Macroeconomic Uncertainty Prices*

when Beliefs are Tenuous

Lars Peter Hansen† Thomas J.Sargent‡

December 30, 2018

Abstract

A representative investor confronts two levels of model uncertainty. The investor has a set of well defined parametric "structured models" but does not know which of them is best. The investor also suspects that all of the structured models are misspecified. These uncertainties about probability distributions of risks give rise to components of equilibrium prices that differ from the well understood risk prices widely used in asset pricing theory. A quantitative example highlights a representative investor's uncertainties about the size and persistence of macroeconomic growth rates. Our model puts nonlinearities into marginal valuations that induce time variations in market prices of uncertainty. These arise because the representative investor especially fears high persistence of low growth rate states and low persistence of high growth rate states.

Keywords— Risk, uncertainty, asset prices, relative entropy, Chernoff entropy, robustness, variational preferences; baseline, structured, and unstructured models

JEL Classification— C52, C58, D81, D84, E7, G12

---

*We thank Lloyd Han for his help with estimation, Victor Zhorin for his help calculating Chernoff entropy, Yiran Fan for help in formulating and solving ode's for robust decision problems, and John Wilson for additional research assistance. Fernando Alvarez, Ben Brooks, Vera Chau, Timothy Christensen, Stavros Panageas, Doron Ravid, and Bálint Szőke provided insightful critical comments on earlier drafts. We acknowledge and thank the Alfred P. Sloan Foundation Grant G-2018-11113 for support.

†University of Chicago, E-mail: lhansen@uchicago.edu.

‡New York University, E-mail: thomas.sargent@nyu.edu.

1 Introduction

This paper describes prices of macroeconomic uncertainty that emerge from how investors evaluate utility consequences of alternative specifications of state dynamics. Movements in these uncertainty prices induce variations in asset values. We construct a quantitative example in which uncertainty about macroeconomic growth rates plays a central role. Because they have especially adverse consequences for discounted expected utilities, a representative investor especially fears growth rate persistence in times of weak growth, but fears absence of growth rate persistence when macroeconomic growth is strong.

To construct uncertainty prices, we posit a stand-in investor who has a parametric family of structured models (with either fixed or time-varying parameters) that we represent with a recursive structure suggested by Chen and Epstein (2002) for continuous time models with Brownian motion information flows. Because our decision maker distrusts all of his structured models, he adds unstructured nonparametric models that reside within statistical neighborhoods of them.[1] To include concerns about such unstructured statistical models, we use a model of preferences proposed by Hansen and Sargent (2018) that extends models by Hansen and Sargent (2001) and Hansen et al. (2006) of a decision maker who expresses distrust of a single probability model by surrounding it with an infinite dimensional family of difficult-to-discriminate unstructured models. The Hansen and Sargent (2018) preferences are a continuous-time version of the dynamic variational preferences of Maccheroni et al. (2006).

The representative agent, who in our quantitative example impersonates "the market," is uncertain about prospective macroeconomic growth rates. Shadow prices that characterize aspects of model specifications that most concern a representative investor equal uncertainty prices that clear competitive security markets. Multiplying an endogenously determined vector of worst-case drift distortions by minus one gives a vector of prices that compensate the representative investor for bearing model uncertainty.[2] Time variations in uncertainty prices emerge endogenously because the representative investor's concerns about the persistence of macroeconomic growth rates make uncertainty prices depend on the state of the economy. These findings extend earlier quantitative results that indicated that investor's responses to modest amounts of model ambiguity can substitute for the

[1] By "structured" we mean more or less tightly parameterized statistical models. Thus, for us "structured models" aren't what econometricians working in the tradition either of the Cowles commission or of rational expectations econometrics would call "structural" models.

[2] This object also played a central role in the analysis of Hansen and Sargent (2010).

implausibly large risk aversions during economic downturns that are required to explain observed market prices of risk.

Section 2 specifies an investor's baseline probability model and perturbations to it, both cast in continuous time for analytical convenience. The representative investor uses particular martingales as perturbations to a baseline model to create his set of structured models, and other martingales as perturbations to those structured models to express his suspicion that all of his structured models are misspecified. Section 3 describes discounted relative entropy, a statistical measure of discrepancy between martingales, and uses it to construct a convex set of probability measures that we impute to our decision maker. A martingale representation proves to be a tractable way for us to formulate a robust decision problem in section 4.

