Boston University

OpenBU **http://open.bu.edu**

Economics BU Open Access Articles

2018-09-11

Aversion to ambiguity and model misspecification in dynamic stochastic environments

This work was made openly accessible by BU Faculty. Please share how this access benefits you. Your story matters.

Version	Published version
Citation (published version):	Lars Peter Hansen, Jianjun Miao. 2018. "Aversion to ambiguity and model misspecification in dynamic stochastic environments." Proceedings of the National Academy of Sciences, Volume 115, Issue 37, pp. 9163 - 9168. https://doi.org/10.1073/pnas.1811243115

https://hdl.handle.net/2144/35332
Boston University

Boston University Libraries
OpenBU

Aversion to ambiguity and model misspecification in dynamic stochastic environments

Lars Peter Hansen[a,b,c,1,2] **and Jianjun Miao**[d,e,f,1]

[a]Department of Economics, University of Chicago, Chicago, IL 60637; [b]Department of Statistics, University of Chicago, Chicago, IL 60637; [c]Booth School of Business, University of Chicago, Chicago, IL 60637; [d]Department of Economics, Boston University, Boston, MA 02215; [e]Institute of Chinese Financial Studies, Southwestern University of Finance and Economics, Chengdu 610074, China; and [f]China Economics and Management Academy, Central University of Finance and Economics, Beijing 100081, China

Contributed by Lars Peter Hansen, July 27, 2018 (sent for review June 29, 2018; reviewed by Fabio Maccheroni and Tomasz Strzalecki)

Preferences that accommodate aversion to subjective uncertainty and its potential misspecification in dynamic settings are a valuable tool of analysis in many disciplines. By generalizing previous analyses, we propose a tractable approach to incorporating broadly conceived responses to uncertainty. We illustrate our approach on some stylized stochastic environments. By design, these discrete time environments have revealing continuous time limits. Drawing on these illustrations, we construct recursive representations of intertemporal preferences that allow for penalized and smooth ambiguity aversion to subjective uncertainty. These recursive representations imply continuous time limiting Hamilton–Jacobi–Bellman equations for solving control problems in the presence of uncertainty.

uncertainty | robustness | misspecification | ambiguity | risk

In statistics, control theory, decision theory, and economics, the question of how to cope with subjective uncertainty comes into play. While researchers have developed many different approaches, it is standard in decision theory to impose axioms on preferences over choices. Axiomatic decision theory justifies representations of preferences that provide applied researchers with alternative ways to capture uncertainty responses. In applications, however, a decision maker must decide how to calibrate preference parameters, including aversion to ambiguity, and to model misspecification. An important issue that arises is how to transport these parameters across alternative environments, a question to which existing axiomatic treatments provide little guidance.

It has become common practice within economics to transport risk-aversion parameters and subjective discount rate parameters from one environment to another. Through some examples, we suggest that a mechanical implementation of the same approach is not appealing for some forms of aversion to ambiguity and model misspecification that interest us. These examples are not only of interest in their own right, but they also provide guidance as to meaningful continuous time limiting counterparts to Hamilton–Jacobi–Bellman (HJB) equations, for which the aversions contribute in the limit.

We build our analysis in multiple steps. We first use a version of Shannon's relative entropy to quantify uncertainty. Specifically, in section 1, we present a relative entropy decomposition of the joint probability distribution over a future observation and an unknown parameter that we find to be revealing when conceptualizing various forms of aversion. In section 2, we pose two alternative static robustness problems using relative entropy penalizations. Both are special cases of the variational preferences axiomatized by refs. 1 and 2. The first underlies the known connection between risk sensitive and robust control dating back to ref. 3. As we illustrate, this problem features the potential misspecification of the predictive density familiar from statistics. A second problem targets a concern about prior misspecification. In section 3, we solve these two problems using examples and explore some limits. By changing the exposure to an underlying shock, we distinguish in a sharp way the contributions from the likelihood and before the relative entropy of the joint distribution. We provide a way to scale aversion parameters so that their impact remains intact, even in the limiting economy.

The limits that we explore in section 3 are, by design, valuable inputs into the formulation and calibration of uncertainty preferences in dynamic stochastic environments. For example, in section 4, we study the impact of compounding the assessments of risk, ambiguity, and the potential for misspecification of subjective probabilities in a sequence of simple recursive learning environments. We also explore a convenient parametrization of smooth ambiguity specification from ref. 4, which we know from ref. 5 sometimes can be motivated by an aversion to prior uncertainty. We use these calculations to shed light on how to parameterize aversion to ambiguity and model misspecification across alternative dynamic environments, including an environment that emerges as a continuous time limit. Finally, in section 5, we propose a more general continuous time specification with distinct forms of aversions to alternative uncertainty components. This provides a counterpart of the discrete time specifications of refs. 5, 6, and 7. Moreover, it provides HJB equations that allow a decision maker to confront uncertainty in its various forms.

1. Entropy Penalization

Relative entropy is an attractive and tractable measure of discrepancy between probability measures. We use it to guide our

Significance

In many dynamic economic settings, a decision maker finds it challenging to quantify the uncertainty or assess the potential for mistakes in models. We explore alternative ways of acknowledging these challenges by drawing on insights from decision theory as conceptualized in statistics, engineering, and economics. We suggest tractable and revealing ways to incorporate behavioral responses to uncertainty, broadly conceived. Our analysis adopts recursive intertemporal preferences for decision makers that allow them to be ambiguity averse and concerned about the potential misspecification of subjective uncertainty. By design, these representations have revealing implications for continuous time environments with Brownian information structures. Problems where uncertainty's structure is obscure, such as macroeconomics, finance, and climate change, are promising areas for application of these tools.

