Economic Theory (2022) 74:335–371
https://doi.org/10.1007/s00199-022-01441-5

RESEARCH ARTICLE

Asset pricing under smooth ambiguity in continuous time

Lars Peter Hansen[1] · **Jianjun Miao**[2]

Received: 22 February 2022 / Accepted: 11 June 2022 / Published online: 30 June 2022

Abstract
We study asset pricing implications of a revealing and tractable formulation of smooth ambiguity investor preferences in a continuous-time environment. Investors do not observe a hidden Markov state and instead make inferences about this state using past data. We show that ambiguity about this hidden state distribution alters investor decisions and equilibrium asset prices. Our continuous-time formulation allows us to apply recursive filtering and Hamilton–Jacobi–Bellman methods to solve the modified decision problem. Using such methods, we show how characterizations of portfolio allocations and local uncertainty-return tradeoffs change when investors are ambiguity-averse.

Keywords Risk · Ambiguity · Robustness · Asset pricing · Portfolio allocation · Continuous time

JEL Classification D81 · G11 · G12

1 Introduction

We consider control and pricing problems with a continuous-time specification when decision makers have preferences that include what is referred to as *smooth ambiguity*. We introduce hidden Markov states that are the source of the intertemporal ambiguity concerns on the part of the decision-maker. We show how these concerns alter portfolio allocations of investors and equilibrium asset prices. While there are

Forthcoming in *Economic Theory*. The GitHub link: https://github.com/lphansen/PortfolioChoice provides details about the computations and code access for the example in Sect. 5.

✉ Lars Peter Hansen
lhansen@uchicago.edu

Jianjun Miao
miaoj@bu.edu

1 University of Chicago, Chicago, USA

2 Boston University, Boston, USA

several applications of the smooth ambiguity approach to decision-making in discrete time, to implement this approach in continuous-time requires rethinking how the ambiguity aversion responds to the more frequent arrival of information. We produce a tractable formulation with a well-posed continuous-time limiting representation extending the discrete-time recursive formulation of Klibanoff et al. (2009), Hayashi and Miao (2011), and Ju and Miao (2012). This extension opens the door to our investigation of the impact of ambiguity aversion on continuous-time portfolio choice and asset pricing.

This paper is a companion to Hansen and Miao (2018). Our previous paper showed how to construct decision-maker preferences that are recursive and express ambiguity aversion and/or misspecification concerns for continuous-time Brownian motion information structures with hidden states or parameters. That paper showed a novel way to embed ambiguity aversion in a discrete-time limit. The focus was on the conceptual framework with little discussion about implementations for economic applications. The aim of this paper is to show how so-called smooth ambiguity preferences, potentially motivated by robustness concerns, alter portfolio construction and equilibrium asset prices.

Our continuous-time formulation of ambiguity aversion has an equivalent formulation as a robust prior/posterior adjustment to the conditional distribution over the hidden Markov state. This extends an insight from discrete-time given in Hansen and Sargent (2007) to continuous-time hidden Markov dynamics. The outcome of the robustness adjustment is an altered probability distribution over the hidden state conditioned on current information that shifts weights towards states that are more concern to the decision-maker. As we discuss and illustrate, this equivalence has three valuable implications: (i) it offers a tractable way to solve dynamic decision problems formulated as max–min problems associated with Hamilton–Jacobi–Bellman (HJB) equations; (ii) it opens the door to assessing or calibrating ambiguity aversion parameters in different ways; and (iii) the altered probability distribution helps to reveal the consequences of the ambiguity aversion in decision and provides a revealing characterization of the equilibrium valuation consequences. While these observations are also central to the Barnett et al. (2020) analysis of the climate change uncertainty for economic dynamics, their paper abstracts from learning.

In this paper we apply the continuous-time ambiguity aversion model of Hansen and Miao (2018) to:

- characterize the ambiguity-averse contribution to the local market price of uncertainty in addition to the familiar risk aversion component;
- show how increases in ambiguity aversion alter the forward-looking hedging demand as well as the static demand for a portfolio of risky and risk-less securities, potentially in opposite directions.

Our paper has been influenced by some important prior contributions of Larry Epstein. First, we use the recursive preferences derived by Epstein and Zin (1989) in discrete time and by Duffie and Epstein (1992) in continuous time designed to include separate preference contributions for risk aversion and intertemporal substitution. Our inclusion of ambiguity aversion can be viewed as a "smooth" counterpart to the (Chen and Epstein 2002) recursive version max–min utility theory in continuous time.

This paper is organized as follows. In Sect. 2, we provide the underlying probabilistic structure assumed in this paper along with two special cases that have been used extensively in applications: Kalman–Bucy filtering and Wonham filtering. We present and discuss our continuous-time representation of the utility recursion in Sect. 3. In Sects. 4 and 5, we combine results from these two initial sections to study equilibrium pricing and a portfolio allocation. In both cases, we identify an explicit impact of ambiguity aversion. Section 6 introduces jump risk into the smooth ambiguity utility model in continuous time. Finally, we provide some concluding remarks in Sect. 7.

2 Stochastic setting

Let $(\Omega, \mathfrak{A}, \mathcal{P})$ be a probability space and time is continuous on $[0, \infty)$. Let $(Y, Z) = \{(Y_t, Z_t) : t \geq 0\}$ be a multi-dimensional partially observed stochastic process, where Z is an unobservable component and Y is an observable component. Let $\mathfrak{F} = \{\mathfrak{F}_t : t \geq 0\}$ be the filtration generated by (Y, Z) and $\mathfrak{G} = \{\mathfrak{G}_t : t \geq 0\}$ the filtration generated by Y alone. The problem of filtering is to construct and update expectations conditioned on $\mathfrak{G}_t$ for $t \geq 0$.

Suppose that Y satisfies

$$Y_t = y_0 + \int_0^t \mu_s^Y (Y_s, Z_s)\, ds + \int_0^t \sigma_s^Y (Y_s)\, dW_s,$$

where $W = \{W_t : t \geq 0\}$ is a multi-dimensional standard Brownian motion defined on $(\Omega, \mathfrak{F}, \mathcal{P})$ and y_0 is an initial condition. The process $\{\sigma_t (Y_t) : t \geq 0\}$ is adapted to the coarser $\mathfrak{G}$ filtration. For simplicity, assume that the diffusion matrix $\sigma_t^Y \left(\sigma_t^Y\right)'$ is nonsingular for all t almost surely.

We follow chapter 8 of Liptser and Shiryaev (2001a) by presenting a general filtering result in continuous time. Let $\bar{\mu}_t^Y = \mathbb{E}\left[\mu_t^Y (Y_t, Z_t) \,|\mathfrak{G}_t\right]$. Liptser and Shiryaev show that under certain technical conditions, the process, $\overline{W} = \left\{\overline{W}_t : t \geq 0\right\}$, defined by

$$\begin{aligned} d\overline{W}_t &\doteq \left[\sigma_t^Y \left(\sigma_t^Y\right)'\right]^{-1/2} \left(dY_t - \bar{\mu}_t^Y dt\right) \\ &= \left[\sigma_t^Y \left(\sigma_t^Y\right)'\right]^{-1/2} \left[\left(\mu_t^Y - \bar{\mu}_t^Y\right) dt + \sigma_t^Y dW_t\right], \end{aligned} \tag{1}$$

is a standard Brownian motion relative to the observation information filtration $\mathfrak{G}$. The Brownian motion $\overline{W}$, called an *innovation process*, plays a key role in filtering theory because it generates $\mathfrak{G}$ and has the same dimension as Y.

We consider two examples of filters in the next two subsections.

2.1 Kalman–Bucy filter

Suppose that the partially observed system is given by

$$\begin{aligned} dY_t &= \left(A_{y0} + A_{yy}Y_t + A_{yz}Z_t\right)dt + B_y dW_t, \\ dZ_t &= \left(A_{z0} + A_{zy}Y_t + A_{zz}Z_t\right)dt + B_z dW_t. \end{aligned} \tag{2}$$

Suppose that all matrices in the above equations are constant and appropriately conformable. Moreover, suppose that $B_y B_y'$ is nonsingular and that Z_0 is normally distributed with mean $\bar{z}_0$ and variance Σ_0. We treat the case in which $B_y B_z{}' = 0$. Otherwise, we can transform the state to remove the conditional correlation by studying the filtering problem:

$$\widetilde{Z}_t = Z_t - B_z B_y{}' \left(B_y B_y'\right)^{-1} Y_t.$$

The state $\widetilde{Z}_t$ has the same conditional covariance matrix as Z_t and a conditional mean that is translated by

$$A_{zz} B_z B_y{}' \left(B_y B_y'\right)^{-1} Y_t.$$

As noted by section 10.3 of Liptser and Shiryaev (2001a), under the limited information filtration $\mathfrak{G}$, $\overline{Z}_t \doteq \mathbb{E}\left(Z_t \mid \mathfrak{G}_t\right)$ is a Gaussian process with mean satisfying the Kalman–Bucy filter:

$$\begin{aligned} dY_t &= \left(A_{y0} + A_{yy}Y_t + A_{yz}\overline{Z}_t\right)dt + \left(B_y B_y'\right)^{1/2} d\overline{W}_t, \\ d\overline{Z}_t &= \left(A_{z0} + A_{zy}Y_t + A_{zz}\overline{Z}_t\right)dt + \Sigma_t A_{yz}' \left(B_y B_y'\right)^{-1/2} d\overline{W}_t, \end{aligned} \tag{3}$$

where conditional covariance matrix Σ_t solves:

$$\frac{d\Sigma_t}{dt} = A_{zz}\Sigma_t + \Sigma_t A_{zz}' + B_z B_z' - \Sigma_t A_{yz}' \left(B_y B_y'\right)^{-1} A_{yz}\Sigma_t, \tag{4}$$

and the innovation increment is given by

$$d\overline{W}_t = \left(B_y B_y'\right)^{-1/2} \left[A_{yz}\left(Z_t - \overline{Z}_t\right)dt + B_y dW_t\right]. \tag{5}$$

The Brownian increment $d\overline{W}_t$ under the $\mathfrak{G}$ filtration has two contributions: the prediction error, $Z_t - \overline{Z}_t$ and the Brownian increment, dW_t, under the original $\mathfrak{F}$ filtration. While (2) gives the Markov state dynamics for the $\mathfrak{F}$ filtration, (3) and (4) give an alternative first-order Markov state dynamics pertinent for the $\mathfrak{G}$ filtration where the date t state vector is taken to be $(Y_t, \overline{Z}_t, \Sigma_t)$.

Example 2.1 *(Parameter estimation)* Suppose $A_{z0} = 0$, $A_{zy} = 0$, and $A_{zz} = 0$ and $B_z = 0$. With these restrictions the Kalman–Bucy filtering model specializes to one with parameter estimation with Z_t being time invariant and $A_{yz}Z_t$ being an unknown parameter vector for the evolution of Y. In this special case, the posterior mean evolves as

$$d\overline{Z}_t = \Sigma_t A'_{yz}\left(B_y B'_y\right)^{-1/2} d\overline{W}_t,$$

which is a martingale under the $\mathfrak{G}$ filtration, and the posterior covariance matrix evolves as:

$$\frac{d\Sigma_t}{dt} = -\Sigma_t A'_{yz}\left(B_y B'_y\right)^{-1} A_{yz}\Sigma_t.$$

2.2 Wonham filter

Suppose that the decision-maker observes information generated by a multi-dimensional regime-switching process Y, that evolves as:

$$dY_t = A_{yy}Y_t dt + A_{yz}Z_t dt + B_y dW_t, \tag{6}$$

where W is a Brownian motion with respect to the filtration $\mathfrak{F}$ and $B_y B'_y$ is nonsingular. The process Z evolves as an n-state Markov chain with a discrete state space given by the coordinate vector of $\mathbb{R}^n$. The intensity matrix for the hidden state process Z is $\Lambda = \left[\lambda_{ij}\right]$ where $\lambda_{ij} \geq 0$ for $j \neq i$ and $\lambda_{ii} = -\sum_{j\neq i}\lambda_{ij}$.[1]

Notice that $\overline{Z}_t = \mathbb{E}\left(Z_t \mid \mathfrak{G}_t\right)$ is the vector of predicted state probabilities. By Wonham (1965) or chapter 9 of Liptser and Shiryaev (2001a), we have that

$$dY_t = A_{yy}Y_t dt + A_{yz}\overline{Z}_t dt + \left(B_y B'_y\right)^{1/2} d\overline{W}_t,$$

where

$$d\overline{W}_t = \left(B_y B'_y\right)^{-1/2}\left[A_{yz}\left(Z_t - \overline{Z}_t\right)dt + B_y dW_t\right]. \tag{7}$$

The dynamics for the predicted state are modified to be:

$$d\overline{Z}_t = \Lambda'\overline{Z}_t dt + \left[\operatorname{diag}\left(\overline{Z}_t\right) - \overline{Z}_t\overline{Z}'_t\right] A'_{yz}\left(B_y B'_y\right)^{-1/2} d\overline{W}_t,$$

where $\operatorname{diag}\left(\overline{Z}_t\right)$ is a diagonal matrix with the entries of $\overline{Z}_t$ on the diagonal.

Example 2.2 *(Parameter estimation)* When the intensity matrix is identically zero, the hidden state is invariant. This limiting case is of particular interest and can be viewed

[1] While we could replace B_y by $\sigma_t^Y(Y_t)$, we suppress such dependence for notational simplicity.

as parameter estimation with $A_{yz}Z_t$ as being the unknown parameter vector. The posterior distribution evolves as a martingale

$$d\overline{Z}_t = \left[\operatorname{diag}\left(\overline{Z}_t\right) - \overline{Z}_t\overline{Z}_t'\right] A_{yz}'\left(B_y B_y'\right)^{-1/2} d\overline{W}_t$$

under the $\mathfrak{G}$ filtration. Since Z_t is one of the coordinate vectors, the columns of A_{yz} give the n possible realizations of the parameter vector.