Section 5 describes and compares relative entropy and Chernoff entropy, each of which measures statistical divergence from a set of martingales. We show how to use these measures 1) to assess plausibility of worst-case models in the spirit of Good (1952), and 2) to calibrate a penalty parameter that we use to represent the investor's preferences. By extending the approach of Hansen et al. (2008), section 6 calculates key objects in a quantitative version of a baseline model together with worst-case probabilities associated with a convex set of alternative models that concern both a robust investor and a robust planner. Section 7 constructs a recursive representation of a competitive equilibrium of an economy with a representative robust investor. Then it links worst-case probabilities that emerge from a robust planning problem to equilibrium uncertainty compensations that the representative investor receives in competitive markets. Section 8 offers concluding remarks.

2 Martingales and probabilities

Martingales play an important role in a large literature on pricing derivative claims. They also play a different role in this paper. This section describes convenient mathematical representations of nonnegative martingales that can be used to modify a baseline probability model. Following Hansen and Sargent (2018), section 3 constructs a set of martingales that specify a set of structured models of interest to a decision maker. We also describe additional martingales that generate statistically nearby unstructured models that also concern the decision maker because he fears that all of the structured models are misspecified.

For concreteness, we use the following *baseline* model of a stochastic process $X \doteq \{X_t :$

$t \geqslant 0\}$: [3]

$$dX_t = \widehat{\mu}(X_t)dt + \sigma(X_t)dW_t, \tag{1}$$

where W is a multivariate standard Brownian motion.[4] A *plan* is a $\{C_t : t \geqslant 0\}$ that is a progressively measurable process with respect to the filtration $\mathcal{F} = \{\mathcal{F}_t : t \geqslant 0\}$ associated with the Brownian motion W augmented by information available at date zero. The date t component C_t is measurable with respect to $\mathcal{F}_t$.

A decision maker cares about plans. Because he does not fully trust baseline model (1), the decision maker explores utility consequences of other probability models that he obtains by multiplying probabilities associated with (1) by likelihood ratios depicted as positive martingales with unit expectations. Following an extensive literature in probability theory, we represent a likelihood ratio by a positive martingale M^U with respect to the baseline Brownian motion specification

$$dM_t^U = M_t^U U_t \cdot dW_t \tag{2}$$

or

$$d \log M_t^U = U_t \cdot dW_t - \frac{1}{2}|U_t|^2 dt, \tag{3}$$

where U is progressively measurable with respect to the filtration $\mathcal{F}$. In the event that

$$\int_0^t |U_\tau|^2 d\tau < \infty \tag{4}$$

with probability one, the stochastic integral $\int_0^t U_\tau \cdot dW_\tau$ is an appropriate probability limit. Imposing the initial condition $M_0^U = 1$, we express the solution of stochastic differential equation (2) as the stochastic exponential

$$M_t^U = \exp\left(\int_0^t U_\tau \cdot dW_\tau - \frac{1}{2}\int_0^t |U_\tau|^2 d\tau\right). \tag{5}$$

As specified so far, M_t^U is a local martingale, but not necessarily a martingale.[5]

[3] We let X denote a stochastic process, X_t the process at time t, and x a realized value of the process.

[4] Although applications typically use a Markov formulation, this restriction is not essential. Our formulation could be generalized to allow other stochastic processes constructed as functions of a Brownian motion information structure.

[5] It is inconvenient here to impose sufficient conditions for the stochastic exponential to be a martingale like Kazamaki's or Novikov's. Instead, we will verify that an extremum of a pertinent optimization problem does indeed result in a martingale.

Definition 2.1. *$\mathcal{M}$ denotes the set of all martingales M^U constructed as stochastic exponentials via representation* (5) *with a U that satisfies* (4) *and is progressively measurable with respect to $\mathcal{F} = \{\mathcal{F}_t : t \geqslant 0\}$.*

Associated with U are probabilities defined by

$$E^U\left[B_t|\mathcal{F}_0\right] = E\left[M_t^U B_t|\mathcal{F}_0\right]$$

for any $t \geqslant 0$ and any bounded $\mathcal{F}_t$-measurable random variable B_t, so the positive random variable M_t^U acts as a Radon-Nikodym derivative for the date t conditional expectation operator $E^U\left[\,\cdot\,|X_0\right]$.