Author contributions: L.P.H. and J.M. designed research, performed research, contributed new reagents/analytic tools, and wrote the paper.

Reviewers: F.M., Bocconi University; and T.S., Harvard University.

The authors declare no conflict of interest.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]L.P.H. and J.M. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: lhansen@uchicago.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1811243115/-/DCSupplemental.

Published online August 28, 2018.

ECONOMIC SCIENCES

APPLIED MATHEMATICS

formulation of preferences for a decision maker with an aversion to model ambiguity and concerns about misspecification. As a precursor to formulating such preferences, we explore the relative entropy relations between priors, likelihoods, and predictive densities in a static setting.

We start with a prior distribution π for hypothetical parameter values θ over a set Θ and a likelihood λ that informs us of the density for possible outcomes $y \in \mathcal{Y}$ given θ (with respect to a measure τ). Let $\mathbf{p}$ denote the implied joint probability measure over $\mathcal{Y} \times \Theta$:

$$\mathbf{p}(dy, d\theta) = \lambda(y|\theta)\tau(dy)\pi(d\theta).$$

The corresponding predictive density for y that integrates over θ is

$$\phi(y) = \int_\Theta \lambda(y|\theta)\pi(d\theta).$$

In our analysis, it is the predictive density that the decision maker cares about when taking actions, but he or she has uncertainty about subjective inputs into its construction. We introduce baseline counterparts with $\hat{}$ and explore discrepancies relative to these baselines.

The relative entropy discrepancy of the joint distribution for outcomes and parameters is defined as

$$D(\mathbf{p} \,|\, \widehat{\mathbf{p}}) = \int_\Theta \log\left[\frac{d\pi}{d\hat{\pi}}(\theta)\right]\pi(d\theta) + \int_\Theta\int_{\mathcal{Y}} \lambda(y|\theta)\log\left[\frac{\lambda(y|\theta)}{\hat{\lambda}(y|\theta)}\right]\tau(dy)\pi(d\theta), \quad \textbf{[1]}$$

where we presume that $\mathbf{p}$ is in a set $\mathcal{P}$ of probability measures that are absolutely continuous with respect to the baseline $\widehat{\mathbf{p}}$. This representation gives us two potential contributions to relative entropy: one coming from differences in the prior and the other coming from differences in the likelihood.

As we show in *SI Appendix*, there is an equivalent way to represent the same relative entropy in terms of predictive densities and posterior distributions, where the posterior distribution, π^+, is

$$\pi^+(d\theta \,|\, y) = \frac{\lambda(y\,|\,\theta)}{\phi(y)}\pi(d\theta)$$

and $\hat{\pi}^+$ is defined analogously.

2. Relative Entropy and Robustness Problem

In this section, we pose and solve two problems that adjust for robustness. These problems are the ingredients to preferences that capture a concern for robustness. While both use relative entropy to constrain the robustness assessment, one does so in a more restricted way to target the robust choice of a prior. The solutions to both are special cases of the solution to a more general problem that entails exponential tilting toward lower expected utilities.

Consider the first robust evaluation of a y-dependent utility $U(y)$.

Problem 2.1.

$$\min_{\mathbf{p}\in\mathcal{P}} \int U(y)d\mathbf{p} + \kappa D(\mathbf{p} \,|\, \widehat{\mathbf{p}}).$$

A more primitive starting point would be to write $\widetilde{U}(a, y)$, where a is chosen before y is observed. We may then rank alternative actions using the corresponding solutions to this problem. Notably, we presume that the utility function does not depend on the unknown parameter.

In this problem, $\kappa > 0$ is a parameter that penalizes the search for robustness in the likelihood and prior. The $\kappa = \infty$ limit enforces a commitment to the baseline specifications. Since the function $U(y)$ depends only on y and not on θ, the solution distorts only the predictive density

$$\phi^*(y) = \frac{\exp\left[-\frac{1}{\kappa}U(y)\right]\hat{\phi}(y)}{\int_{\mathcal{Y}}\exp\left[-\frac{1}{\kappa}U(y)\right]\hat{\phi}(y)\tau(dy)}$$

and not the posterior π^+ of θ given y. The resulting minimized objective is

$$-\kappa\log\int_{\mathcal{Y}}\exp\left[-\frac{1}{\kappa}U(y)\right]\hat{\phi}(y)\tau(dy).$$

The exponential tilting toward outcomes y with lower utility is familiar from the extensive literature applying relative entropy penalizations in control and estimation problems. In particular, *Problem 2.1* has the same solution as

$$\min_\phi \int_{\mathcal{Y}} U(y)\phi(y)\tau(dy) + \kappa\int_{\mathcal{Y}}\left[\log\phi(y) - \log\hat{\phi}(y)\right]\phi(y)\tau(dy).$$

This latter problem is the static counterpart to the dynamic recursive specification used extensively in robust control theory: for instance, refs. 3, 8, 9, 10, and 11.* While the construction of the predictive density $\hat{\phi}$ embeds the reference prior $\hat{\pi}$, prior sensitivity is only confronted in an indirect way by *Problem 2.1* through potential modifications in the predictive density.

As an alternative, we target prior robustness by restricting $\lambda = \hat{\lambda}$. This eliminates specification concerns about the likelihood. One justification for this omission is that uncertainty about how a prior weights alternative likelihoods already gives us a way to capture many forms of likelihood uncertainty. Let Π denote the set of priors that are absolutely continuous with respect to $\hat{\pi}$ and solve *Problem 2.2*.

