Research*, 28, 154-180, 2003.

[36] J.-F. Le Gall, *Brownian motion, martingales, and stochastic calculus*, Springer, New York, 2016.

[37] J-P. Lepeltier and J. San Martín, Existence for BSDE with superlinear-quadratic coefficient, *Stochastics: An International Journal of Probability and Stochastic Processes*, 63, 227-240, 1998.

[38] M. Marinacci and L. Montrucchio, Unique solutions for stochastic recursive utilities, *Journal of Economic Theory,* 145.5, 1776-1804, 2010.

[39] F. Maccheroni, M. Marinacci and D. Ruffino, Alpha as ambiguity: robust mean-variance portfolio analysis, *Econometrica,* 81, 1075-1113, 2013.

[40] M. Marinacci, Model uncertainty, *Journal of the European Economic Association*, 13, 998-1076, 2015.

[41] H. McKean, *Stochastic integrals*, AMS, Providence, 1969.

[42] M. Nisio, *Stochastic control theory: dynamic programming principle*, Springer, New York, 2014.

[43] E. Pardoux and S. Peng, Adapted solution of a backward stochastic differential equation, *Systems and Control Letters*, 14, 55-61, 1990.

[44] L.J. Savage, *The Foundations of Statistics*, Wiley, New York, 1954.

[45] M. Schroder and C. Skiadas, Lifetime consumption-portfolio choice under trading constraints, recursive preferences, and nontradeable income, *Stochastic Processes and their Applications*, 115, 1-30, 2005.

[46] R. L. Schilling, *Measures, integrals and martingales*, 2nd ed. Cambridge University Press, Cambridge, 2017.

[47] M. Siniscalchi, Vector expected utility and attitudes toward variation, *Econometrica*, 77.3 , 801-855, 2009.

[48] C. Skiadas, Smooth ambiguity aversion toward small risks and continuous-time recursive utility, *Journal of Political Economy*, 121, 775-792, 2013.

[49] R. Tevzadze, Solvability of backward stochastic differential equations with quadratic growth, *Stochastic Processes and their Applications*, 118, 503-515, 2008.

[50] N. Touzi, *Optimal stochastic control, stochastic target problems, and backward SDE*, Springer, New York, 2013.