3 Smooth ambiguity preferences

We represent intertemporal preferences through the use of continuation values. Denote the decision-maker's continuation utility process $V = \{V_t : t \geq 0\}$ for a $\mathfrak{G}$-adapted consumption process $C = \{C_t : t \geq 0\}$. The process V solves a backward stochastic differential equation (BSDE) implied by the preferences. Such a differential equation has a terminal condition imposed at some date T. In our characterization, we will effectively take limits as T tends to infinity to sidestep horizon dependence in the solution.

3.1 Restricting the local evolution of *V*

Write the evolution in terms of the $\mathfrak{G}$ filtration as

$$dV_t = V_t\bar{\mu}_t^V dt + V_t\bar{\sigma}_t^V d\overline{W}_t,$$

where both $\bar{\mu}_t^V$ and $\bar{\sigma}_t^V$ are adapted to the $\mathfrak{G}$ filtration. Then we may also use (1) to write an $\mathfrak{F}$ filtration counterpart as

$$dV_t = V_t\mu_t^V dt + V_t\sigma_t^V dW_t,$$

where

$$\mu_t^V - \bar{\mu}_t^V = \bar{\sigma}_t^V \left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\left(\mu_t^Y - \bar{\mu}_t^Y\right), \tag{8}$$

$$\sigma_t^V = \bar{\sigma}_t^V \left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\sigma_t^Y. \tag{9}$$

While σ_t^V given by (9) is $\mathfrak{G}_t$ measurable, μ_t^V given by (8) is only restricted to be $\mathfrak{F}_t$ measurable. Taking expectations on the two sides of (8) conditioned on $\mathfrak{G}_t$, we have $\bar{\mu}_t^V = \mathbb{E}\left(\mu_t^V \mid \mathfrak{G}_t\right)$. Importantly, armed with a solution to the filtering problem, $(\bar{\mu}_t^V, \bar{\sigma}_t^V)$ determines (μ_t^V, σ_t^V) . In order to produce a formula for $\bar{\mu}_t^V$, the local $\mathfrak{F}$ evolution will come into play as dW_t contributes "risk" and $\mu_t^V - \bar{\mu}_t^V$ contributes ambiguity induced by the hidden state.

The smooth ambiguity preferences in continuous time impose a restriction on $(\bar{\mu}_t^V, \bar{\sigma}_t^V)$ expressed in terms of the pair (μ_t^V, σ_t^V). Hansen and Miao (2018) derived this equation by taking limits of a discrete-time model related to Hansen and Sargent (2007), Hansen and Sargent (2011), Klibanoff et al. (2009), Hayashi and Miao (2011), and Ju and Miao (2012):

$$V_t = \left\{\left[1 - \exp(-\delta\epsilon)\right](C_t)^{1-\rho} + \exp(-\delta\epsilon)(\mathcal{A}_t)^{1-\rho}\right\}^{\frac{1}{1-\rho}}, \tag{10}$$

where $\delta > 0$ is the subjective discount rate, $\frac{1}{\rho} > 0$ is an intertemporal elasticity of substitution, and $\epsilon > 0$ denotes a time increment. The two conditional certainty equivalent operators are defined as

$$\mathcal{R}_t = \left\{\mathbb{E}\left[(V_{t+\epsilon})^{1-\gamma} | \mathfrak{F}_t\right]\right\}^{\frac{1}{1-\gamma}}, \gamma > 0, \tag{11}$$

$$\mathcal{A}_t = \left\{\mathbb{E}\left[(\mathcal{R}_t)^{-\frac{\alpha}{\epsilon}} | \mathfrak{G}_t\right]\right\}^{-\frac{\epsilon}{\alpha}}, \alpha > 0, \tag{12}$$

where the parameters γ and α capture risk aversion and ambiguity aversion, respectively. The contribution of the certainty equivalent $\mathcal{R}_t$ is familiar from the work of Kreps and Porteus (1978) and Epstein and Zin (1989). The certainty equivalent adjustment $\mathcal{A}_t$ is familiar from the work of Klibanoff et al. (2009) except for the impact of ϵ. The ϵ contribution for ambiguity aversion is motivated explicitly in Hansen and Sargent (2011) and Hansen and Miao (2018).[2] Conveniently, the composite aggregator (10) is homogeneous in degree one in $(C_t, V_{t+\epsilon})$. Hansen and Miao (2018) take limits of this composite aggregator and derive a continuous-time counterpart that we use in this paper.

Initially, replace the ambiguity certainty equivalent given in (12) by assuming that

$$\mathcal{A}_t = \mathcal{R}_t = \left\{\mathbb{E}\left[(V_{t+\epsilon})^{1-\gamma} | \mathfrak{F}_t\right]\right\}^{\frac{1}{1-\gamma}}.$$

Consider a solution in terms of BSDE expressed in terms of the filtration $\mathfrak{F}$. Consistent with Duffie and Epstein (1992), it follows that by taking an $\epsilon \downarrow 0$ limit

$$0 = \frac{\delta}{1-\rho}\left[\left(\frac{C_t}{V_t}\right)^{1-\rho} - 1\right] + \mu_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2. \tag{13}$$

Notice that this equation entails both the drift contribution μ_t^V and the diffusion contribution $\left|\sigma_t^V\right|^2$. The latter is scaled by the risk aversion parameter γ. An important special case is when intertemporal elasticity of substitution is unity ($\rho = 1$), in which case the limiting utility recursion is

[2] Skiadas (2013) adopts a different scaling using α in place of $\frac{\alpha}{\epsilon}$ and shows that smooth ambiguity adjustment vanishes in the limit. See Hansen and Sargent (2011) and Hansen and Miao (2018) for further discussion.

$$0 = \delta\left(\log C_t - \log V_t\right) + \mu_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2.$$

Under the partial information filtration $\mathfrak{G}$, the BSDE for V still satisfies (13) but with μ_t^V replaced by $\bar{\mu}_t^V$.[3]

Consider next the change when induced by the smooth adjustment for ambiguity. Then the drift contribution μ_t^V in (13) is replaced by the ambiguity adjustment derived in equation [11] of Hansen and Miao (2018):

$$-\frac{1}{\alpha}\log \mathbb{E}\left[\exp\left(-\alpha\mu_t^V\right) \mid \mathfrak{G}_t\right] \leq \mathbb{E}\left(\mu_t^V \mid \mathfrak{G}_t\right) = \bar{\mu}_t^V, \tag{14}$$

which gives the ambiguity-adjusted modification to equation (13):

$$\begin{aligned} 0 = \frac{\delta}{1-\rho}\left[\left(\frac{C_t}{V_t}\right)^{1-\rho} - 1\right] + \bar{\mu}_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2 \\ -\frac{1}{\alpha}\log \mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right). \end{aligned} \tag{15}$$

The $\rho = 1$ limiting recursion is now

$$\begin{aligned} 0 = \delta\left(\log C_t - \log V_t\right) + \bar{\mu}_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2 \\ -\frac{1}{\alpha}\log \mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right). \end{aligned} \tag{16}$$

Notice from formula (9) that the risk adjustment can be expressed as

$$\frac{\gamma}{2}\left|\sigma_t^V\right|^2 = \frac{\gamma}{2}\bar{\sigma}_t^V\left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\sigma_t^Y\left(\sigma_t^Y\right)'\left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\left(\bar{\sigma}_t^V\right)' = \frac{\gamma}{2}\left|\bar{\sigma}_t^V\right|^2,$$

and from formula (8), the ambiguity adjustment can be expressed as

$$\begin{aligned} &-\frac{1}{\alpha}\log \mathbb{E}\left[\exp\left(-\alpha\mu_t^V\right) \mid \mathfrak{G}_t\right] \\ &= \bar{\mu}_t^V - \frac{1}{\alpha}\log \mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right) \\ &= \bar{\mu}_t^V - \frac{1}{\alpha}\log \mathbb{E}\left(\exp\left[-\alpha\left(\bar{\sigma}_t^V\left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\left(\mu_t^Y - \bar{\mu}_t^Y\right)\right)\right] \mid \mathfrak{G}_t\right). \end{aligned}$$

Thus, we may rewrite formula (15) as

$$0 = \frac{\delta}{1-\rho}\left[\left(\frac{C_t}{V_t}\right)^{1-\rho} - 1\right] + \bar{\mu}_t^V - \frac{\gamma}{2}\left|\bar{\sigma}_t^V\right|^2$$

[3] As we will see in the argument that follows, $\left|\sigma_t^V\right|^2 = \left|\bar{\sigma}_t^V\right|^2$.

$$-\frac{1}{\alpha}\log\mathbb{E}\left(\exp\left[-\alpha\left(\bar{\sigma}_t^V\left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\left(\mu_t^Y-\bar{\mu}_t^Y\right)\right)\right]\mid\mathfrak{G}_t\right). \quad (17)$$

Equation (17) restricts the local evolution $(\bar{\mu}_t^V, \bar{\sigma}_t^V)$ of the continuation value given the current level, V_t along with date t consumption and the local dynamics of the solution to the filtering problem. This equation will give us a direct input our solutions for equilibrium prices and portfolio allocations. Notice that the adjustment in the second line of the equation is recognizable as a "smooth" adjustment for aversion to ambiguity in the drift of the continuation value, in contrast to the max–min recursive utility specification of Chen and Epstein (2002).

3.2 A robustness interpretation

To obtain a more direct generalization of the max–min recursive utility formulation, we follow Hansen and Sargent (2011) and Hansen and Miao (2018) by motivating this adjustment as an outcome of a relative-entropy minimization problem:

$$\begin{aligned}&-\xi\log\mathbb{E}\left[\exp\left[-\frac{\left(\mu_t^V-\bar{\mu}_t^V\right)}{\xi}\right]\mid\mathfrak{G}_t\right]\\&=\min_{M_t\geq 0,\mathbb{E}(M_t\mid\mathfrak{G}_t)=1}\mathbb{E}\left[M_t\left(\mu_t^V-\bar{\mu}_t^V\right)\mid\mathfrak{G}_t\right]+\xi\mathbb{E}\left(M_t\log M_t\mid\mathfrak{G}_t\right)\end{aligned} \quad (18)$$

subject to $\mathbb{E}\left(M_t\mid\mathfrak{G}_t\right)=1$ where $\xi>0$. It suffices to search over random variables that are $\mathfrak{F}_t$ measurable given that the drift μ_t^V satisfies this restriction. The random variable M_t is used as a robustness adjustment for probabilities of $\mathfrak{F}_t$ measurable events conditioned on $\mathfrak{G}_t$. The parameter ξ is a penalty parameter and $\mathbb{E}\left(M_t\log M_t\mid\mathfrak{G}_t\right)$ is a measure of entropy relative to a baseline probability distribution.

The minimizing M_t^* is given by familiar exponential tilting formula:

$$M_t^*=\frac{\exp\left[-\frac{\left(\mu_t^V-\bar{\mu}_t^V\right)}{\xi}\right]}{\mathbb{E}\left(\exp\left[-\frac{\left(\mu_t^V-\bar{\mu}_t^V\right)}{\xi}\right]\mid\mathfrak{G}_t\right)}=\frac{\exp\left(-\frac{\mu_t^V}{\xi}\right)}{\mathbb{E}\left[\exp\left(-\frac{\mu_t^V}{\xi}\right)\mid\mathfrak{G}_t\right]},$$

where probabilities are tilted towards adverse outcomes as quantified by the surprise movement in the drift $\mu_t^V-\bar{\mu}_t^V$. The left side of (18) is equivalent to the smooth ambiguity adjustment where $\xi=\frac{1}{\alpha}$. Thus an alternative way to express (17) is

$$\begin{aligned}0=&\frac{\delta}{1-\rho}\left[\left(\frac{C_t}{V_t}\right)^{1-\rho}-1\right]+\bar{\mu}_t^V-\frac{\gamma}{2}\left|\bar{\sigma}_t^V\right|^2\\&+\min_{M_t\geq 0,\mathbb{E}(M_t\mid\mathfrak{G}_t)=1}\mathbb{E}\left[M_t\left(\mu_t^V-\bar{\mu}_t^V\right)\mid\mathfrak{G}_t\right]+\xi\mathbb{E}\left(M_t\log M_t\mid\mathfrak{G}_t\right)\end{aligned}$$

for $\left(\mu_t^V-\bar{\mu}_t^V\right)$ given by (8).

The minimized objective gives the equivalent smooth ambiguity adjustment where $\xi = \frac{1}{\alpha}$. A large ambiguity aversion parameter α acts as a small penalization parameter when making robust adjustments to learning or filtering problems. Reducing ξ (increasing α) reduces the objective to be minimized and consequently makes the robustness/ambiguity adjustment more pronounced. This formulation can be viewed as a continuous-time counterpart to the recursive formulation of variational preferences in Maccheroni et al. (2006).[4] This robustness interpretation (i) alters how we think of plausible settings for the ambiguity aversion parameter, (ii) provides additional interpretative insights into applications of smooth ambiguity preferences, and (iii) offers alternative tractable ways for computing numerical solutions to decision problems. In terms of (i) and (ii) the robustness interpretation opens the door to assess the plausibility of alternative choices of α following the robust Bayesian perspective of Good (1952) and the statistical detection methods featured by Anderson et al. (2003). In terms of (iii) see Barnett et al. (2020) for a computation approach that implements ambiguity aversion through minimization over probabilities.