Problem 2.2.

$$\min_{\pi\in\Pi}\int\overline{U}(\theta)\pi(d\theta) + \kappa\int_\Theta\log\left[\frac{d\pi}{d\hat{\pi}}(\theta)\right]\pi(d\theta),$$

where

$$\overline{U}(\theta) \equiv \int_{\mathcal{Y}} U(y)\hat{\lambda}(y|\theta)\tau(dy).$$

Notice that $\overline{U}$ is constructed without likelihood uncertainty. In using the formulation to investigate prior sensitivity, we penalize only the prior contribution to relative entropy. This formulation follows ref. 5.

The solution to this problem gives the worst case prior

$$\pi^*(d\theta) = \frac{\exp\left[-\frac{1}{\kappa}\overline{U}(\theta)\right]}{\int_\Theta\exp\left[-\frac{1}{\kappa}\overline{U}(\theta)\right]\hat{\pi}(d\theta)}\hat{\pi}(d\theta),$$

provided that the right-hand side integral is finite. The optimized objective is

$$-\kappa\log\int_\Theta\exp\left[-\frac{1}{\kappa}\overline{U}(\theta)\right]\hat{\pi}(d\theta). \quad \textbf{[2]}$$

The solution to this particular problem is recognizable as a smooth ambiguity objective and a special case of ref. 4.†

*Fabio Maccheroni reminded us that an alternative way to connect with this robust control theory literature is to impose a prior with a point mass on a single value of θ and targeting only likelihood uncertainty.

†Refs. 12 and 13 discuss earlier motivations for smooth ambiguity decision problems.

3. Revealing Family of Stochastic Environments

We study *Problems 2.1 and 2.2* for a family of stochastic environments to illustrate the impacts of the penalization. We design these illustrations to serve as a precursor to formulating continuous time limits to be used in a dynamic setting. They will provide input into formulations of HJB equations for continuous time decision problems.

Consider a parameterized family of log-normal random variables indexed by ϵ, and assume a logarithmic utility function U:

$$U(Y_\epsilon) = \log Y_\epsilon,\ \log Y_\epsilon = y_0 + \mu(\theta)\epsilon - \frac{1}{2}|\varsigma|^2\epsilon + \varsigma\cdot Z_\epsilon, \quad [3]$$

where μ is a measurable function of θ and Z_ϵ is a normally distributed random vector with mean 0 and covariance matrix ϵI.[‡] We find it convenient to think of $\epsilon > 0$ as an interval of time and $\{Z_\epsilon : 0 \le \epsilon \le 1\}$ as a multivariate standard Brownian motion.

Under the $\hat{\phi}_\epsilon$ predictive density, $\log Y_\epsilon$ is distributed as a mixture of a prior distribution $\hat{\pi}$ for θ and a normal distribution with mean

$$y_o + \epsilon\mu(\theta) - \frac{\epsilon}{2}|\varsigma|^2$$

and variance $\epsilon|\varsigma|^2$. For pedagogical simplicity, we consider the small ϵ approximation, whereby the distribution for $\log Y_\epsilon$ implied by $\hat{\phi}_\epsilon$ is normal with mean $y_o + \epsilon\int_\Theta \mu(\theta)\hat{\pi}(d\theta) - \frac{\epsilon}{2}|\varsigma|^2$ and variance $\epsilon|\varsigma|^2$. The minimized objective in *Problem 2.1* is

$$\begin{aligned} &-\kappa \log \int_{\mathcal{Y}} \exp\left[-\frac{1}{\kappa}U(y)\right]\hat{\phi}_\epsilon(y)\tau(dy) \\ &\approx y_o + \epsilon\left[\int_\Theta \mu(\theta)\hat{\pi}(d\theta) - \left(\frac{1}{2} + \frac{1}{2\kappa}\right)|\varsigma|^2\right]. \end{aligned}$$

Thus, up to the first-order approximation in ϵ, the decision maker targets concerns about misspecified likelihoods in contrast to misspecified priors.

Consider now *Problem 2.2*, where

$$\overline{U}_\epsilon(\theta) \equiv \int_{\mathcal{Y}} \log(y)\hat{\lambda}_\epsilon(y|\theta)\tau(dy). \quad [4]$$

The minimized objective is given by

$$\begin{aligned} &-\kappa \log \int_\Theta \exp\left[-\frac{1}{\kappa}\overline{U}_\epsilon(\theta)\right]\hat{\pi}(d\theta) \\ &\approx y_o + \epsilon\left[\int_\Theta \mu(\theta)\hat{\pi}(d\theta) - \frac{|\varsigma|^2}{2}\right], \end{aligned}$$

which is the same as the expected utility level under $\hat{\pi}$ up to the first-order approximation in ϵ. Thus, the robustness adjustment vanishes to the first order in ϵ. In contrast, the risk-aversion adjustment associated with logarithmic utility remains present.

Why is this the case? The probability distribution for Y_ϵ becomes concentrated as ϵ becomes small conditioned on θ. The same is true of the predictive distributions for alternative priors. By contrast, the prior divergence contribution to relative entropy does not depend on ϵ. This difference in how the prior and likelihoods behave as a function of ϵ causes the entropy penalty relative to the utility distortion to converge to infinity as ϵ goes to zero. Thus, the solution to *Problem 2.2* implies no robustness adjustment up to a first-order approximation.

[‡] We use $|\cdot|$ to denote the Euclidean norm and I to denote an identity matrix.