Under baseline model (1), W is a standard Brownian motion, but under the alternative U model, it has increments

$$dW_t = U_t dt + dW_t^U, \tag{6}$$

where W^U is now a standard Brownian motion. Furthermore, under the M^U probability measure, $\int_0^t |U_\tau|^2 d\tau$ is finite with probability one for each t. While (3) expresses the evolution of $\log M^U$ in terms of increment dW, the evolution in terms of dW^U is

$$d\log M_t^U = U_t \cdot dW_t^U - \frac{1}{2}|U_t|^2 dt. \tag{7}$$

In light of (7), we can write model (1) as:

$$dX_t = \widehat{\mu}(X_t)dt + \sigma(X_t)\cdot U_t dt + \sigma(X_t)dW_t^U.$$

3 Measuring statistical discrepancies

We use a log-likelihood ratio to construct entropy relative to a probability specification affiliated with a martingale M^S defined by a drift distortion process S. Rather than using a log likelihood ratio $\log M_t^U$ with respect to the baseline model, we use a log likelihood ratio $\log M_t^U - \log M_t^S$ with respect to the M_t^S model to arrive at:

$$E\left[M_t^U\left(\log M_t^U - \log M_t^S\right)|\mathcal{F}_0\right] = \frac{1}{2}E\left(\int_0^t M_\tau^U |U_\tau - S_\tau|^2 d\tau \middle| \mathcal{F}_0\right).$$

When the following limits exist, a notion of relative entropy appropriate for stochastic processes is:

$$
\begin{aligned}
\lim_{t\to\infty}\frac{1}{t}E\left[M_t^U\left(\log M_t^U-\log M_t^S\right)\Big|\mathcal{F}_0\right] &= \lim_{t\to\infty}\frac{1}{2t}E\left(\int_0^t M_\tau^U|U_\tau-S_\tau|^2d\tau\Big|\mathcal{F}_0\right)\\
&= \lim_{\delta\downarrow 0}\frac{\delta}{2}E\left(\int_0^\infty \exp(-\delta\tau)M_\tau^U|U_\tau-S_\tau|^2d\tau\Big|\mathcal{F}_0\right).
\end{aligned}
$$

The second line is a limit of Abel integral averages, where scaling by δ makes the weights $\delta\exp(-\delta\tau)$ integrate to one. To assess model misspecification, instead of undiscounted relative entropy, we shall use Abel averages with a discount rate equal to the subjective rate that discounts expected utility flows. With that in mind, we define a discrepancy between two martingales M^U and M^S as:

$$
\Delta\left(M^U;M^S|\mathcal{F}_0\right)=\frac{\delta}{2}\int_0^\infty \exp(-\delta t)E\left(M_t^U \mid U_t-S_t\mid^2\Big|\mathcal{F}_0\right)dt.
$$

We start from a convex set $M^S\in\mathcal{M}^o$ of *structured* models that we represent by martingales with respect to one such baseline model. We use undiscounted relative entropy to constrain this set.

We describe how to form $\mathcal{M}^o$ in subsection 3.1. Structured models in $\mathcal{M}^o$ are alternatives to the baseline model that are of particular interest to a decision maker. For a real number $\theta>0$, define a scaled discrepancy of martingale M^U from a set of martingales $\mathcal{M}^o$ as

$$
\Theta(M^U|\mathcal{F}_0)=\theta\inf_{M^S\in\mathcal{M}^o}\Delta\left(M^U;M^S|\mathcal{F}_0\right). \tag{8}
$$

Scaled discrepancy $\Theta(M^U|\mathcal{F}_0)$ equals zero for M^U in $\mathcal{M}^o$ and is positive for M^U not in $\mathcal{M}^o$. We use discrepancy $\Theta(M^U|\mathcal{F}_0)$ to define a set of unstructured models that are near the set $\mathcal{M}^o$; our decision maker wants to understand utility consequences of these nearby models too. The scaling parameter θ measures how heavily an expected utility maximizing decision maker penalizes an expected utility minimizing agent for distorting probabilities relative to models in $\mathcal{M}^o$.

3.1 A family $\mathcal{M}^o$ of structured models

We construct a family of structured probabilities by forming a set of martingales M^S with respect to a baseline probability associated with model (1). Formally,

$$\mathcal{M}^o = \left\{ M^S \in \mathcal{M} \text{ such that } S_t \in \Xi_t \text{ for all } t \geqslant 0 \right\} \tag{9}$$

where Ξ is a process of convex sets adapted to the filtration $\mathcal{F}$. Chen and Epstein (2002) also used an instant-by-instant constraint like (9) to construct a set of probability models.[6]