3.3 Examples

We now revisit the two filtering examples that introduced in Sect. 2

Example 3.1 *(Kalman–Bucy filtering)* For Kalman–Bucy filtering, $\mu_t^Y - \bar{\mu}_t^Y$ is conditionally normally distributed with mean zero and nonstochastic covariance matrix $A_{yz}\Sigma_t A_{yz}'$ and $\sigma_t^Y = B_y$. As a consequence,

$$\log \mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right) = \frac{\alpha^2}{2}\bar{\sigma}_t^V \left(B_y B_y'\right)^{-1/2} A_{yz}\Sigma_t A_{yz}' \left(B_y B_y'\right)^{-1/2} \left(\bar{\sigma}_t^V\right)',$$

since the left side variable is distributed as a log normal with mean zero and variance $|\bar{\sigma}_t^V|^2$ expressed in logarithms. Thus the smooth ambiguity adjustment is

$$\begin{aligned} &-\frac{1}{\alpha}\log \mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right) \\ &= -\frac{\alpha}{2}\bar{\sigma}_t^V \left(B_y B_y'\right)^{-1/2} A_{yz}\Sigma_t A_{yz}' \left(B_y B_y'\right)^{-1/2} \left(\bar{\sigma}_t^V\right)'. \end{aligned}$$

Combining the risk and smooth ambiguity contributions in (17) gives

$$-\frac{1}{2}\bar{\sigma}_t^V \left[\alpha\left(B_y B_y'\right)^{-1/2} A_{yz}\Sigma_t A_{yz}' \left(B_y B_y'\right)^{-1/2} + \gamma I\right] \left(\bar{\sigma}_t^V\right)',$$

where the first term in $[\cdot]$ is a covariance matrix adjustment for state estimation and the second term is a Brownian risk adjustment, each weighted by distinct aversion

[4] Variational preferences are a generalization of max–min utility, which allows for penalization. Strictly speaking, Maccheroni et al. (2006) presume a form of time separability, which is true in our setting when $\rho = \gamma$.

parameters. The distorted state distribution is normal with a conditional mean given by

$$\overline{Z}_t - \alpha \Sigma_t A'_{yz} \left(B_y B'_y \right)^{-1/2} \left(\sigma_t^V \right)'$$

and an undistorted conditional covariance matrix Σ_t.

In the special case of a scalar observation process Y, the ambiguity adjustment is equivalent to an enhanced risk aversion parameter that could be time varying and given by

$$\alpha \frac{A_{yz} \Sigma_t \left(A_{yz} \right)'}{|B_y|^2} + \gamma.$$

Equivalently, a robust adjustment to the state distribution imitates risk aversion in preference.

Example 3.2 *(Wonham filtering)* For Wonham filtering, the state distribution for Z_t conditioned on $\mathfrak{G}_t$ is not normally distributed since the realizations of Z_t are coordinate vectors. While the log–exp ambiguity adjustment may be computed numerically, the resulting ambiguity adjustment cannot be captured by a simple covariance correction and takes a different form than that for the risk correction.

4 Pricing uncertainty

In this section, we show how to incorporate smooth ambiguity into the local prices of uncertainty expressed as expected returns compensations. Let $S = \{S_t : t \geq 0\}$ be the stochastic discount factor process used to represent the shadow price process implied by a consumer's optimization problem. Write the continuous-time evolution for $(\log C, \log V, \log S)$ as

$$d \log C_t = \left(\bar{\mu}_t^C - \frac{1}{2} |\bar{\sigma}_t^C|^2 \right) dt + \bar{\sigma}_t^C d\overline{W}_t, \quad (19)$$

$$d \log V_t = \left(\bar{\mu}_t^V - \frac{1}{2} |\bar{\sigma}_t^V|^2 \right) dt + \bar{\sigma}_t^V d\overline{W}_t, \quad (20)$$

$$d \log S_t = \left(\bar{\mu}_t^S - \frac{1}{2} |\bar{\sigma}_t^S|^2 \right) dt + \bar{\sigma}_t^S d\overline{W}_t. \quad (21)$$

Our aim is to produce formulas for $(\bar{\mu}_t^S, \bar{\sigma}_t^S)$ taking the consumption dynamics as given, say as the outcome of a dynamic resource allocation problem. For the purposes of the discussion in this section, we may think of the dynamics in (19) as derived from applying a Sect. 2 recursive filtering solution to:

$$d \log C_t = \left(\mu_t^C - \frac{1}{2} |\sigma_t^C|^2 \right) dt + \sigma_t^C dW_t. \quad (22)$$

4.1 Local risk-return tradeoff

Consider a positive cumulative return process Q:

$$dQ_t = Q_t \bar{\mu}_t^Q dt + Q_t \bar{\sigma}_t^Q d\overline{W}_t. \tag{23}$$

This process presumes all cash flows are reinvested, and thus the local return is $\frac{dQ}{Q}$. Frictionless asset pricing theory implies that QS is a martingale under the $\mathfrak{G}$ filtration. Applying Ito's Lemma, the local martingale restriction implies that

$$Q_t S_t \left(\bar{\mu}_t^Q + \bar{\mu}_t^S + \bar{\sigma}_t^S \cdot \bar{\sigma}_t^Q \right) = 0.$$

Thus the instantaneous expected return satisfies

$$\bar{\mu}_t^Q = -\bar{\mu}_t^S - \bar{\sigma}_t^S \cdot \bar{\sigma}_t^Q. \tag{24}$$

From this formula, we see that the instantaneous risk-free rate is $-\bar{\mu}_t^S$ and the local price vector expressed as expected return compensation for exposure to Brownian risk is $-\bar{\sigma}_t^S$. This computation is well known from continuous-time asset pricing theory.

The continuous-time filtering problem provides some extra structure to this pricing formula. Recall that the Brownian increment of relevance to a decision maker is

$$d\overline{W}_t = \left[\sigma_t^Y \left(\sigma_t^Y \right)' \right]^{-1/2} \left[\left(\mu_t^Y - \bar{\mu}_t^Y \right) dt + \sigma_t^Y dW_t \right].$$

Thus exposure to $d\overline{W}_t$ is a bundle of the exposure to the Brownian increment dW_t and to $\left(\mu_t^Y - \bar{\mu}_t^Y \right) dt$. In what follows, we will deduce compensations for each of these two exposures as part of our construction of $\bar{\sigma}_t^S$.

4.2 Recursive risk and ambiguity adjustments

To set the stage for characterizing the contribution of ambiguity aversion to the local (in time) asset prices, we first derive the stochastic discount factor for the discrete-time approximation in the utility recursions (10), (11), and (12). We then take limits as the time increment approaches zero. Our derivation will be admittedly heuristic, but is designed to motivate the extension to ambiguity aversion.[5] This derivation can be viewed as a way to construct an informed guess for the equilibrium asset prices, and the formal verification will require more specificity about the underlying economic environment.

We start by computing two marginal utilities of consumption at distinct but nearby points in time as implied by the utility recursions. One is for the current consumption,

[5] For a more rigorous derivation in continuous time that abstracts from ambiguity aversion see, for instance (Skiadas 2007).

and the other is future consumption from the perspective of the current period for a discrete-time process interval of length ϵ:

$$MC_{t+\epsilon}^{\epsilon} = \exp(-\delta\epsilon)[1-\exp(-\delta\epsilon)]\left(\frac{C_{t+\epsilon}}{V_{t+\epsilon}}\right)^{-\rho}\left(\frac{V_{t+\epsilon}}{\mathcal{R}_t}\right)^{-\gamma}\left(\frac{\mathcal{R}_t}{\mathcal{A}_t}\right)^{-\frac{(\alpha+\epsilon)}{\epsilon}}\left(\frac{\mathcal{A}_t}{V_t}\right)^{-\rho},$$

$$MC_t^0 = [1-\exp(-\delta\epsilon)]\left(\frac{C_t}{V_t}\right)^{-\rho},$$

where the risk adjustment is

$$\mathcal{R}_t = \left(\mathbb{E}\left[(V_{t+\epsilon})^{1-\gamma} \mid \mathfrak{F}_t\right]\right)^{\frac{1}{1-\gamma}},$$

and the ambiguity adjustment is

$$\mathcal{A}_t = \left(\mathbb{E}\left[(\mathcal{R}_t)^{-\frac{\alpha}{\epsilon}} \mid \mathfrak{G}_t\right]\right)^{-\frac{\epsilon}{\alpha}}.$$

As in the discrete-time model of Ju and Miao (2012), the marginal utility calculations use familiar formulas for CES functional forms and are computed by differentiating through the three recursions, (10) , (11), and (12).

The marginal utility ratio evaluated at equilibrium outcomes gives the stochastic discount factor increment:

$$\frac{S_{t+\epsilon}}{S_t} = \frac{MC_{t+\epsilon}^{\epsilon}}{MC_t^0} = \left(\frac{\widetilde{S}_{t,t+\epsilon}}{\widetilde{S}_t}\right)\left(\frac{\mathcal{R}_t}{\mathcal{A}_t}\right)^{-\frac{\alpha}{\epsilon}}, \tag{25}$$

where

$$\frac{\widetilde{S}_{t+\epsilon}}{\widetilde{S}_t} = \exp(-\delta\epsilon)\left(\frac{C_{t+\epsilon}}{C_t}\right)^{-\rho}\left(\frac{V_{t+\epsilon}}{V_t}\right)^{\rho-\gamma}\left(\frac{\mathcal{R}_t}{V_t}\right)^{\gamma-1}\left(\frac{\mathcal{A}_t}{V_t}\right)^{1-\rho}, \tag{26}$$

where the two components $\frac{\widetilde{S}_{t+\epsilon}}{\widetilde{S}_t}$ and $\left(\frac{\mathcal{R}_t}{\mathcal{A}_t}\right)^{-\frac{\alpha}{\epsilon}}$ to the stochastic discount factor in Eq. (25) behave qualitatively different as ϵ declines to zero. For instance, the small ϵ limit of $\frac{\widetilde{S}_{t+\epsilon}}{\widetilde{S}_t}$ is one, while we will derive a nonzero limit of the second term.

Consider first the local evolution of the $\widetilde{S}$ process under the $\mathfrak{F}$ filtration, which we represent as:

$$d\widetilde{S}_t = \widetilde{S}_t\widetilde{\mu}_t^S dt + \widetilde{S}_t\widetilde{\sigma}_t^S dW_t$$

implying a logarithmic counterpart:

$$d\log\widetilde{S}_t = \left(\widetilde{\mu}_t^S - \frac{1}{2}\left|\widetilde{\sigma}_t^S\right|^2\right)dt + \widetilde{\sigma}_t^S dW_t.$$

From formula (26), the local exposure of $d \log \widetilde{S}_t$ to dW_t is

$$\widetilde{\sigma}_t^S = -\rho\sigma_t^C + (\rho - \gamma)\sigma_t^V. \tag{27}$$

since the terms $\left(\frac{\mathcal{R}_t}{V_t}\right)^{\gamma-1}$ and $\left(\frac{\mathcal{A}_t}{V_t}\right)^{1-\rho}$ only contribute to the drift $\widetilde{\mu}_t^S$. In "Appendix A", we deduce the following local representation for this drift:

$$\begin{aligned}\tilde{\mu}_t^S =& -\delta - \rho\mu_t^C + \frac{1}{2}\left[\left(\rho^2+\rho\right)\left|\sigma_t^C\right|^2 - 2\rho(\rho-\gamma)\sigma_t^C\cdot\sigma_t^V + (\rho-\gamma)(\rho-1)\left|\sigma_t^V\right|^2\right]\\ &+ (\rho-1)\left(\mu_t^V + \frac{1}{\alpha}\log\mathbb{E}\left[\exp\left(-\alpha\mu_t^V\right) \mid \mathfrak{G}_t\right]\right).\end{aligned} \tag{28}$$

The contribution

$$\left(\frac{\mathcal{R}_t}{\mathcal{A}_t}\right)^{-\frac{\alpha}{\epsilon}} = \left(\frac{\mathcal{R}_t/V_t}{\mathcal{A}_t/V_t}\right)^{-\frac{\alpha}{\epsilon}}$$

to the stochastic discount factor requires special attention. This random variable is $\mathfrak{F}_t$ measurable and has expectation conditioned on $\mathfrak{G}_t$ equal to one. Thus it acts as a change in the probability measure for $\mathfrak{F}_t$ measurable events conditioned on $\mathfrak{G}_t$. In the appendix we characterize the limit as:

$$\widetilde{M}_t = \frac{\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right]}{\mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right)},$$

which induces the same probability adjustment as the robust interpretation given in Sect. 3. We denote the corresponding conditional expectation as $\widetilde{\mathbb{E}}\left(\cdot \mid \mathfrak{G}_t\right)$.