As an alternative, we alter *Problem 2.2* by letting the prior divergence scale in ϵ so that the consequence and cost of altering priors have comparable magnitudes up to the first order in ϵ. Thus, we allow κ to depend on ϵ by setting $\kappa(\epsilon) = \kappa_d\epsilon$ for some constant $\kappa_d > 0$ and consider *Problem 3.1*.

Problem 3.1.

$$\min_{\pi\in\Pi} \int_\Theta \overline{U}_\epsilon(\theta)\pi(d\theta) + \epsilon\kappa_d \int_\Theta \log\left[\frac{d\pi}{d\hat{\pi}}(\theta)\right]\pi(d\theta).$$

The minimized objective is now given by

$$-\epsilon\kappa_d \log \int_\Theta \exp\left[-\frac{1}{\epsilon\kappa_d}\overline{U}_\epsilon(\theta)\right]\hat{\pi}(d\theta).$$

Plugging Eq. **4** into this expression yields

$$y_0 + \epsilon\left\{-\kappa_d \log \int_\Theta \exp\left[-\frac{1}{\kappa_d}\mu(\theta)\right]\hat{\pi}(d\theta)\right\} - \frac{\epsilon}{2}|\varsigma|^2.$$

We may interpret the term in $\{\cdot\}$ as the certainty equivalent of $\mu(\theta)$ adjusted for robustness. By Jensen's Inequality, it is lower than the Bayesian mean of $\mu(\theta)$ under the baseline prior. This captures utility cost associated with the robust prior adjustment. The worst case prior is given by

$$\begin{aligned} \pi^*_\epsilon(d\theta) &= \frac{\exp\left[-\frac{1}{\epsilon\kappa_d}\overline{U}_\epsilon(\theta)\right]\hat{\pi}(d\theta)}{\int_\Theta \exp\left[-\frac{1}{\epsilon\kappa_d}\overline{U}_\epsilon(\theta)\right]\hat{\pi}(d\theta)} \\ &\to \frac{\exp\left[-\frac{1}{\kappa_d}\mu(\theta)\right]\hat{\pi}(d\theta)}{\int_\Theta \exp\left[-\frac{1}{\kappa_d}\mu(\theta)\right]\hat{\pi}(d\theta)} \end{aligned}$$

as $\epsilon \to 0$. Therefore, by reducing the penalization proportionally to ϵ, the exponential tilting applies to the first derivative of the objective function U with respect to ϵ. In so doing, we are left applying a particular smooth ambiguity adjustment to the local mean of $U(Y)$ conditioned on θ. The worst case prior puts more weight on the lower values of the mean $\mu(\theta)$.

We see little motivation for holding fixed the robustness penalization as we change ϵ. While the suggested adjustment is at the moment ad hoc, we will provide additional motivation in our subsequent analysis. In particular, this scaling lends itself to the study of prior robustness within a continuous time setting for dynamic decision problems.

The same issue emerges with other approaches to ambiguity as in the smooth ambiguity model in ref. 4. With that perspective, we view $-\frac{1}{\kappa(\epsilon)}$ as an exponential adjustment for ambiguity aversion capturing a differential preference response for exposure to ambiguity induced by a prior over θ, in contrast to risk conditioned on the parameter θ. By holding $\kappa(\epsilon)$ fixed independent of ϵ, we reproduce a claim in ref. 14. By instead using $\kappa(\epsilon) = \epsilon\kappa_d$, we obtain a limiting hyperbolic parameterization of smooth ambiguity aversion.

4. Recursive Risk and Smooth Ambiguity Aversion

We now explore consequences of compounding risk, ambiguity, and misspecification aversion over time. We continue to find the environment in section 3 captured in ref. 3 to be revealing from a pedagogical standpoint, allowing us to draw on some of our previous insights. The calculations in this section illustrate how risk compounds in rather different ways than uncertainty about priors over parameters. The resulting differences are pertinent to how we conceive of ambiguity and misspecification aversion in dynamic environments in general and in continuous time limit

environments in particular.[§] Given that there are known ways for ambiguity and misspecification aversion to have an impact on preferences in discrete time, we find it most appealing to adopt parameterizations with more meaningful and revealing continuous time limits.

In what follows, partition the interval $[0,1]$ into subintervals of length $1/n$, and let $\epsilon = 1/n$. Let $\mathcal{F}^{\epsilon}_{j\epsilon}$ be the sigma algebra generated by θ and $\{Y_0, Y_{1\epsilon}, Y_{2\epsilon}, \ldots, Y_{j\epsilon}\}$ for $j = 0, 1, 2, \ldots, n-1$.

A. Risk Aversion. We first consider a recursive construction of risk aversion over subintervals. Introduce a certainty equivalent operator applied to a positive random variable X that is $\mathcal{F}^{\epsilon}_{t+\epsilon}$ measurable:

$$\mathbb{R}_{\epsilon,t}(X) \equiv \left(E\left[X^{1-\gamma} \mid \mathcal{F}^{\epsilon}_t\right]\right)^{\frac{1}{1-\gamma}},$$

where $t = j\epsilon$ and $\gamma > 0$ represents the risk-aversion parameter. For $\gamma = 1$, $\mathbb{R}_{\epsilon,t}(X) \equiv \exp\left[E\left(\log X \mid \mathcal{F}^{\epsilon}_t\right)\right]$. Note that

$$\log \mathbb{R}_{\epsilon,t}\left(Y_{t+\epsilon}\right) = \log Y_t + \epsilon\left[\mu(\theta) - \frac{\gamma}{2}|\varsigma|^2\right]$$

for $\epsilon > 0$. The risk adjustment $\epsilon\gamma|\varsigma|^2/2$ scales with ϵ.