We form a set of structured models by restricting relative entropies. The (undiscounted) entropy for a stochastic process M^S relative to the baseline is:

$$\varepsilon(M^S) = \lim_{t \to \infty} \frac{1}{2t} \int_0^t E\left(M^S_\tau |S_\tau|^2 \middle| \mathcal{F}_0 \right) d\tau.$$

Notice that ε is the limit as $t \to +\infty$ of a process of mathematical expectations of time series averages

$$\frac{1}{2t} \int_0^t |S_\tau|^2$$

under the probability measure implied by M^S. Suppose that M^S is defined by the drift distortion $S_t = \eta(Z_t)$, where Z is an ergodic Markov process formed as an invariant function of the state X with autonomous transition probabilities that converge to a unique well-defined stationary distribution Q under the M^S probability. In this case, we can use Q to evaluate relative entropy by computing:

$$\frac{1}{2} \int |\eta|^2 dQ.$$

We represent the instantaneous counterpart to the one-period transition distribution for a Markov process in terms of an infinitesimal generator. A generator tells how conditional expectations of the Markov state evolve locally and can be derived informally by differentiating the family of conditional expectation operators with respect to the gap of elapsed time. For a diffusion, the infinitesimal generator $\mathcal{A}$ of transitions under the M^S

[6] Anderson et al. (1998) also explored consequences of a constraint like (9) but without the state dependence in Ξ. Allowing for state dependence is important in the applications featured in this paper.

probability is the second-order differential operator:

$$\mathcal{A}^s \rho = \frac{\partial \rho}{\partial z} \cdot (\widehat{\mu} + \sigma s) + \frac{1}{2} \text{trace} \left(\sigma' \frac{\partial^2 \rho}{\partial z \partial z'} \sigma \right)$$

for $s = \eta(z)$, where the test function ρ resides in an appropriately defined domain of the generator $\mathcal{A}$. Then

$$\mathcal{A}^s \rho = \frac{\mathsf{q}^2}{2} - \frac{|s|^2}{2}, \tag{10}$$

where relative entropy $\varepsilon(M^S) = \frac{\mathsf{q}^2}{2}$ and q measures the magnitude of the corresponding drift distortion. To compute relative entropy associated with a process defined by generator $\mathcal{A}^s$, we solve equation (10) simultaneously for q and the function ρ. The function ρ is defined only up to translation by a constant and is a long-horizon refinement of relative entropy:

$$\rho(z) - \int \rho dQ = \lim_{t \to \infty} \frac{1}{2} \int_0^t E \left(M_\tau^S |S_\tau|^2 - \mathsf{q}^2 \mid Z_0 = z \right),$$

where Q is the stationary distribution for the probability associated with the $S_t = \eta(Z_t)$ probability model.

Having indicated how to compute relative entropy for a Markov process, we want to use similar calculations to restrict a family of potential structured models in terms of their relative entropies $\varepsilon(M^S)$. Instead of specifying $\frac{\mathsf{q}}{2}$ only, we now also specify ρ *a priori*. For reasons discussed in Hansen and Sargent (2018), restricting q alone is too weak to allow us to get a set expressible as in (9). Therefore, we work with the more restrictive set:

$$\Xi_t = \left\{ s : \mathcal{A}^s \rho(Z_t) \leqslant \frac{\mathsf{q}^2}{2} - \frac{|s|^2}{2} \right\}$$

for a given choice of (q, ρ). The boundary of the set defined in this way includes models with the same long-horizon relative entropy as well as having the same refined approximation to relative entropy: $\rho(z) - \int \rho dQ$. Since we specify ρ only up to a scale, we are not compelled to specify $\int \rho dQ$ *a priori*. The boundary of the set Ξ_t is easy to characterize because it is represented as quadratic function in s given Z_t.

3.2 Misspecification of structured models

Our decision maker wants to evaluate utility consequences of the structured models in $\mathcal{M}^o$ and also of unstructured models that statistically are difficult to distinguish from them.

For that purpose, he employs the scaled statistical discrepancy measure $\Theta(M^U|\mathcal{F}_0)$ defined in (8).[7] The decision maker uses the scaling parameter $\theta < \infty$ and the relative entropy that it implies to calibrate a set of nearby unstructured models. The decision maker solves a minimization problem in which θ serves as a penalty parameter that punishes exploring utility consequences of unstructured probabilities that are statistically too far from the structured models. The minimization problem induces a preference ordering that belongs to a broader class of dynamic variational preferences that Maccheroni et al. (2006) showed are dynamically consistent.