Consider a local return:

$$\frac{dQ_t}{Q_t} = \bar{\mu}_t^Q dt + \bar{\sigma}_t^Q d\overline{W}_t.$$

Given the ambiguity aversion we distinguish the risk exposure from the ambiguity exposure by rewriting this local evolution as:

$$\frac{dQ_t}{Q_t} = \bar{\mu}_t^Q dt + \bar{\sigma}_t^Q\left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\left(\mu_t^Y - \bar{\mu}_t^Y\right)dt + \bar{\sigma}_t^Q\left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\sigma_t^Y dW_t.$$

For the pricing, we first take account of the local contribution of $\widetilde{S}_t$ conditioned on the finer sigma algebra $\mathfrak{F}_t$:

$$\begin{aligned}\mathbb{E}\left[d\left(\widetilde{S}_t Q_t\right) \mid \mathfrak{F}_t\right] =& \ \widetilde{S}_t Q_t \bar{\mu}_t^Q dt + \widetilde{S}_t Q_t \widetilde{\mu}_t^S dt\\ &+ \widetilde{S}_t Q_t \bar{\sigma}_t^Q\left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\left(\mu_t^Y - \bar{\mu}_t^Y\right)dt\end{aligned}$$

$$+ \widetilde{S}_t Q_t \bar{\sigma}_t^Q \left[\sigma_t^Y \left(\sigma_t^Y \right)' \right]^{-1/2} \sigma_t^Y \left(\tilde{\sigma}_t^S \right)' dt$$

where $\tilde{\sigma}_t^S$ satisfies (27). Moreover, from (1) and (27),

$$\left[\sigma_t^Y \left(\sigma_t^Y \right)' \right]^{-1/2} \sigma_t^Y \left(\tilde{\sigma}_t^S \right)' = \left[-\rho \bar{\sigma}_t^C + (\rho - \gamma) \bar{\sigma}_t^V \right]'.$$

Next we use the change of probability measure that we derived for the ambiguity adjustment represented as a conditional expectation $\widetilde{\mathbb{E}}\left(\cdot \mid \mathfrak{G}_t\right)$ as input into the martingale pricing restriction under the $\mathfrak{G}$ filtration:

$$\bar{\mu}_t^Q = -\widetilde{\mathbb{E}}\left(\widetilde{\mu}_t^S \mid \mathfrak{G}_t \right) - \bar{\sigma}_t^Q \cdot \bar{\sigma}_t^S,$$

where

$$\overline{\sigma}_t^S = \left[-\rho \bar{\sigma}_t^C + (\rho - \gamma) \bar{\sigma}_t^V \right] + \left[\widetilde{\mathbb{E}}\left(\mu_t^Y \mid \mathfrak{G}_t \right) - \mathbb{E}\left(\mu_t^Y \mid \mathfrak{G}_t \right) \right]' \left[\sigma_t^Y \left(\sigma_t^Y \right)' \right]^{-1/2}.$$

This verifies that the vector $-\overline{\sigma}_t^S$ of uncertainty prices of exposure to the Brownian increment $d\overline{W}_t$ has two components, one that is familiar from recursive utility theory and other that depends on the difference between the $\widetilde{\mathbb{E}}$ and $\mathbb{E}$ expectations of μ_t^Y conditioned on $\mathfrak{G}_t$. This second contribution is novel and comes from the ambiguity adjustment.

Claim 4.1 *The risk-free rate is $-\widetilde{\mathbb{E}}\left(\widetilde{\mu}_t^s \mid \mathfrak{G}_t \right)$ where $\widetilde{\mu}_t^S$ is given by (28). Local prices of uncertainty for $d\overline{W}_t$ are the sum of the following two components:*

1. *(risk) $\rho \bar{\sigma}_t^C + (\gamma - \rho) \bar{\sigma}_t^V$;*
2. *(ambiguity)*

$$\left[\mathbb{E}\left(\mu_t^Y \mid \mathfrak{G}_t \right) - \widetilde{\mathbb{E}}\left(\mu_t^Y \mid \mathfrak{G}_t \right) \right]' \left[\sigma_t^Y \left(\sigma_t^Y \right)' \right]^{-1/2},$$

where the ambiguity-adjusted conditional expectation of $\mathfrak{F}_t$ events conditioned on $\mathfrak{G}_t$ information has Radon–Nikodym derivative

$$\frac{\exp\left[-\alpha \left(\mu_t^V - \bar{\mu}_t^V \right) \right]}{\mathbb{E}\left(\exp\left[-\alpha \left(\mu_t^V - \bar{\mu}_t^V \right) \right] \mid \mathfrak{G}_t \right)}.$$

4.3 Long-run uncertainty

We illustrate with the use of these formulas with a hidden state counterpart to the model of Hansen et al. (2008). The hidden state captures long-run risk as in the Bansal and

Yaron (2004).[6] In addition to a hidden state that evolves as a first-order autoregression, we include a second hidden state that is invariant over time and reflects uncertainty about mean return. The presence of both components adds to the long-term uncertainty. In contrast to the extensive "long-run risk" literature our decision maker confronts hidden state ambiguity because of a lack of confidence in the subjective specification of growth-rate dynamics. While our long-run uncertainty states are disguised to investors, we allow for multiple macroeconomic indicators of these states.[7]

Suppose that

$$
\begin{aligned}
dY_t &= A_{yy}\left(Y_t - \mathsf{u}_1 Z_t^1\right)dt + \mathsf{u}_1 dZ_t^1 + B_{yy} dW_t,\\
dZ_t &= A_{zz} Z_t dt + B_z dW_t,
\end{aligned}
$$

where $Z_t = \left(Z_t^1, Z_t^2, Z_t^3\right)'$ and u_1 is a column vector of ones with the same number of coordinates as Y. Assume that the eigenvalues of A_{yy} all have strictly negative real parts. The hidden process Z^1 captures a common stochastic trend for the macroeconomy, and the process $Y - \mathsf{u}_1 Z^1$ is asymptotically stationary. The process Z^1 has the familiar Bansal and Yaron dynamics. In particular,

$$
A_{zz} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & -\eta & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad B_z = \begin{bmatrix} B_{z1} \\ B_{z2} \\ 0 \end{bmatrix},
$$

where η is a positive scalar. The hidden state process Z^2 is recognizable as an Ornstein–Uhlenbeck process, a continuous-time version of scalar autoregression, and the process Z^3 is invariant over time and contributes an unknown time trend of the macroeconomy.

To match our previous specification of the Kalman–Bucy dynamics, we rewrite the first equation as

$$
dY_t = A_{yy} Y_t dt + A_{yz} Z_t dt + B_y dW_t,
$$

where

$$
\begin{aligned}
A_{yz} &= D_{yz} A_{zz} - A_{yy} D_{yz},\\
B_y &= B_{yy} + D_{yz} B_z,\\
D_{yz} &= \begin{bmatrix} \mathsf{u}_1 & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

As a precursor to computing a value function, it is straightforward to show that

$$
\begin{aligned}
\mathbb{E}\left(Y_{t+\tau} \mid Y_t, Z_t\right) = {} & \exp\left(A_{yy}\tau\right)\left(Y_t - D_{yz} Z_t\right) + \mathsf{u}_1 Z_t^1 \\
& + \frac{1}{\eta}\left[1 - \exp(-\eta\tau)\right]\mathsf{u}_1 Z_t^2 + \tau \mathsf{u}_1 Z_t^3.
\end{aligned}
$$

6 In contrast to the specification in this example, Bansal and Yaron (2004) treat their long-run risk process as observable to investors. By making long-run risk hidden to the investor, we open the door to an additional ambiguity adjustment asset values. While Bansal and Yaron (2004) allow for a stochastic volatility state, we abstract from such a state in this example.

7 For a related approach, see Hansen et al. (2008), who use such indicators in their quantitative analysis of macroeconomic uncertainty. In contrast to their formulation, we include hidden growth rate states.

The expected exponentially weighted expectation conditioned on (Y_t, Z_t) is

$$\delta\left[\mathbb{E}\int_0^\infty \exp(-\delta\tau)Y_{t+\tau}d\tau \mid Y_t, Z_t\right] = \delta\left(\delta I - A_{yy}\right)^{-1}\left(Y_t - D_{yz}Z_t\right) + \mathsf{u}_1 Z_t^1 + \left(\frac{1}{\eta+\delta}\right)\mathsf{u}_1 Z_t^2 + \frac{1}{\delta}\mathsf{u}_1 Z_t^3.$$

The first term on the right side is contribution from the transient macro dynamics and the second one from the long-run risk contribution. By the Law of Iterated Expectations,

$$\delta\left[\mathbb{E}\int_0^\infty \exp(-\delta\tau)Y_{t+\tau}d\tau \mid Y_t, \overline{Z}_t\right] = \delta\left(\delta I - A_{yy}\right)^{-1}\left(Y_t - D_{yz}\overline{Z}_t\right) + \mathsf{u}_1\left[1\ \tfrac{1}{\eta+\delta}\ \tfrac{1}{\delta}\right]\overline{Z}_t.$$

Since the matrix A_{yy} is nonsingular, when δ declines to zero, $\delta\left(-\delta I + A_{yy}\right)^{-1}$ converges to zero. Thus for small values of δ, the remaining terms dominate.[8] These are terms that feature growth rate uncertainty.

For preferences, we presume that $\rho = 1$ and that consumption satisfies

$$\log C_t = \mathsf{u}_c Y_t,$$

where u_c is a row vector of zeros with a one in the first position. Guess a value function of the form

$$\log V_t = \mathsf{v}_0(\Sigma_t) + \mathsf{v}_y Y_t + \mathsf{v}_z\overline{Z}_t,$$

where $\mathsf{v}_y Y_t + \mathsf{v}_z\overline{Z}_t$ is the exponentially weighted average of future expected log consumption given current information

$$\mathsf{v}_y = \delta\mathsf{u}_c\left(\delta I - A_{yy}\right)^{-1},$$
$$\mathsf{v}_z = -\delta\mathsf{u}_c\left(\delta I - A_{yy}\right)^{-1}D_{yz} + \left[1\ \tfrac{1}{\eta+\delta}\ \tfrac{1}{\delta}\right].$$

See "Appendix A" for details. Armed with these computations, the vector of risk prices is

$$\rho\bar{\sigma}_t^C + (\gamma-\rho)\bar{\sigma}_t^V = \left[\rho\mathsf{u}_c + (\gamma-\rho)\mathsf{v}_y\right]\left(B_y B_y'\right)^{1/2} + (\gamma-\rho)\mathsf{v}_z\Sigma_t A_{yz}'\left(B_y B_y'\right)^{-1/2}.$$

Consider next the ambiguity prices. Observe that

$$\mu_t^V - \bar{\mu}_t^V = \left[\mathsf{v}_y + \mathsf{v}_z\Sigma_t A_{yz}'\left(B_y B_y'\right)^{-1}\right]A_{yz}\left(Z_t - \overline{Z}_t\right), \quad (29)$$

[8] This is of relevance because much of the empirical asset pricing literature that studies long-run risk imposes a very small value of δ.

where $A_{yz}\left(Z_t - \overline{Z}_t\right)$ is distributed as a multivariate normal with conditional mean zero and conditional covariance matrix $A_{yz}\Sigma_t A_{yz}'$. In light of formula (29), the ambiguity adjusted probability maintains the normal distribution and the covariance matrix Σ_t, but the conditional mean is now

$$-\alpha A_{yz}\Sigma_t A_{yz}'\left[\mathsf{v}_y + \mathsf{v}_z\Sigma_t A_{yz}'\left(B_y B_y'\right)^{-1}\right]'.$$

This can be verified by a straightforward "complete-the-squares" argument. The corresponding contribution to the vector of uncertainty prices is

$$\begin{aligned}&\left[\mathbb{E}\left(\mu_t^Y \mid \mathfrak{G}_t\right) - \widetilde{\mathbb{E}}\left(\mu_t^Y \mid \mathfrak{G}_t\right)\right]'\left[\sigma_t^Y\left(\sigma_t^Y\right)'\right]^{-1/2}\\ &= -\alpha\left[\mathsf{v}_y + \mathsf{v}_z\Sigma_t A_{yz}'\left(B_y B_y'\right)^{-1}\right]A_{yz}\Sigma_t A_{yz}'\left(B_y B_y'\right)^{-1/2}.\end{aligned}$$

Notice that the altered mean in the normal distribution for the ambiguity adjustment to the Kalman–Bucy filtering depends only on Σ_t and not on either Y_t or Z_t. This result, however, is very special. The linearity of $\log V_t$ in $(Y_t, \overline{Z}_t)$ exploits both the presumed linear state dynamics and the presumption of a unitary elasticity of intertemporal substitution, $\rho = 1$. More generally, the value functions will not be linear in the stochastically varying state vector. Even in the Wonham filtering, the ambiguity adjustment will depend on the predicted state vector and the distributional adjustment for the hidden state will cease to be normal both under the original baseline distribution and the ambiguity adjusted distribution.

Collin-Dufresne et al. (2016) studied parameter learning in a model with learning about parameters governing long-run risk. In their framework, investors learn about parameters over time while committing to a unique prior. Here we show how to go further and incorporate ambiguity aversion about the subjective probabilities in a continuous-time environment.[9]

5 Portfolio choice

In this section we apply our utility model to a portfolio choice problem over an infinite horizon. We consider an environment where an investor can only invest in a risk-less asset with constant return r, and a risky stock with stochastic returns.

5.1 Uncertain return

The scalar Z is the instantaneous mean return on a risky asset, but this mean is not directly observed by the investor. Instead the investor uses past data to infer what this

[9] Ai (2010) explored long-run risk in a production economy with incomplete information. The methods described here open the door to extensions with investor ambiguity aversion.

mean is using observations on Y.[10] From the standpoint of Kalman–Bucy filtering, this is a special case of parameter estimation as described in Example 2.1 with $A_{yz} = 1$ and B_y a row vector with the same number of entries as the Brownian motion W. Formally, let the scalar cumulative return process Y evolve as

$$\begin{aligned} dY_t &= Zdt + B_y dW_t \\ &= \overline{Z}_t dt + \left(Z - \overline{Z}_t\right) dt + B_y dW_t \\ &= \overline{Z}_t dt + |B_y| d\overline{W}_t, \end{aligned}$$

where

$$\begin{aligned} d\overline{Z}_t &= \frac{\Sigma_t}{|B_y|^2}\left(dY_t - \overline{Z}_t dt\right) = \frac{\Sigma_t}{|B_y|^2}\left[\left(Z - \overline{Z}_t\right) dt + B_y dW_t\right], \\ d\Sigma_t &= -\frac{(\Sigma_t)^2}{|B_y|^2} dt. \end{aligned}$$

The solution to the second equation is

$$\Sigma_t = \frac{|B_y|^2 \Sigma_0}{t\Sigma_0 + \left|B_y\right|^2}.$$

As to be expected, for $\Sigma_0 > 0$, Σ_t declines to zero as t gets large. In other words, $\overline{Z}_t$ converges in mean square to the time invariant Z. Moreover,

$$dY_t - \overline{Z}_t dt = \left(Z - \overline{Z}_t\right) dt + B_y dW_t,$$

which isolates two component pertinent to investor preferences:

$$\underbrace{\left(Z - \overline{Z}_t\right) dt}_{\text{ambiguity}}, \underbrace{B_y dW_t}_{\text{risk}}.$$

The wealth of the investor is an endogenous state variable and evolves according to

$$\begin{aligned} dX_t &= \Psi_t X_t dY_t + (1 - \Psi_t) X_t r dt - C_t dt \\ &= X_t r dt - C_t dt + \Psi_t X_t \left[Zdt - rdt + B_y dW_t\right], \end{aligned}$$

where Ψ_t is the portfolio weight on the risky security and r is a riskless return, which is assumed to be constant. Applying Ito's Lemma,

$$d \log X_t = rdt - \left(\frac{C_t}{X_t}\right) dt - \frac{(\Psi_t)^2}{2} |B_y|^2 dt + \Psi_t \left(Zdt - rdt + B_y dW_t\right)$$

[10] This extends the analysis in Feldman (1992) who featured the case in which $\rho = \gamma = 1$ and $\alpha = 0$.

$$= r dt - \left(\frac{C_t}{X_t}\right) dt - \frac{(\Psi_t)^2}{2} |B_y|^2 dt + \Psi_t \left(\overline{Z}_t - r\right) dt + \Psi_t \left[\left(Z - \overline{Z}_t\right) dt + B_y dW_t\right].$$

We suppose that there are two control variables at each date: $(\Psi_t, \frac{C_t}{X_t})$. Set $\rho = 1$ and consider three contributions to the HJB equation from the continuous-time utility recursion using (16):

1. $\delta\left(\log C_t - \log V_t\right) + \bar{\mu}_t^V - \frac{1}{2}|\sigma_t^V|^2$;
2. $\frac{(1-\gamma)}{2}\left|\sigma_t^V\right|^2$;
3. $-\frac{1}{\alpha}\log \mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right)$.