Introduce a recursive construction for the certainty equivalent of the terminal consumption Y_1,

$$U_t = \mathbb{R}_{\epsilon,t}\left(U_{t+\epsilon}\right),\ U_1 = Y_1 \qquad \textbf{[5]}$$

for $t = 0, \epsilon, \ldots, (n-1)\epsilon$. We show that

$$U_0 = \mathbb{R}_{1,0}\left(Y_1\right) = Y_0 \exp\left[\mu(\theta) - \frac{\gamma}{2}|\varsigma|^2\right]$$

for any value of ϵ. Notice that, in the recursive construction on the right-hand side of Eq. **5**, we use the same risk-aversion coefficient γ in the certainty equivalent operator for any ϵ. This simple example shows why the common practice of holding γ fixed across environments (in our case, indexed by ϵ) can be sensible, even as we shrink the exposure to risk by looking at small time intervals, due to the law of iterated expectations.

B. Parameter Learning. Before investigating a recursive specification of smooth ambiguity, we remind readers of Bayesian updating within this setting. Analogous to our previous discussion, we let $\mathcal{G}^{\epsilon}_{j\epsilon}$ be the sigma algebra generated by $\{Y_0, Y_{1\epsilon}, Y_{2\epsilon}\ldots, Y_{j\epsilon}\}$ for $j = 0, 1, 2, \ldots, n-1$, but we exclude the random parameter θ from the construction. The resulting family $\{\mathcal{G}^{\epsilon}_{j\epsilon} : j = 0, 1, 2, \ldots, n-1\}$ captures the reduced information structure for a decision maker that does not know the parameter realization. For simplicity, let $\mu(\theta) = \theta$.

Suppose that the date 0 prior for θ is normal with mean m_0 and variance q_0. Only data on Y are used to make inferences about θ. The posterior conditioned on $\mathcal{G}^{\epsilon}_t$ is normal with mean m_t and variance q_t. The Bayesian recursive updating for $m_{t+\epsilon}$ is

$$m_{t+\epsilon} - m_t = \frac{q_t}{|\varsigma|^2 + \epsilon q_t}\left(\log Y_{t+\epsilon} - \log Y_t - \epsilon m_t + \frac{\epsilon}{2}|\varsigma|^2\right), \qquad \textbf{[6]}$$

and the conditional variance update is

$$q_{t+\epsilon} - q_t = \frac{\epsilon (q_t)^2}{|\varsigma|^2 + \epsilon q_t}. \qquad \textbf{[7]}$$

Both equations have well-known continuous time limits. The limiting version of Eq. **6** for the conditional mean is

$$dm_t = \frac{q_t}{|\varsigma|^2}\left(d\log Y_t - m_t\,dt + \frac{|\varsigma|^2}{2}dt\right),$$

and the limiting version of Eq. **7** is

$$\frac{dq_t}{dt} = -\frac{(q_t)^2}{|\varsigma|^2}.$$

C. Recursive Robust Priors. Analogous to ref. 2, define a certainty equivalent robustness adjustment operator

$$\mathbb{B}_{\epsilon,t}(X) \equiv -\kappa(\epsilon)\log E\left[\exp\left(-\frac{1}{\kappa(\epsilon)}X\right) \mid \mathcal{G}^{\epsilon}_t\right]$$

for a random variable X that is $\mathcal{F}^{\epsilon}_t$ measurable. Here, we allow the robustness parameter $\kappa(\epsilon) > 0$ to depend on ϵ. Motivated by the static model in sections 3 and 4, we define the recursive adjustment for robust priors for log utility over the terminal consumption Y_1,

$$\log B_t = \mathbb{B}_{\epsilon,t}\left[E\left(\log B_{t+\epsilon} \mid \mathcal{F}^{\epsilon}_t\right)\right],\ B_1 = Y_1, \qquad \textbf{[8]}$$

for $t = 0, \epsilon, \ldots, (n-1)\epsilon$.

We show in *SI Appendix* that

$$\log B_t = \log Y_t + (1-t)m_t + \mathsf{b}_t,$$

where

$$\mathsf{b}_{t+\epsilon} - \mathsf{b}_t = \frac{\epsilon^2}{2\kappa(\epsilon)}q_t\left(\frac{|\varsigma|^2 + (1-t)q_t}{|\varsigma|^2 + \epsilon q_t}\right)^2 + \frac{\epsilon|\varsigma|^2}{2}.$$

From this formula, we see that the worst case prior/posterior for θ has a mean given by

$$m_t - \frac{\epsilon}{\kappa(\epsilon)}q_t\left(\frac{|\varsigma|^2 + (1-t)q_t}{|\varsigma|^2 + \epsilon q_t}\right).$$

If the robustness parameter $\kappa(\epsilon)$ is held constant independent of ϵ, then robustness vanishes in the continuous time limit as $\epsilon \to 0$. By contrast, we let $\kappa(\epsilon) = \kappa_d\epsilon$ as in *Problem 3.1* so that penalization used in making robust prior/posterior adjustments scales with ϵ. Take small ϵ limits, and find

$$\frac{d\mathsf{b}_t}{dt} = \frac{1}{2\kappa_d}q_t\left(\frac{|\varsigma|^2 + (1-t)q_t}{|\varsigma|^2}\right)^2 + \frac{|\varsigma|^2}{2}.$$

The first term on the right-hand side is the recursive robustness adjustment that remains present in a continuous time limit because of how we scale the robustness penalty as a function of ϵ.