To understand how our formulation relates to dynamic variational preferences, notice how structured models represented in terms of their drift distortion processes S_t appear separately from unstructured models represented in terms of drift distortion U_t on the right side of the statistical discrepancy measure

$$\Delta\left(M^U; M^S|\mathcal{F}_0\right) = \frac{\delta}{2}\int_0^\infty \exp(-\delta t) E\left(M_t^U \mid U_t - S_t \mid^2 \Big| \mathcal{F}_0\right) dt.$$

Specification (8) leads to a conditional discrepancy

$$\xi_t(U_t) = \inf_{S_t \in \Xi_t} |U_t - S_t|^2$$

and an associated scaled integrated discounted discrepancy

$$\Theta\left(M^U|\mathcal{F}_0\right) = \frac{\theta\delta}{2}\int_0^\infty \exp(-\delta t) E\left[M_t^U \xi_t(U_t) \Big| \mathcal{F}_0\right] dt. \tag{11}$$

Our decision maker wants to know utility consequences of statistically close unstructured models that he describes with the discrepancy measure $\Theta\left(M^U|\mathcal{F}_0\right)$. Therefore, he ranks alternative hypothetical state-contingent and date-contingent consumption plans by the minimized value of a discounted sum of expected utilities plus a θ-scaled relative entropy penalty $\Theta(M^U|\mathcal{F}_0)$, where minimization is over the implied set of models.

[7] Watson and Holmes (2016) and Hansen and Marinacci (2016) discuss several misspecification challenges confronted by statisticians and economists.

4 Recursive Representation of Preferences and Decisions

By describing a set of structured models and a continuation value process over consumption plans, this section prepares the way for the quantitative application in section 6. A scalar continuation value stochastic process ranks alternative consumption plans. The date t continuation values tell a decision maker's date t ranking. Continuation value processes have a recursive structure that makes preferences be dynamically consistent. For Markovian decision problems, a Hamilton-Jacobi-Bellman (HJB) equation describes the evolution of continuation values.

4.1 Continuation values

For a consumption plan $\{C_t\}$, the continuation value process $\{V_t\}_{t=0}^{\infty}$ is

$$V_t = \min_{\{U_\tau : t \leqslant \tau < \infty\}} E\left(\int_0^\infty \exp(-\delta\tau) \left(\frac{M_{t+\tau}^U}{M_t^U} \right) \left[\psi(C_{t+\tau}) + \left(\frac{\theta\delta}{2} \right) \xi_{t+\tau}(U_{t+\tau}) \right] d\tau \mid \mathcal{F}_t \right) \tag{12}$$

where ψ is an instantaneous utility function. We will set $\psi = \log$ in computations that follow. Equation (12) builds in a recursive structure that can be expressed as

$$\begin{aligned} V_t = \min_{\{U_\tau : t \leqslant \tau < t+\epsilon\}} \Bigg\{ & E\left[\int_0^\epsilon \exp(-\delta\tau) \left(\frac{M_{t+\tau}^U}{M_t^U} \right) \left[\psi(C_{t+\tau}) + \left(\frac{\theta\delta}{2} \right) \xi_{t+\tau}(U_{t+\tau}) \right] d\tau \mid \mathcal{F}_t \right] \\ & + \exp(-\delta\epsilon) E\left[\left(\frac{M_{t+\epsilon}^U}{M_t^U} \right) V_{t+\epsilon} \mid \mathcal{F}_t \right] \Bigg\} \end{aligned} \tag{13}$$

for $\epsilon > 0$.

Heuristically, we can "differentiate" the right-hand side of (13) with respect to ϵ to obtain an instantaneous counterpart to a Bellman equation. Viewing the continuation value process $\{V_t\}$ as an Ito process, write:

$$dV_t = \nu_t dt + \varsigma_t \cdot dW_t. \tag{14}$$

A local counterpart to (13) is:

$$0 = \min_{U_t} \left[\psi(C_t) - \frac{\theta\delta}{2} \xi_t(U_t) - \delta V_t + U_t \cdot \varsigma_t + \nu_t \right], \tag{15}$$

where U_t is restricted to be $\mathcal{F}_t$ measurable. The term $U_t \cdot \varsigma_t$ comes from an Ito adjustment to the local covariance between $\frac{dM_t^U}{M_t^U}$ and dV_t. It is an adjustment to the drift ν_t of dV_t that is induced by using martingale M^U to change the probability measure. Preferences ranked by continuation value processes V_t are continuous-time counterparts to the dynamic variational preferences of Maccheroni et al. (2006).