If we were to only consider the first contribution, the resulting HJB equation would be that for the discounted expected logarithmic utility. The additional contributions reflect the adjustments for risk aversion that exceeds that of the discounted log utility benchmark and for ambiguity aversion.

We now construct the above contributions in turn. We use guess and verify for constructing the value function. In so doing, suppose

$$\log V_t = \log X_t + J_0\left(\Sigma_t\right) + \frac{1}{2} J_2\left(\Sigma_t\right)\left(\overline{Z}_t - r\right)^2$$

for functions (J_0, J_2). Let (x, z, s) denote the potential realizations of the state $(X_t, \overline{Z}_t, \Sigma_t)$ and (ψ, c) denote the potential realizations for the controls $(\Psi_t, \frac{C_t}{X_t})$.

- Using our value function guess, rewrite contribution 1) as

$$\begin{aligned} H_1(\psi, c \mid x, z, s) = \delta &\left[\log c - J_0(s) - \frac{1}{2} J_2(s)(z-r)^2\right] \\ &+ r - c - \frac{(\psi)^2}{2}|B_y|^2 + \psi(z-r) \\ &+ \frac{1}{2} J_2(s) \frac{s^2}{|B_y|^2} \\ &- \frac{d}{ds}\left[J_0(s) + \frac{1}{2} J_2(s)(z-r)^2\right] \frac{s^2}{|B_y|^2}. \end{aligned}$$

- Next consider the adjustment for risk given by contribution 2). Notice that

$$\sigma_t^V = \Psi_t B_y + J_2(s)\Sigma_t\left(\overline{Z}_t - r\right)\frac{\Sigma_t}{|B_y|^2} B_y.$$

This leads to depict the second contribution as:

$$H_2(\psi \mid x, z, s) = \frac{1-\gamma}{2}\left(\psi + J_2(s)(z-r)\frac{s}{|B_y|^2}\right)^2 |B_y|^2.$$

- Finally, consider the adjustment for ambiguity given by contribution 3). Notice that

$$\mu_t^V - \overline{\mu}_t^V = \left[\Psi_t + J_2\left(\Sigma_t\right)\left(\overline{Z}_t - r\right)\frac{\Sigma_t}{|B_y|^2} \right]\left(Z_t - \bar{Z}_t\right).$$

  Thus the random variable on the left side is normally distributed conditioned $\mathfrak{G}_t$ with mean zero. Compute the adjustment for ambiguity by using the formula for the mean of a log-normally distributed random variable

$$H_3(\psi \mid x, z, s) = -\frac{\alpha}{2}\left[\psi + J_2\left(s\right)\left(z - r\right)\frac{s}{|B_y|^2} \right]^2 s.$$

Combining these contributions, extending an argument in Duffie and Epstein (1992) implies that the HJB equation of interest is:

$$0 = \max_{\psi, c} H_1(\psi, c \mid x, z, s) + H_2(\psi \mid x, z, s) + H_3(\psi \mid x, z, s).$$

The first-order condition for ψ is

$$0 = z - r - |B_y|^2\psi + \left[\psi + J_2\left(s\right)\left(z - r\right)\frac{s}{|B_y|^2} \right]\left[(1-\gamma)|B_y|^2 - \alpha s \right].$$

Therefore, the maximizing choice is

$$\psi^* = \left[\frac{z - r}{\gamma|B_y|^2 + \alpha s} \right] - J_2\left(s\right)\frac{s}{|B_y|^2}\left[\frac{(\gamma - 1)|B_y|^2 + \alpha s}{\gamma|B_y|^2 + \alpha s} \right](z - r). \tag{30}$$

The portfolio choice has a structure very much analogous to that found in Merton (1971) and extended by Gennotte (1986) and Brennan (1998) to include learning. The first term in the square brackets on the right side of (30) captures the myopic mean-variance component and the second term reflects a hedging component. The formulas differ because Merton, Gennotte and Brennan presume that $\rho = \gamma$ and abstract from ambiguity aversion. Since we are interested in the case in which $\gamma \geq 1$, the directional impact of the second term depends on the sign of $J_2\left(s\right)(z - r)$. We will explore this impact in the numerical computations that follow.

Notice that what matters in these formulas is the composite uncertainty adjustment $\gamma|B_y|^2 + \alpha s$ and not the relative importance of the risk and ambiguity adjustments. Substituting Σ_t for s and factoring, we write:

$$\gamma\left|B_y\right|^2 + \alpha\Sigma_t = \left[\gamma + \alpha\left(\frac{\Sigma_t}{\left|B_y\right|^2}\right) \right]\left|B_y\right|^2, \tag{31}$$

where we plugged in Σ_t for s. The right side of (31) suggests that the ambiguity aversion can imitate enhanced risk aversion, albeit in a manner that is time dependent. Specifically, notice that α multiplies the variance ratio: $\frac{\Sigma_t}{|B_y|^2}$.

In terms of a robust Bayesian interpretation: the ambiguity aversion adjustment is the outcome of a minimization problem implying a distorted conditional mean return given by

$$\overline{Z}_t - \alpha \Sigma_t \left[\psi^* \left(\overline{Z}_t - r, \Sigma_t \right) + J_2 \left(\Sigma_t \right) \left(\overline{Z}_t - r \right) \frac{\Sigma_t}{|B_y|^2} \right]. \tag{32}$$

This opens the door to assess the plausibility of alternative choices of α following the robust Bayesian perspective of Good (1952).

Since we impose a unitary elasticity of intertemporal substitution, ($\rho = 1$), the optimized consumption–wealth ratio is constant:

$$\log C_t - \log X_t = \log \delta.$$

To obtain formulas for (J_0, J_2) we substitute the optimal decision rules in the HJB equation and obtain two equations to be solved. While there are typically multiple solutions, only one will be of interest and interpreted as the limit of finite-horizon solutions.

5.2 A quantitative illustration

To illustrate the impact of ambiguity aversion, we solve some numerical examples. We set parameter values as follows: $\gamma = 5$, $\rho = 1$, $\delta = 0.01$, $r = 0.02$, $B_y = 0.18$, $\Sigma_0 = 1\%$. The computations we report impose a terminal condition $\log V_T - \log C_T = 0$ for $T = 25$, which introduces additional time dependence into the continuation value process relative to the formulas in this section. See the "Appendix B" for details about how we solve the problem. By extending the decision horizon, value function and decision eventually converges to an infinite horizon portfolio problem with $\Sigma_t = 0$, rendering ambiguity aversion irrelevant. In "Appendix B", we also show results for the infinite-horizon problem.

Figure 1 presents the hedging, myopic, and total demand for the stock as functions of the expected excess returns $(\overline{Z} - r)$ at date zero for three values of ambiguity aversion $\alpha = 0, 3, 6$ and $\Sigma_0 = 1\%$. The case of $\alpha = 0$ corresponds to the ambiguity-neutral utility model of Duffie and Epstein (1992), adjusted to include learning. For simplicity, we focus on the case of positive expected excess returns because the case of negative expected excess returns is the mirror image. In the former case, the myopic demand is always positive.

Figure 1 shows that ambiguity aversion lowers the myopic demand as ambiguity aversion has the comparable effect of raising risk aversion. The hedging demand is negative for positive expected excess returns. Intuitively, bad news about stock returns following a negative shock is even worse as it also implies that expected future returns are low due to learning. Given this impact of news, the demand for the risky stock is diminished relative to the myopic (long) position. Increasing investor ambiguity aversion reduces the (short) hedging position compared with the Duffie and Epstein. While a more ambiguity averse investor has a more muted hedging

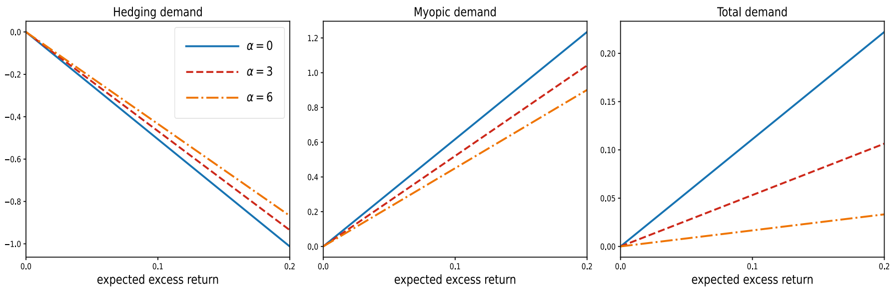

Fig. 1 The impact of ambiguity aversion at time zero on the demand for the risky asset for $\Sigma_0 = 1\%$. The three panels depict the myopic, hedging and total demands, respectively

response, the total demand declines for the larger value of α. To understand better, the numerical magnitudes of the two values of α, consider the implications of formula (32) for the robust-adjusted probabilities. As reported in "Appendix B", the worst-case probabilities reduce the expected excess return by about 20% for $\alpha = 3$ and by a little over 30% for $\alpha = 6$.

Figure 2 presents the impact of prior uncertainty Σ_0 for the Duffie and Epstein model ($\alpha = 0$) and for the smooth ambiguity model with $\alpha = 3$. For the $\alpha = 0$ model, increasing Σ_0 enhances the negative contribution of the hedging demand. The myopic demand does not depend on Σ_0. Thus the total demand declines as prior uncertainty is increased, but it remains positive for the values of Σ_0 that we consider. By contrast, for an ambiguity averse investor, the hedging demand is not monotonic with the prior uncertainty. Intuitively there are two opposite effects: higher prior uncertainty generates a larger short hedging position, but it also makes the impact of ambiguity aversion larger and hence lowers this short position. The latter effect is represented by the term $\frac{\alpha \Sigma_0}{(B_y)^2}$ and can dominate the first effect depending on parameter values. Because an increase in Σ_0 raises the impact of ambiguity aversion and hence effective risk aversion, it also reduces the myopic demand. As a result, the total demand for an ambiguity averse investor declines with Σ_0 and becomes negative for the largest of the three values of Σ_0 that we consider.

5.3 Modifications and extensions

We briefly describe some alternative economic environments and discuss the consequences of these extensions.

Remark 5.1 When $\rho \neq 1$, we consider continuation values of the more general form:

$$\log V_t = \log X_t + J\left(\Sigma_t, \overline{Z}_t\right)$$

While $\log V_t$ is no longer quadratic conditioned on $\overline{Z}_t$, the formula for the optimal portfolio has a direct extension involving the corresponding partial derivatives of J. The consumption wealth ratio will no longer be constant, but the decision rule will satisfy: For this more general specifications of ρ,

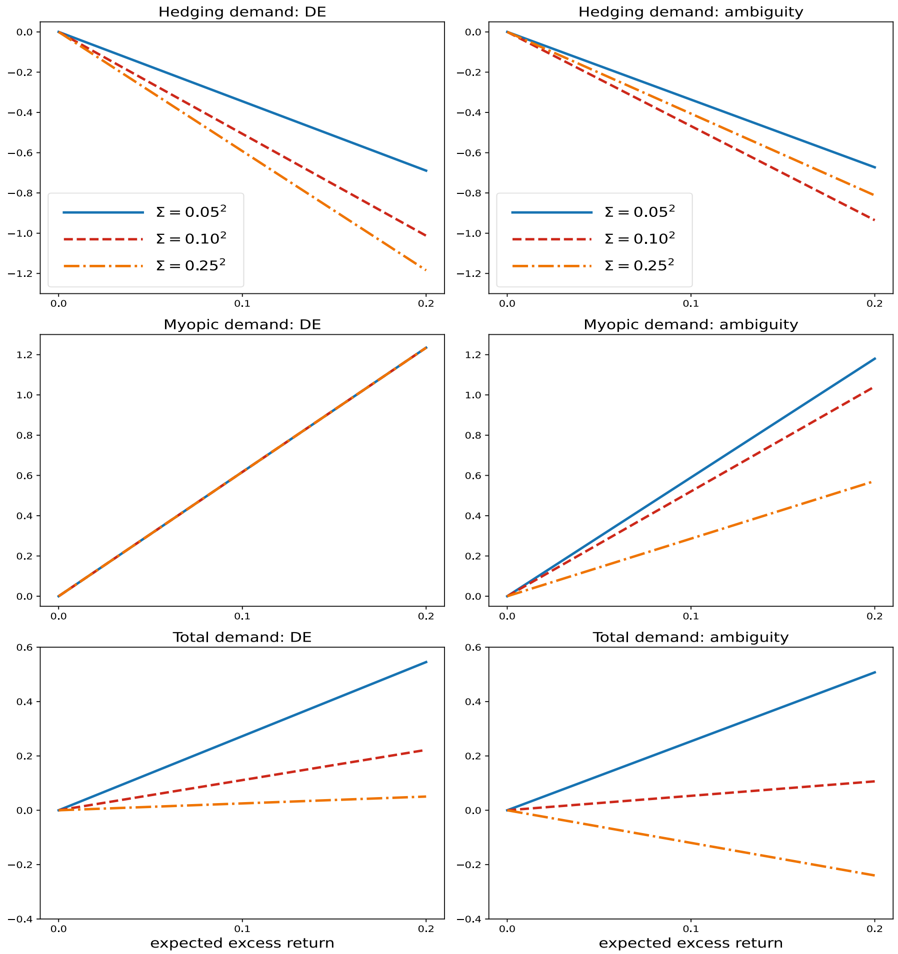

Fig. 2 The impact of prior uncertainty as captured by three alternative choices of Σ_0. The results depicted in the first column are for the model of Duffie and Epstein (1992) implemented by setting $\alpha = 0$. The results depicted in the second column presume that $\alpha = 3$. The rows give the alternative contributions to the demand for the risky asset

$$\log C_t - \log X_t = \frac{\log \delta}{\rho} + \left(\frac{\rho - 1}{\rho}\right)(\log V_t - \log X_t)$$

Because the continuation value is forward looking, the sign of $\rho - 1$ determines qualitatively how beliefs about the future, as captured by $\log V_t$, vary with the consumption–wealth ratio. This is true at least locally in $\rho - 1$.