D. Risk Aversion and Smooth Ambiguity. We next explore "smooth ambiguity adjustments" from ref. 4 over different time intervals, including ones that are arbitrarily small. For X that is $\mathcal{F}^{\epsilon}_t$ measurable, define a certainty equivalent operator

$$\mathbb{A}_{\epsilon,t}(X) \equiv \left[E\left(X^{-\alpha(\epsilon)}|\mathcal{G}^{\epsilon}_t\right)\right]^{-\frac{1}{\alpha(\epsilon)}},$$

[§] The calculations build in part on the prior work in ref. 15, which first discussed productive ways for robustness concerns to persist in a continuous time limiting specification. Our analysis here addresses these issues more generally and forges connections to a broader collection of previous contributions.

where $\alpha(\epsilon)>0$ may depend on ϵ and captures aversion to the uncertainty about the unknown parameter θ.

We now investigate how recursive smooth ambiguity behaves as a function of ϵ. Define smooth ambiguity adjustments recursively for utility over the terminal consumption Y_1,

$$A_t = \mathbb{A}_{\epsilon,t}\left[\mathbb{R}_{\epsilon,t}\left(A_{t+\epsilon}\right)\right], \; A_1 = Y_1. \quad \textbf{[9]}$$

In *SI Appendix*, we show that

$$\log A_t = \log Y_t + (1-t)m_t + \mathsf{a}_t,$$

where

$$\begin{aligned}\mathsf{a}_{t+\epsilon} - \mathsf{a}_t = &\\ &\frac{1}{2}\left(\frac{|\varsigma|^2+(1-t)q_t}{|\varsigma|^2+\epsilon q_t}\right)^2\left[\epsilon(\gamma-1)|\varsigma|^2+\alpha(\epsilon)\epsilon^2 q_t\right]\\ &+\frac{\epsilon|\varsigma|^2}{2}\end{aligned}$$

and $\mathsf{a}_1 = 0$.

Consistent with our earlier discussion of smooth ambiguity over small time intervals, by letting $\alpha(\epsilon)$ be the same for all $\epsilon>0$, smooth ambiguity contributes only a second-order adjustment.[¶]

In contrast, by adopting a hyperbolic parameterization $\alpha(\epsilon)=\frac{\alpha_h}{\epsilon}$, the smooth ambiguity contribution becomes first order and does not vanish in a continuous time limit. With this specification, the continuous time limiting equation becomes

$$\frac{d\mathsf{a}_t}{dt} = \frac{1}{2}\left(\frac{|\varsigma|^2+(1-t)q_t}{|\varsigma|^2}\right)^2\left[\alpha_h q_t + (\gamma-1)|\varsigma|^2\right] + \frac{|\varsigma|^2}{2}.$$

It matches our robust prior recursion $A_t = B_t$ by setting $\gamma=1$ and $\alpha_h = \frac{1}{\kappa_d}$. Thus, the smooth ambiguity adjustments are equivalent to robust prior adjustments when using a logarithmic specification of risk aversion.

In obtaining the limiting recursion, we held the risk-aversion parameter the same over environments indexed by ϵ, while we scaled the ambiguity aversion or robustness parameter with ϵ. While there is no perfect way to transform ambiguity-aversion parameters across environments with different discrete time increments, the embedding that we suggest has the virtue of possessing a tractable continuous time limit. While we may think that setting $\alpha(\epsilon)=\frac{\alpha_h}{\epsilon}$ amounts to imposing "infinite" ambiguity aversion in the continuous time limit, this is misguided in our view. Notice that, in the continuous time limit, α_h scales the variance associated with estimation. As an analogy to our scaling, let $0<\beta<1$ denote the discount factor over a unit of time. Then, we think of $\beta^{\frac{1}{n}}$ as the discount factor over an interval $\frac{1}{n}$, which in the large n limit, implies a unit discount factor in the limit. However, in fact, the rate $-\log\beta$ continues to play a role of discounting in the continuous time limit.

5. Implications for Stochastic Control

Let $(\Omega,\mathcal{F},\mathbf{P})$ be a probability space, and time is continuous over $[0,T]$. Let $\{Y_t : t\geq 0\}$ be an observable stochastic process and θ be an unknown parameter. Let $\{\mathcal{F}_t\}$ be the filtration generated by current and past Y_t and θ, and let $\{\mathcal{G}_t\}$ be the filtration generated by current and past Y_t only. The decision maker's continuation value process (V_t) is adapted to $\{\mathcal{G}_t\}$ and satisfies a backward stochastic differential equation:

$$dV_t = V_t\nu_t(\theta)\,dt + V_t\varsigma_t\cdot dW_t, \; V_T \text{ given},$$

where $\{W_t : t\geq 0\}$ is a standard Brownian motion relative to $(\Omega,\mathcal{F},\{\mathcal{F}_t\},\mathbf{P})$. The continuation value process will be used to give recursive construction of the decision maker's preference over consumption processes. We will derive an HJB equation for the $\{V_t\}$ to incorporate risk aversion, smooth ambiguity aversion, and/or prior misspecification aversion.

We find it most convenient to use the logarithm of the continuation value $v_t=\log V_t$, which with an application of Ito's Lemma, satisfies

$$dv_t = \nu_t(\theta)\,dt - \frac{1}{2}|\varsigma_t|^2\,dt + \varsigma_t\cdot dW_t.$$

This evolution captures the familiar lognormal adjustment.