Remark 5.2 As an alternative specifications of the learning problem with $\rho = 1$, we replace the Kalman–Bucy filtering dynamics with the Wonham counterpart. Recall that these dynamics are:

$$d\overline{Z}_t = \Lambda'\overline{Z}_t dt + \left[\operatorname{diag}\left(\overline{Z}_t\right) - \overline{Z}_t\overline{Z}_t'\right] A_{yz}' \left|B_y\right|^{-1} d\overline{W}_t$$
$$= \Lambda'\overline{Z}_t dt + \left[\operatorname{diag}\left(\overline{Z}_t\right) - \overline{Z}_t\overline{Z}_t'\right] A_{yz}' \left|B_y\right|^{-2} \left[A_{yz}\left(Z_t - \overline{Z}_t\right) dt + B_y dW_t\right],$$

where the return evolution is:

$$dY_t = A_{yz} Z_t dt + B_y dW_t = A_{yz}\overline{Z}_t dt + A_{yz}\left(Z_t - \overline{Z}_t\right) dt + B_y dW_t.$$

The corresponding wealth dynamics are:

$$d\log X_t = r dt - \left(\frac{C_t}{X_t}\right) dt - \frac{(\Psi_t)^2}{2}$$
$$|B_y|^2 dt + \psi_t \left[A_{yz}\overline{Z}_t dt - r dt + A_{yz}\left(Z_t - \overline{Z}_t\right) dt + B_y dW_t\right].$$

We again suppose that the continuation value is additively separable, and write:

$$V_t = X_t + J\left(\overline{Z}_t\right).$$

The construction of the HJB equation is analogous to the Kalman–Bucy filtering specification with modifications in the state evolution for $\overline{Z}$. It is straightforward to show that

$$\mu_t^V - \mathbb{E}\left(\mu_t^V \mid \mathfrak{G}_t\right) = \left[\psi_t + \left[\frac{\partial J}{\partial z}\left(\overline{Z}_t\right)\right]' \left[\operatorname{diag}\left(\overline{Z}_t\right) - \overline{Z}_t\overline{Z}_t'\right] A_{yz}' \left|B_y\right|^{-2}\right] A_{yz}\left(Z_t - \overline{Z}_t\right),$$

which conditionally linear in $Z_t - \overline{Z}_t$. The random vector, $Z_t - \overline{Z}_t$ is discrete in contrast to the Kalman–Bucy filtering dynamics. While the ambiguity computation:

$$-\frac{1}{\alpha}\log\mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right)$$

is straightforward, the resulting adjustment cannot be captured by a simple variance correction in contrast to the Kalman–Bucy filtering model.

Remark 5.3 The $\rho = 1$ model computation has a direct extension to the environment like that in Sect. 4.3 with long-run uncertainty and Kalman–Bucy filtering. Instead of positing a process for log consumption, we suppose that the logarithm of a valuation process formed by reinvesting any dividends back into the risk security has a long-term uncertainty component with hidden states. The implied instantaneous return replaces $r dt + B_y dW_t$ in the wealth dynamics. This extension allows for there to be predictability in the return process as has been documented in the empirical literature. The hidden state variable is now three dimensional. To accommodate the additional hidden states, the continuation value now takes the form:

$$\log V_t = \log X_t + J_0\left(\Sigma_t\right) + J_1\left(\Sigma_t\right)\overline{Z}_t + \frac{1}{2}\left(\overline{Z}_t\right)' J_2\left(\Sigma_t\right)\overline{Z}_t,$$

where J_1 is a row vector and J_2 is a symmetric matrix.

Remark 5.4 While the example takes the return process as exogenously specified, we could instead introduce production as part of social planner's problem, and explore capital accumulation and the implied equilibrium asset prices.[11] For instance, see Barnett et al. (2020), who study a climate change problem in presence of ambiguity aversion. It can be convenient for such a model to construct a max–min version of the HJB equation in which the minimization is represented as implied by (18).[12] The resulting HJB equation can then be formulated as a zero-sum, two-player differential game along the lines of Fleming and Souganidis (1989) and others.

6 Unknown jump processes

While our paper focuses on the case of Brownian information structure, we comment briefly on how to extend this analysis to accommodate jumps. Suppose that Y evolves according to one of n possible jump processes. We allow for a jump to take place at any t, and we use the notation t- to denote pre-jump values. That is, we follow the usual convention that processes are CADLAG (right continuous with left limits). For instance, if there is a jump at t, V_t is the post-jump continuation value and V_{t-} is the pre-jump continuation value. We represent the local evolution for process j with an intensity, $\mathcal{I}^j_{t-}$, implying that the approximate probability of a jump over a small interval ϵ is $\epsilon\mathcal{I}^j_{t-}$. We represent the jump distribution conditioned on a jump happening using an expectation operator $\mathbb{T}^j_{t-}$ for $j = 1, 2, ..., n$. Conditioning on $\mathfrak{F}_t$ at any date includes knowledge of which of the n possible stochastic process specifications governs the actual data evolution. This knowledge is not included in $\mathfrak{G}_t$. Instead the decision maker uses historical data to make inferences about the plausibility each of the j's.

We start by deducing a local representation of the recursive utility risk adjustment for each of the jump processes. That is, we use the $\mathfrak{F}$ filtration. Again our derivation will be heuristic providing candidate formulas to be formally justified with a more complete specification of the decision problem. Let $V_{t+\epsilon}$ be the continuation value at date $t+\epsilon$. Then the approximate conditional expectation of $(V_{t+\epsilon})^{1-\gamma}$ over an interval ϵ for jump process j is

$$\mathbb{E}\left[(V_{t+\epsilon})^{1-\gamma} \mid \mathfrak{F}_{t-}\right] \approx \left(1-\epsilon\mathcal{I}^j_{t-}\right)(V_{t-})^{1-\gamma} + \epsilon\mathcal{I}^j_{t-}\mathbb{T}^j_{t-}\left[(V_{t+\epsilon})^{1-\gamma}\right].$$

The small ϵ limit of interest is

$$\lim_{\epsilon\downarrow 0}\frac{\mathbb{E}\left[(V_{t+\epsilon})^{1-\gamma} \mid \mathfrak{F}_{t-}\right]-(V_{t-})^{1-\gamma}}{\epsilon} = \mathcal{I}^j_{t-}\left(\mathbb{T}^j_{t-}\left[(V_t)^{1-\gamma}\right]-(V_{t-})^{1-\gamma}\right).$$

[11] For instance, we could extend the results of Detemple (1986) who analyzed a production economy with learning in the case in which $\rho = \gamma = 1$ and $\alpha = 0$.

[12] In contrast to this paper, Barnett et al. (2020) abstract from learning.

By the "chain rule," two certainty equivalent counterpart approximations are

$$\lim_{\epsilon \downarrow 0} \frac{\left(\mathbb{E}\left[(V_{t+\epsilon})^{1-\gamma} \mid \mathfrak{F}_{t-}\right]\right)^{\frac{1}{1-\gamma}} - V_{t-}}{\epsilon} = \frac{\mathcal{I}^j_{t-}}{1-\gamma}\left(\mathbb{T}^j_{t-}\left[(V_t)^{1-\gamma}\right] - (V_{t-})^{1-\gamma}\right)(V_{t-})^{\gamma},$$

and

$$\lim_{\epsilon \downarrow 0} \frac{\frac{1}{1-\gamma}\log \mathbb{E}\left[(V_{t+\epsilon})^{1-\gamma} \mid \mathfrak{F}_{t-}\right] - \log V_{t-}}{\epsilon} = \frac{\mathcal{I}^j_{t-}}{1-\gamma}\left[\frac{\mathbb{T}^j_{t-}(V_t)^{1-\gamma}}{(V_{t-})^{1-\gamma}} - 1\right].$$

Using this computation, if the consumer/investor conditions on model j, the counterpart to (13) in the local construction of investor preferences with jump risk is

$$0 = \frac{\delta}{1-\rho}\left[\left(\frac{C_{t-}}{V_{t-}}\right)^{1-\rho} - 1\right] + \frac{\mathcal{I}^j_{t-}}{1-\gamma}\left[\frac{\mathbb{T}^j_{t-}(V_t)^{1-\gamma}}{(V_{t-})^{1-\gamma}} - 1\right].$$

The ambiguity neutral counterpart averages over the different possible jump processes conditioned on date t information:

$$0 = \frac{\delta}{1-\rho}\left[\left(\frac{C_{t-}}{V_{t-}}\right)^{1-\rho} - 1\right] + \frac{1}{1-\gamma}\sum_{j=1}^{n}\pi^j_{t-}\mathcal{I}^j_{t-}\left[\frac{\mathbb{T}^j_{t-}(V_t)^{1-\gamma}}{(V_{t-})^{1-\gamma}} - 1\right], \quad (33)$$

where π^j_{t-} is the probability of model j conditioned on $\mathfrak{G}_{t-}$. Because $\mathcal{I}^j_{t-}$ is in the conditioning information set of $\mathbb{T}^j_{t-}$, the expectations implied by π^j_{t-} and $\mathbb{T}^j_{t-}$ can be collapsed to one as in Bayesian analysis.[13]

By (12), our discrete-time approximation for the logarithm of the smooth ambiguity aversion adjustment is:[14]

$$\begin{aligned}
&-\frac{\epsilon}{\alpha}\log\sum_{j=1}^{n}\pi^j_{t-}\left[\left(\mathbb{E}\left[(V_{t+\epsilon})^{1-\gamma} \mid \mathfrak{G}_{t-}\right]\right)^{\frac{1}{1-\gamma}}\right]^{-\frac{\alpha}{\epsilon}} \\
&\quad \approx -\frac{\epsilon}{\alpha}\log\sum_{j=1}^{n}\pi^j_{t-}\exp\left[-\frac{\alpha}{\epsilon}\left(\log V_{t-} + \frac{\epsilon\mathcal{I}^j_{t-}}{1-\gamma}\left[\frac{\mathbb{T}^j_{t-}(V_t)^{1-\gamma}}{(V_{t-})^{1-\gamma}} - 1\right]\right)\right]
\end{aligned}$$

13 For instance, we can rewrite (33) in terms of an average intensity $\sum_{j=1}^{n}\pi^j_{t-1}\mathcal{I}^j_{t-}$ and corresponding average expectation conditioned on a jump as

$$\sum_{j=1}^{n}\left(\frac{\pi^j_{t-}\mathcal{I}^j_{t-}}{\sum_{j=1}^{n}\pi^j_{t-1}\mathcal{I}^j_{t-}}\right)\mathbb{T}^j_{t-}.$$

14 Analogous to the Brownian motion specification, Skiadas (2013) used α in place of $\frac{\alpha}{\epsilon}$ in the formula that follows when considering smooth ambiguity with jumps. As he shows, the smooth ambiguity adjustment vanishes in the limit, resulting in (33) in continuous time.

$$= \log V_{t-} - \frac{\epsilon}{\alpha} \log \sum_{j=1}^{n} \pi_{t-}^{j} \exp\left(-\frac{\alpha \mathcal{I}_{t-}^{j}}{1-\gamma}\left[\frac{\mathbb{T}_{t-}^{j}(V_t)^{1-\gamma}}{(V_{t-})^{1-\gamma}} - 1\right]\right).$$

As a consequence, our continuous-time smooth ambiguity adjustment alters the local representation of preferences to be:

$$0 = \min_{\tilde{\pi}^j \geq 0, \sum_{j=1}^{n} \tilde{\pi}^j = 1} \frac{\delta}{1-\rho}\left[\left(\frac{C_{t-}}{V_{t-}}\right)^{1-\rho} - 1\right] + \sum_{j=1}^{n} \tilde{\pi}^j \left(\frac{\mathcal{I}_{t-}^{j}}{1-\gamma}\right)\left[\frac{\mathbb{T}_{t-}^{j}(V_t)^{1-\gamma}}{(V_{t-})^{1-\gamma}} - 1\right]$$

$$+ \xi \sum_{j=1}^{n} \tilde{\pi}^j \left(\log \tilde{\pi}^j - \log \pi_{t-}^{j}\right)$$

$$= \frac{\delta}{1-\rho}\left[\left(\frac{C_{t-}}{V_{t-}}\right)^{1-\rho} - 1\right] - \frac{1}{\alpha} \log \sum_{j=1}^{n} \pi_{t-}^{j} \exp\left(-\frac{\alpha \mathcal{I}_{t-}^{j}}{1-\gamma}\left[\frac{\mathbb{T}_{t-}^{j}(V_t)^{1-\gamma}}{(V_{t-})^{1-\gamma}} - 1\right]\right).$$

where, as before, $\alpha = \frac{1}{\xi}$. In particular, the robust prior interpretation carries over to the inclusion of jumps.