Motivated by the analyses in sections 3–5, we define risk and ambiguity adjustment operators:

$$\begin{aligned}\mathbb{R}^*_{\epsilon,t}\left(v_{t+\epsilon}\right) &\equiv \frac{1}{1-\gamma}\log E\left(\exp\left[(1-\gamma)v_{t+\epsilon}\right]|\mathcal{F}_t\right),\\ \mathbb{B}^*_{\epsilon,t}\left(w_{\epsilon,t}\right) &\equiv -\frac{\epsilon}{\alpha_h}\log E\left[\exp\left(-\frac{\alpha_h}{\epsilon}w_{\epsilon,t}\right)|\,\mathcal{G}_t\right],\end{aligned}$$

where $w_{\epsilon,t}$ is $\mathcal{F}_t$ measurable. In what follows, let

$$v^*_{\epsilon,t} = \mathbb{B}^*_{\epsilon,t}\circ\mathbb{R}^*_{\epsilon,t}\left(v_{t+\epsilon}\right)$$

and $V^*_{\epsilon,t} = \exp\left(v^*_{\epsilon,t}\right)$.

We consider a preference specification that includes intermediate consumption and subjective discounting. Suppose that V_t satisfies the commonly used and convenient recursion

$$V_t = \left(\left[1-\exp(-\delta\epsilon)\right](C_t)^{1-\rho} + \exp(-\delta\epsilon)\left(V^*_{\epsilon,t}\right)^{1-\rho}\right)^{\frac{1}{1-\rho}},$$

where $\delta>0$ is the subjective discount rate, $\{C_t : t\geq 0\}$ is a consumption process adapted to $\{\mathcal{G}_t : t\geq 0\}$, and $\frac{1}{\rho}$ is an intertemporal elasticity of substitution. Using the homogeneity property of this recursion and taking logarithms, write

$$\begin{aligned}0 = \frac{1}{1-\rho}\log&\left[\left[1-\exp(-\delta\epsilon)\right]\left(\frac{C_t}{V_t}\right)^{1-\rho}\right.\\ &\left.+\exp(-\delta\epsilon)\left(\frac{V^*_{\epsilon,t}}{V_t}\right)^{1-\rho}\right]. \quad \textbf{[10]}\end{aligned}$$

We derive (heuristically) the limit of the right-hand side of the equal sign of Eq. **10**. This will characterize a restriction on local evolution of the continuation value process. Notice that

$$\begin{aligned}\mathbb{R}^*_{\epsilon,t}\left(v_{t+\epsilon}-v_t\right) &= \mathbb{R}^*_{\epsilon,t}\left(v_{t+\epsilon}\right) - v_t,\\ \mathbb{B}^*_{\epsilon,t}\left(w_{\epsilon,t}-v_t\right) &= \mathbb{B}^*_{\epsilon,t}\left(w_{\epsilon,t}\right) - v_t,\end{aligned}$$

where v_t is $\mathcal{G}_t$ measurable. Extending our previous calculations, we first study the local risk adjustment, $\mathbb{R}^*_{\epsilon,t}\left(v_{t+\epsilon}-v_t\right)$:

$$\begin{aligned}&\lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\left(\frac{1}{1-\gamma}\right)\log E\left(\exp\left[(1-\gamma)(v_{t+\epsilon}-v_t)\right]|\mathcal{F}_t\right)\\ &= \nu_t(\theta) - \frac{\gamma}{2}|\varsigma_t|^2.\end{aligned}$$

Let $w_{\epsilon,t} = \mathbb{R}^*_{\epsilon,t}\left(v_{t+\epsilon}\right)$ and $w_{0,t} = v_t$, and compute

$$\begin{aligned}&\lim_{\epsilon\downarrow 0}\frac{1}{\epsilon}\mathbb{B}^*_{\epsilon,t}\left(w_{\epsilon,t}-v_t\right)\\ &= -\frac{1}{\alpha_h}\log E\left(\exp\left[-\alpha_h\nu_t(\theta)\right]|\mathcal{G}_t\right) - \frac{\gamma}{2}|\varsigma_t|^2.\end{aligned}$$

[¶] Researchers (16) have a revealing but different type of limiting characterization of risk aversion and smooth ambiguity in a static environment.

With these intermediate calculations, we construct the local counterpart to Eq. **10** by dividing both sides of the equation by ϵ and taking small ϵ limits, resulting in

$$0 = \frac{\delta}{1-\rho}\left[\left(\frac{C_t}{V_t}\right)^{1-\rho} - 1\right] - \frac{1}{\alpha_h}\log E\left(\exp\left[-\alpha_h \nu_t(\theta)\right]|\mathcal{G}_t\right) - \frac{\gamma}{2}|\varsigma_t|^2. \qquad [11]$$

The first term on the right-hand side comes from discounting, intermediate consumption, and intertemporal substitution. The second term captures uncertainty to the drift induced by the unknown parameter vector θ. We motivate this either as a robust prior/posterior adjustment or as an aversion to "smooth ambiguity" in the unknown drift. The third term captures the local risk adjustment coming from exposure to the underlying Brownian motion.

Using this limiting recursion for an HJB equation requires that we use a value function and its derivatives to deduce formulas for $\nu_t(\theta)$ and ς_t as functions of the relevant Markov-state vector.

We conclude with four connections to related literature in control theory and economics.

Remark 5.1: Recursion Eq. **11** provides a continuous time analog to discrete time recursions in refs. 6 and 7. Moreover, it shows how to incorporate ambiguity aversion into the continuous time specifications of refs. 17 and 18.