This specification of preferences implies counterparts to HJB equations. As in the diffusion case, such equations take as inputs information state variables that capture the learning dynamics along with other state variables pertinent to the analysis. The formulation has direct extensions to more general filtering problems and to problems that include Brownian shocks along with jumps. A more full analysis of such problems is beyond the scope of this paper.

7 Conclusion

In this paper, we investigate dynamic decision problems that feature both ambiguity aversion and risk aversion posed in a continuous-time environment. These preferences extend specifications proposed by Duffie and Epstein (1992) by including a continuous-time counterpart to the recursive smooth ambiguity model of Klibanoff et al. (2009). The smooth ambiguity adjustment has a robust Bayesian interpretation whereby the decision-maker is unsure about the current period posterior distribution over hidden states or parameters. As in the robust Bayesian literature, we use this interpretation to produce a probabilistic adjustment to the posterior that captures the decision-maker's ambiguity. We show how both ambiguity and risk aversion reflect this probabilistic adjustment.

Acknowledgements We thank Shirui Chen and Zhenhuan Xie for research assistance and Diana Petrova for editorial assistance and two referees for helpful comments on an earlier draft of this paper. We also thank the Alfred P. Sloan Foundation (Grant G-2018-11113) for the financial support toward this research.

Declarations

Conflict of interest The authors do not have any financial or non-financial interests that are directly related to this research

A Pricing derivation

In this appendix, we provide some derivations that support the analysis in Sect. 4. By (25), we have

$$d\log S_t \approx \log S_{t+\epsilon} - \log S_t = \log \widetilde{S}_{t+\epsilon} - \log \widetilde{S}_t - \frac{\alpha}{\epsilon}\log\left(\frac{\mathcal{R}_t}{\mathcal{A}_t}\right).$$

Thus, to compute the drift and diffusion terms of $\log S_t$, we only need to compute the local mean and local Brownian motion exposure for $d\log \widetilde{S}_t$ and $\frac{\alpha}{\epsilon}\log\left(\frac{\mathcal{R}_t}{\mathcal{A}_t}\right)$.

We compute the local mean $\widetilde{\mu}_t^S$ and local Brownian exposure $\widetilde{\sigma}_t^S$ with respect to the $\mathfrak{F}$ filtration where

$$d\log \widetilde{S}_t = \left(\tilde{\mu}_t^S - \frac{1}{2}|\tilde{\sigma}_t^S|^2\right)dt + \tilde{\sigma}_t^S dW_t.$$

Write the logarithm:

$$\begin{aligned}\log \widetilde{S}_{t+\epsilon} - \log \widetilde{S}_t = & \left[-\delta\epsilon - \rho\left(\log C_{t+\epsilon} - \log C_t\right) + (\rho-\gamma)\left(\log V_{t+\epsilon} - \log V_t\right)\right] \\ & + \left[(\gamma-1)\left(\log \mathcal{R}_t - \log V_t\right)\right] + \left[(1-\rho)\left(\log \mathcal{A}_t - \log V_t\right)\right].\end{aligned}$$

We consider separately the local contributions of the three terms in $[\cdot]$. We use the following intuitive approximation of (22)

$$d\log C_t \approx \log C_{t+\epsilon} - \log C_t = \left(\mu_t^C - \frac{1}{2}\left|\sigma_t^C\right|^2\right)\epsilon + \sigma_t^C dW_t,$$

where dW_t is Gaussian with mean zero and covariance matrix ϵI conditioned on $\mathfrak{F}_t$. Similar approximations apply to other variables.

The first $[\cdot]$ term has local evolution:

$$\left[-\delta - \rho\mu_t^C + (\rho-\gamma)\mu_t^V + \frac{\rho\left|\sigma_t^C\right|^2}{2} - \frac{(\rho-\gamma)\left|\sigma_t^V\right|^2}{2}\right] dt + \widetilde{\sigma}_t^S dW_t,$$

where

$$\widetilde{\sigma}_t^S = -\rho\sigma_t^C + (\rho-\gamma)\sigma_t^V.$$

Write the second $[\cdot]$ term as

$$(\gamma-1)\left(\log \mathcal{R}_t - \log V_t\right) = -\log \mathbb{E}\left(\exp\left[(1-\gamma)(\log V_{t+\epsilon} - \log V_t)\right] \mid \mathfrak{F}_t\right).$$

The local (in time) approximation is

$$(\gamma-1)\left(\mu_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2\right) dt.$$

Write the third $[\cdot]$ term as

$$(1-\rho)\left(\log \mathcal{A}_t - \log V_t\right) = -\frac{\epsilon(1-\rho)}{\alpha}\log\left[\mathbb{E}\left(\exp\left[-\frac{\alpha}{\epsilon}\left(\log \mathcal{R}_t - \log V_t\right)\right] \mid \mathfrak{G}_t\right)\right].$$

The local in time approximation is

$$\begin{aligned}
&\frac{(\rho-1)}{\alpha}\log \mathbb{E}\left[\exp\left(-\alpha\left[\mu_t^V - \left(\frac{\gamma}{2}\right)\left|\sigma_t^V\right|^2\right]\right) \mid \mathfrak{G}_t\right] dt \\
&= \frac{(\rho-1)}{\alpha}\log \mathbb{E}\left[\exp\left(-\alpha\mu_t^V\right) \mid \mathfrak{G}_t\right] dt + \frac{(\rho-1)\gamma}{2}\left|\sigma_t^V\right|^2 dt.
\end{aligned}$$

Finally, to make the transformation back to levels we add the local variance adjustment

$$\frac{1}{2}\left|\widetilde{\sigma}_t^S\right|^2 = \frac{1}{2}\left[\rho^2\left|\sigma_t^C\right|^2 - 2\rho(\rho-\gamma)\sigma_t^C\cdot\sigma_t^V + (\rho-\gamma)^2\left|\sigma_t^V\right|^2\right].$$

Combining the local volatility adjustments gives

$$\begin{aligned}
&\frac{1}{2}\left[\left(\rho^2+\rho\right)\left|\sigma_t^C\right|^2 - 2\rho(\rho-\gamma)\sigma_t^C\cdot\sigma_t^V + \left[(\rho-\gamma)^2 - (\rho-\gamma) + (\rho-\gamma)\gamma\right]\left|\sigma_t^V\right|^2\right] \\
&= \frac{1}{2}\left[\left(\rho^2+\rho\right)\left|\sigma_t^C\right|^2 - 2\rho(\rho-\gamma)\sigma_t^C\cdot\sigma_t^V + (\rho-\gamma)(\rho-1)\left|\sigma_t^V\right|^2\right].
\end{aligned}$$

Combining all of the dt terms gives

$$\widetilde{\mu}_t^S = -\delta - \rho\mu_t^C$$

$$+ (\rho - 1)\left[\mu_t^V + \frac{1}{\alpha}\log \mathbb{E}\left[\exp\left(-\alpha\mu_t^V\right) \mid \mathfrak{G}_t\right]\right]$$
$$+ \frac{1}{2}\left[\left(\rho^2 + \rho\right)\left|\sigma_t^C\right|^2 - 2\rho(\rho-\gamma)\sigma_t^C \cdot \sigma_t^V + (\rho-\gamma)(\rho-1)\left|\sigma_t^V\right|^2\right].$$

To produce a limiting representation of the probability adjustment, write

$$\left(\frac{\mathcal{R}_t}{\mathcal{A}_t}\right)^{-\frac{\alpha}{\epsilon}} = \frac{\left(\frac{\mathcal{R}_t}{V_t}\right)^{-\frac{\alpha}{\epsilon}}}{\left(\frac{\mathcal{A}_t}{V_t}\right)^{-\frac{\alpha}{\epsilon}}}.$$

Note that

$$\begin{aligned}
&-\frac{\alpha}{\epsilon}\left(\log \mathcal{R}_t - \log V_t\right)\\
&= -\frac{\alpha}{\epsilon}\left[\frac{1}{1-\gamma}\log \mathbb{E}\left(\exp\left((1-\gamma)\log V_{t+\epsilon}\right) \mid \mathfrak{F}_t\right) - \log V_t\right]\\
&\approx -\alpha\left(\mu_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2\right).
\end{aligned}$$

Thus

$$\left(\frac{\mathcal{R}_t}{V_t}\right)^{-\frac{\alpha}{\epsilon}} \approx \exp\left[-\alpha\left(\mu_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2\right)\right]$$

and

$$\left(\frac{\mathcal{A}_t}{V_t}\right)^{-\frac{\alpha}{\epsilon}} = \mathbb{E}\left[\left(\frac{\mathcal{R}_t}{V_t}\right)^{-\frac{\alpha}{\epsilon}} \mid \mathfrak{G}_t\right] \approx \mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2\right)\right] \mid \mathfrak{G}_t\right).$$

Since $\bar{\mu}_t^V$ and $\left|\sigma_t^V\right|^2$ are $\mathfrak{G}_t$ measurable, we have

$$\begin{aligned}
\left(\frac{\mathcal{R}_t}{\mathcal{A}_t}\right)^{-\frac{\alpha}{\epsilon}} &\approx \frac{\exp\left[-\alpha\left(\mu_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2\right)\right]}{\mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \frac{\gamma}{2}\left|\sigma_t^V\right|^2\right)\right] \mid \mathfrak{G}_t\right)}\\
&= \frac{\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right]}{\mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right)}.
\end{aligned}$$

We use the right side random variable as a change of measure for a continuous-time limit.

Consider next the Kalman–Bucy example. We construct an Hamilton–Jacobi–Bellman (HJB) equation based on the recursive preference representation:

$$0 = \delta\left(\log C_t - \log V_t\right) + \bar{\mu}_t^V - \frac{1}{\alpha}\log \mathbb{E}\left(\exp\left[-\alpha\left(\mu_t^V - \bar{\mu}_t^V\right)\right] \mid \mathfrak{G}_t\right) - \frac{\gamma}{2}\left|\sigma_t^V\right|^2.$$

We use lower case variables (y, z, s) to denote potential realizations of $(Y_t, \overline{Z}_t, \Sigma_t)$. The HJB equation inclusive of ambiguity aversion is

$$
\begin{aligned}
0 = &\ \delta\left(\mathsf{u}_c y - \mathsf{v}_0(s) - \mathsf{v}_y y - \mathsf{v}_z z\right) \\
&+\mathsf{v}_y\left(A_{y0} + A_{yy}y + A_{yz}z\right) + \mathsf{v}_z\left(A_{z0} + A_{zy}y + A_{zz}z\right) \\
&-\frac{\alpha}{2}\left[\mathsf{v}_y\ \mathsf{v}_z\right]\begin{bmatrix} I \\ \left(sA'_{yz} + B_zB'_y\right)\left(B_yB'_y\right)^{-1}\end{bmatrix} A_{yz}sA'_{yz}\left[I\ \left(B_yB'_y\right)^{-1}\left(A_{yz}s + B_yB'_z\right)\right]\begin{bmatrix}\mathsf{v}'_y \\ \mathsf{v}'_z\end{bmatrix} \\
&-\frac{\gamma}{2}\left[\mathsf{v}_y\ \mathsf{v}_z\right]\begin{bmatrix} I \\ \left(sA'_{yz} + B_zB'_y\right)\left(B_yB'_y\right)^{-1}\end{bmatrix} B_yB'_y\left[I\ \left(B_yB'_y\right)^{-1}\left(A_{yz}s + B_yB'_z\right)\right]\begin{bmatrix}\mathsf{v}'_y \\ \mathsf{v}'_z\end{bmatrix} \\
&+\frac{\partial \mathsf{v}_0(s)}{\partial \text{vec}(s)'}\text{vec}\left[A_{zz}s + sA'_{zz} + B_zB'_z - sA'_{yz}\left(B_yB'_y\right)^{-1}A_{yz}\right],
\end{aligned}
$$

where vec(·) forms a column vector by stacking the non-redundant rows of the symmetric matrix argument. The HJB equation implies that the coefficient vectors v_y and v_z satisfy the linear equation:

$$
\left[\mathsf{v}_y\ \mathsf{v}_z\right]\begin{bmatrix} -\delta I + A_{yy} & A_{yz} \\ A_{zy} & -\delta I + A_{zz}\end{bmatrix} = -\left[\delta \mathsf{u}_c\ 0\right].
$$

The solution to this equation is also the solution to the exponentially weighted integral:

$$
\delta \mathsf{u}_c\left[\mathbb{E}\int_0^\infty \exp(-\delta\tau)Y_{t+\tau}d\tau \mid Y_t = y, Z_t = z\right]
$$

studied in Sect. 3. Finally, the remaining contributions isolate a differential equation for v_0. This function also solves the forward integral equation:

$$
\begin{aligned}
\mathsf{v}_0(\Sigma_t) = \frac{\delta}{2}\int_0^\infty &\exp(-\delta\tau)\left[\mathsf{v}_y\ \mathsf{v}_z\right]\begin{bmatrix} I \\ \left(\Sigma_{t+\tau}A'_{yz} + B_zB'_y\right)\left(B_yB'_y\right)^{-1}\end{bmatrix} \\
&\left[\alpha A_{yz}\Sigma_{t+\tau}A'_{yz} + \gamma B_yB'_y\right] \\
&\left[I\ \left(B_yB'_y\right)^{-1}\left(A_{yz}\Sigma_{t+\tau} + B'_yB_z\right)\right]\begin{bmatrix}\mathsf{v}'_y \\ \mathsf{v}'_z\end{bmatrix}d\tau.
\end{aligned}
$$

B Portfolio solution

For the results in Sect. 5, we solved the HJB equation numerically. While we posed the HJB equation in Sect. 5 using Σ_t as a state variable, our computations start from a single initialization. Given the initial Σ_0, Σ_t is strictly decreasing in t. Thus, for computational purposes, we express the value function in terms t instead of Σ_t, and solve the corresponding PDE. We derive this transformed PDE in the remainder of this appendix.

Write the value function guess as:

$$K(x, z, t) = x + K_0(t) + \frac{1}{2}K_2(t)(z - r)^2.$$

Then H_1, H_2, and H_3 are replaced by

$$\begin{aligned}\widetilde{H}_1(\psi, c \mid x, z, t) = \delta\left[\log c - K_0(t) - \frac{1}{2}K_2(t)(z - r)^2\right] \\ + r - c - \frac{(\psi)^2}{2}|B_y|^2 + \psi(z - r) \\ + \frac{1}{2}K_2(t)\left[\frac{(\Sigma_t)^2}{|B_y|^2}\right] \\ + \frac{d}{dt}\left[K_0(t) + \frac{1}{2}K_2(t)(z - r)^2\right],\end{aligned}$$

$$\widetilde{H}_2(\psi \mid x, z, t) = \frac{1-\gamma}{2}\left[\psi + K_2(t)(z - r)\left(\frac{\Sigma_t}{|B_y|^2}\right)\right]^2 |B_y|^2,$$

and

$$\widetilde{H}_3(\psi \mid x, z, t) = -\frac{\alpha}{2}\left[\psi + K_2(t)(z - r)\left(\frac{\Sigma_t}{|B_y|^2}\right)\right]^2 \Sigma_t.$$

The HJB equation of interest is:

$$0 = \max_{c,\psi} \widetilde{H}_1(\psi, c \mid x, z, t) + \widetilde{H}_2(\psi \mid x, z, t) + \widetilde{H}_3(\psi \mid x, z, t).$$

The first-order conditions for c imply that

$$c^* = \delta.$$

The first-order conditions for optimal portfolio choice is:

$$\left(\gamma|B_y|^2 + \alpha\Sigma_t\right)\psi^*(z, t) = z - r - K_2(t)(z - r)\left(\frac{\Sigma_t}{|B_y|^2}\right)\left[(\gamma - 1)|B_y|^2 + \alpha\Sigma_t\right].$$

Given these conditions, the combined linear and quadratic terms in the optimal portfolio weight are:

$$\frac{1}{2}\frac{\left[z - r - K_2(t)(z - r)\left(\frac{\Sigma_t}{|B_y|^2}\right)\left[(\gamma - 1)|B_y|^2 + \alpha\Sigma_t\right]\right]^2}{\gamma|B_y|^2 + \alpha\Sigma_t}.$$

Thus, we may rewrite the optimized objective as:

$$
\begin{aligned}
0 = & \delta \log \delta - \delta + r + \frac{1}{2} K_2(t) \left[\frac{(\Sigma_t)^2}{|B_y|^2} \right] \\
& + \frac{d}{dt} \left[K_0(t) + \frac{K_2(t)}{2} (z - r)^2 \right] - \delta \left[K_0(t) + \frac{K_2(t)}{2} (z - r)^2 \right] \\
& - \frac{1}{2} \left[K_2 (t) (z - r) \left(\frac{\Sigma_t}{|B_y|^2} \right) \right]^2 \left[(\gamma - 1) |B_y|^2 + \alpha \Sigma_t \right] \\
& + \frac{1}{2} \frac{\left[z - r - K_2 (t) (z - r) \left(\frac{\Sigma_t}{|B_y|^2} \right) \left[(\gamma - 1) |B_y|^2 + \alpha \Sigma_t \right] \right]^2}{\gamma |B_y|^2 + \alpha \Sigma_t}.
\end{aligned}
$$

Terms involving $(z - r)^2$ give rise to the following differential equation:

$$
\begin{aligned}
0 = & \frac{d}{dt} K_2(t) + \frac{1}{\gamma |B_y|^2 + \alpha \Sigma_t} \\
& - \delta K_2(t) - 2 \frac{\frac{\Sigma_t}{|B_y|^2} \left[(\gamma - 1) |B_y|^2 + \alpha \Sigma_t \right]}{\gamma |B_y|^2 + \alpha \Sigma_t} K_2(t) \\
& - \frac{\frac{(\Sigma_t)^2}{|B_y|^2} \left[(\gamma - 1) |B_y|^2 + \alpha \Sigma_t \right]}{\gamma |B_y|^2 + \alpha \Sigma_t} K_2(t)^2.
\end{aligned}
$$

Table 1 Slopes in Fig. 1

	Terminal condition 1	Terminal condition 2
Hedging demand		
$\alpha = 0$	− 5.52	− 5.06
$\alpha = 3$	− 5.13	− 4.68
$\alpha = 6$	− 4.78	− 4.34
Myopic demand		
$\alpha = 0$	6.17	6.17
$\alpha = 3$	5.21	5.21
$\alpha = 6$	4.50	4.50
Total demand		
$\alpha = 0$	0.65	1.11
$\alpha = 3$	0.08	0.53
$\alpha = 6$	− 0.28	0.17

Terminal condition 1: $T = 100{,}000$ and limiting value as terminal condition
Terminal condition 2: $T = 25$ and 0 as terminal condition

Table 2 Slopes for Fig. 2

	Terminal condition 1	Terminal condition 2
(a) DE ($\alpha = 0$)		
Hedging demand		
$\Sigma_0 = 0.05^2$	− 4.58	− 3.45
$\Sigma_0 = 0.10^2$	− 5.52	− 5.06
$\Sigma_0 = 0.25^2$	− 6.01	− 5.92
Myopic demand		
$\Sigma_0 = 0.05^2$	6.17	6.17
$\Sigma_0 = 0.10^2$	6.17	6.17
$\Sigma_0 = 0.25^2$	6.17	6.17
Total demand		
$\Sigma_0 = 0.05^2$	1.60	2.73
$\Sigma_0 = 0.10^2$	0.65	1.11
$\Sigma_0 = 0.25^2$	0.17	0.25
(b) Ambiguity ($\alpha = 3$))		
Hedging demand		
$\Sigma_0 = 0.05^2$	− 4.49	− 3.36
$\Sigma_0 = 0.10^2$	− 5.13	− 4.68
$\Sigma_0 = 0.25^2$	− 4.14	− 4.06
Myopic demand		
$\Sigma_0 = 0.05^2$	5.90	5.90
$\Sigma_0 = 0.10^2$	5.21	5.21
$\Sigma_0 = 0.25^2$	2.86	2.86
Total demand		
$\Sigma_0 = 0.05^2$	1.41	2.54
$\Sigma_0 = 0.10^2$	0.08	0.53
$\Sigma_0 = 0.25^2$	− 1.28	− 1.20

The remaining terms imply the following differential equation:

$$0 = \frac{d}{dt} K_0(t) - \delta K_0(t) + \delta \log \delta - \delta + r + \frac{1}{2} K_2(t) \frac{(\Sigma_t)^2}{|B_y|^2}.$$

We solve the two differential equations in sequence, plugging the solution to the first one into the second. When computing solutions, we use terminal conditions at date T. For the results reported in Figs. 1 and 2, we imposed that $K_2(25) = 0$. Here we also report results for the infinite-horizon counterpart by set T very large and impose $K_2(T)$ implied be $\Sigma_T = 0$.

$$0 = \frac{1}{\gamma |B_y|^2} - \delta K_2(T).$$

Table 3 Proportional reduction in the expected excess return under the worst-case model

	Terminal condition 1	Terminal condition 2
$\alpha = 3$	0.19	0.18
$\alpha = 6$	0.32	0.31

This terminal conditions coincides with the solution to an infinite horizon, recursive utility portfolio problem without learning and exposure to ambiguity. The corresponding terminal condition for $K_0(T)$ solves:

$$0 = \delta \log \delta - \delta + r - \delta K_0(T).$$

Tables 1 and 2a give the slopes of the portfolio rules depicted in Figs. 1 and 2, respectively, in comparison to the slopes implied by the infinite-horizon problem. The total demand slopes are lower for the infinite-horizon problem with the $\alpha = 6$ slope actually negative. See Table 1. The hedging demand remains non-monotone under ambiguity aversion as we vary Σ_0 for the infinite-horizon problem. See Table 2a for $\alpha = 3$.

Table 3 applies formula (32) to computes the proportional reduction in the expected excess return shows under the implied worst-case probabilities. The Table reports the implied slope (as a function of $\overline{Z}_t - r$) the worst-case increment:

$$\alpha \Sigma_t \left[\psi^* \left(\overline{Z}_t - r, \Sigma_t \right) + J_2 \left(\Sigma_t \right) \left(\overline{Z}_t - r \right) \frac{\Sigma_t}{|B_y|^2} \right].$$

This adjustment lowers the expected excess return by about 20% for $\alpha = 3$ and by a little over 30% for $\alpha = 6$ when $\Sigma_0 = .01$. As can be seen by the numbers reported in table, this conclusion is not very sensitive to whether we limit the decision horizon to be 25 years or allow it to be infinite.

While the appendix computes continuation values by replacing $s = \Sigma_t$ by t, the functions J_0 and J_2 can be inferred from the infinite-horizon solution described here by noting that $J_2(0) = K_2(\infty)$, $J_0(0) = K_0(\infty)$, and using the formula for $\frac{d\Sigma_t}{dt}$.

References

Ai, H.: Information quality and long-run risk: asset pricing implications. J. Finance **65**, 1333–1367 (2010)

Anderson, E.W., Hansen, L.P., Sargent, T.J.: A quartet of semigroups for model specification, robustness, prices of risk, and model detection. J. Eur. Econ. Assoc. **1**, 68–123 (2003)

Bansal, R., Yaron, A.: Risks for the long run: a potential resolution of asset pricing puzzles. J. Finance **59**(4), 1481–1509 (2004)

Barnett, M., Brock, W.A., Hansen, L.P.: Pricing uncertainty induced by climate change. Rev. Financ. Stud. **33**, 1024–1066 (2020)

Brennan, M.J.: The role of learning in dynamic portfolio decisions. Rev. Finance **1**, 295–306 (1998)

Chen, Z., Epstein, L.G.: Ambiguity, risk, and asset returns in continuous time. Econometrica **70**, 1403–1443 (2002)

Collin-Dufresne, P., Johannes, M., Lochstoer, L.A.: Parameter learning in general equilibrium: The asset pricing implications. Am. Econ. Rev. **106**, 664–698 (2016)

Detemple, J.B.: Asset pricing in a production economy with incomplete information. J. Financ. **41**, 383–391 (1986)
Duffie, D., Epstein, L.G.: Stochastic differential utility. Econometrica **60**, 353–394 (1992)
Epstein, L.G., Zin, S.E.: Substitution, risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica **57**, 937–969 (1989)
Feldman, D.: Logarithmic preferences, myopic decisions, and incomplete information. J. Financ. Quant. Anal. **27**, 619–629 (1992)
Fleming, W.H., Souganidis, P.E.: On the existence of value functions of two-player, zero-sum stochastic differential games. Indiana Univ. Math. J. **38**, 293–314 (1989)
Gennotte, G.: Optimal portfolio choice under incomplete information. J. Financ. **41**, 733–749 (1986)
Good, I.J.: Rational decisions. J. R. Stat. Soc. Ser. B (Methodol.) **14**, 107–114 (1952)
Hansen, L.P., Heaton, J.C., Li, N.: Consumption strikes back? Measuring long-run risk. J. Polit. Econ. **116**, 260–302 (2008)
Hansen, L.P., Miao, J.: Aversion to ambiguity and model misspecification in dynamic stochastic environments. Proc. Natl. Acad. Sci. **115**, 9163–9168 (2018)
Hansen, L.P., Sargent, T.J.: Recursive robust estimation and control without commitment. J. Econ. Theory **136**, 1–27 (2007)
Hansen, L.P., Sargent, T.J.: Robustness and ambiguity in continuous time. J. Econ. Theory **146**, 1195–1223 (2011)
Hayashi, T., Miao, J.: Intertemporal substitution and recursive smooth ambiguity. Theor. Econ. **6**, 423–472 (2011)
Ju, N., Miao, J.: Ambiguity, learning, and asset returns. Econometrica **80**, 559–591 (2012)
Klibanoff, P., Marinacci, M., Mukerji, S.: Recursive smooth ambiguity preferences. J. Econ. Theory **144**, 930–976 (2009)
Kreps, D.M., Porteus, E.L.: Temporal resolution of uncertainty and dynamic choice. Econometrica **46**, 185–200 (1978)
Liptser, R.S., Shiryaev, A.N.: Statistics of Random Processes I: General Theory, 2nd edn. Springer, New York (2001)
Liptser, R.S., Shiryaev, A.N.: Statistics of Random Processes II: Applications, 2nd edn. Springer, New York (2001)
Maccheroni, F., Marinacci, M., Rustichini, A.: Dynamic variational preferences. J. Econ. Theory **128**, 4–44 (2006)
Merton, R.C.: Optimum consumption and portfolio rules in a continuous-time model. J. Econ. Theory **3**, 373–413 (1971)
Skiadas, C.: Dynamic portfolio choice and risk aversion. In: Birge, J., Linetsky, V. (eds.) Financial Engineering, volume 15 of Handbooks in Operations Research and Management Science, pp. 789–843. Elsevier, New York (2007)
Skiadas, C.: Smooth ambiguity aversion toward small risks and continuous-time recursive utility. J. Polit. Econ. **121**(4), 775–792 (2013)
Wonham, W.M.: Some applications of stochastic differential equations to optimal nonlinear filtering. J. Soc. Ind. Appl. Math. Ser. A Control **2**(3), 347–369 (1965)