Remark 5.2: The ambiguity adjustment

$$-\frac{1}{\alpha_h}\log E\left(\exp\left[-\alpha_h \nu_t(\theta)\right]|\mathcal{G}_t\right) \qquad [12]$$

for the drift is a smooth counterpart to a continuous time ambiguity adjustment with period-by-period constraints when the random set $H_t = \{\nu_t(\theta) : \theta \in \Theta\}$ is a compact subset of a Euclidean space with probability 1. As α_h becomes arbitrarily large, the smooth ambiguity adjustment converges to the minimum over H_t. The result is a continuation value recursion of the form considered in ref. 19.

More generally, $E\left(\exp\left[-\alpha_h \nu_t(\theta)\right]|\mathcal{G}_t\right)$ is the moment-generating function for ν_t when viewed as a function of θ conditioned on $\mathcal{G}_t$. In general, a moment-generating function is not guaranteed to be finite for all α_h. Even when the moment-generating function is finite for small α_h, there may only be a compact interval of α_h for which the function remains finite.

The existence of a finite upper bound for α_h has a recognizable connection to breakdown points in dynamic control theory.

Remark 5.3: Recursive Eq. **11** also gives the continuous time counterpart to a discrete time recursion in ref. 5 that captures two forms of robustness.

As we discussed previously, smooth ambiguity adjustment Eq. **12** has a dual interpretation as the outcome of prior/posterior uncertainty adjustment. For the likelihood adjustment, there is a well-known link between risk sensitivity and robustness dating back to ref. 3.[#] Capture the likelihood uncertainty by representing $dW_t = h_t dt + dW_t^h$ under a change of probability measure, for which dW_t^h is a Brownian increment and h_t is a local drift distortion. This adjustment alters the implied value function ν_t by $\varsigma_t \cdot h_t$. Minimizing the local evolution of the value function by the choice of h_t subject to a penalization, $\frac{\kappa_f}{2}|h_t|^2$, illustrates a link between robustness and risk sensitivity that is familiar in the control literature.

Remark 5.4: To construct the filtration $\{\mathcal{G}_t : t \geq 0\}$ in practice and use it for solving a control problem, we must produce a recursive solution for a filtering or estimation problem. While we posed the analysis as one for which θ is an unknown parameter, in fact, this parameter could be an actual process designed to capture time variation in some underlying parameters. The sigma algebra $\mathcal{G}_t$ could condition on the entire process or just the process up to time t. Examples of recursive constructions include Zakai equations, Kalman filtering, or Wohham filtering depending on the application.

ACKNOWLEDGMENTS. We thank Michael Barnett, William Brock, Peter Hansen, Massimo Marinacci, and Grace Tsiang for helpful comments. L.P.H. acknowledges Alfred P. Sloan Foundation Grant G-2018-11113 for the support.

[#] A continuous time analysis is, for instance, in ref. 9.

1. Maccheroni F, Marinacci M, Rustinchini A (2006) Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica* 74:1147–1498.
2. Strzalecki T (2011) Axiomatic foundations of multiplier preferences. *Econometrica* 79:47–73.
3. Jacobson DH (1973) Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Trans Automatic Contr* 18:1124–1131.
4. Klibanoff P, Marinacci M, Mukerji S (2005) A smooth model of decision making under uncertainty. *Econometrica* 73:1849–1892.
5. Hansen LP, Sargent TJ (2007) Recursive robust estimation and control without commitment. *J Econ Theor* 136:1–27.
6. Hayashi T, Miao J (2011) Intertemporal substitution and recursive smooth ambiguity preferences. *Theor Econ* 6:423–472.
7. Ju N, Miao J (2012) Ambiguity, learning, and asset returns. *Econometrica* 80:559–591.
8. Doyle JC, Glover K, Khargonekar PP, Francis B (1989) State space solutions for h_2 and h_∞ control problems. *IEEE Trans Automat Contr* 34:831–847.
9. James MR (1992) Asymptotic analysis of nonlinear stochastic risk-sensitive control and differential games. *Math Control Signals Syst* 5:401–417.
10. Basar T, Bernhard P (1995) *H^∞-Optimal Control and Related Minimax Problems: A Dynamic Game Theory Approach* (Birkhauser, Boston), 2nd Ed.
11. Petersen IR, James MR, Dupuis P (2000) Minimax optimal control of stochastic uncertain systems with relative entropy constraints. *Automatic Control IEEE Trans* 45:398–412.
12. Segal U (1990) Two-stage lotteries without the reduction axiom. *Econometrica* 58:349–377.
13. Davis DB, Paté-Cornell M-E (1994) A challenge to the compound lottery axiom: A two-stage normative structure and comparison to other theories. *Theor Decis* 37:267–309.
14. Skiadas C (2013) Smooth ambiguity aversion toward small risks and continuous-time recursive utility. *J Polit Econ* 121:775–792.
15. Hansen LP, Sargent TJ (2011) Robustness and ambiguity in continuous time. *J Econ Theor* 146:1195–1223.
16. Maccheroni F, Marinacci M, Ruffino D (2013) Alpha as ambiguity: Robust mean-variance portfolio analysis. *Econometrica* 81:1075–1113.
17. Duffie D, Epstein L (1992) Stochastic differential utility. *Econometrica* 60:353–394.
18. Duffie D, Lions P-L (1992) PDE solutions of stochastic differential utility. *J Math Econ* 21:577–606.
19. Chen Z, Epstein L (2002) Ambiguity, risk, and asset returns in continuous time. *Econometrica* 70:1403–1443.