4.2 Markovian decision problem

By ranking consumption processes with continuation value processes satisfying (15), a decision maker evaluates utility consequences of unstructured models that our relative entropy measure asserts are difficult to distinguish from members of the set of structured models $\mathcal{M}^o$.

To construct a set of models, the decision maker:

1) Begins with a Markovian baseline model.

2) Creates a set $\mathcal{M}^o$ of *structured* models by naming a sequence of closed convex sets $\{\Xi_t\}$ and associated drift distortion processes $\{S_t\}$ that satisfy structured model constraint (9).

3) Augments $\mathcal{M}^o$ with additional *unstructured* models that violate (9) but according to discrepancy measure (8) are statistically close to models that do satisfy it.

For step 1, the decision maker uses the diffusion (1) as a Markovian baseline model. Step 3 includes statistically similar models that are not necessarily Markovian. We now describe how to implement some of these steps for the quantitative model to be used in section 6. We begin by describing the baseline model used by a key decision maker in that application, a robust planner.

4.2.1 Step 1

We use a single capital version of an Eberly and Wang (2011) model with a long-term risk state z. A robust planner faces an AK model subject to adjustment costs with capital evolution:

$$dK_t = K_t \left(\left[\widehat{\alpha}_k + \widehat{\beta}_k Z_t + \frac{I_t}{K_t} - \phi\left(\frac{I_t}{K_t}\right) \right] dt + \sigma_k \cdot dW_t \right), \tag{16}$$

where ϕ is convex with $\phi(0) = 0$, I_t is investment, and W is a 2×1 Brownian motion. It is convenient to use $\log K$ as the endogenous state variable process, where by Ito's formula it follows that

$$d \log K_t = \left[\widehat{\alpha}_k + \widehat{\beta}_k Z_t + \frac{I_t}{K_t} - \phi\left(\frac{I_t}{K_t}\right) - \frac{|\sigma_k|^2}{2} \right] dt + \sigma_k \cdot dW_t.$$

Consumption is restricted by

$$C_t = \kappa K_t - I_t.$$

The process Z evolves according to

$$dZ_t = \left(\widehat{\alpha}_z - \widehat{\beta}_z Z_t \right) dt + \sigma_z \cdot dW_t,$$

which implies that a stationary distribution for Z is normal with mean $\bar{z} = \widehat{\alpha}_z / \widehat{\beta}_z$ and variance $|\sigma_z|^2/(2\widehat{\beta}_z)$. Let

$$X = \begin{bmatrix} \log K \\ Z \end{bmatrix}$$

and stack the two state equations:

$$\begin{aligned} d \log K_t &= \left[\widehat{\alpha}_k + \widehat{\beta}_k Z_t + \frac{I_t}{K_t} - \phi\left(\frac{I_t}{K_t}\right) - \frac{|\sigma_k|^2}{2} \right] dt + \sigma_k \cdot dW_t \\ dZ_t &= \left(\widehat{\alpha}_z - \widehat{\beta}_z Z_t \right) dt + \sigma_z \cdot dW_t. \end{aligned} \tag{17}$$

4.3 Step 2

A planner chooses I not knowing which of a set of structured model is best while suspecting that all of his structured models are misspecified. We focus on the following collection of structured parametric models:

$$\begin{aligned} d \log K_t &= \left[\alpha_k + \beta_k Z_t + \frac{I_t}{K_t} - \phi\left(\frac{I_t}{K_t}\right) - \frac{|\sigma_k|^2}{2} \right] dt + \sigma_k \cdot dW_t^S \\ dZ_t &= (\alpha_z - \beta_z Z_t) \, dt + \sigma_z \cdot dW_t^S, \end{aligned} \tag{18}$$

where $(\alpha_k, \beta_k, \alpha_z, \beta_z)$ are parameters that distinguish the structured models (18) from the baseline model, (σ_k, σ_z) are parameters common to models (17) and (18), W^S is a 2×1

Brownian motion, and the Brownian motions W and W^S are related by

$$dW_t = S_t dt + dW_t^S, \tag{19}$$

where S_t is the drift distortion implied by parameter values $(\alpha_k, \beta_k, \alpha_z, \beta_z)$. Collection (18) nests the baseline model (17).

We represent members of a parametric class defined by (18) in terms of our section 2 structure with drift distortions S of the form

$$S_t = \eta(Z_t) \equiv \eta_0 + \eta_1(Z_t - \bar{z}),$$

then use (1), (19), and (18) to deduce the following restrictions on η_1:

$$\sigma \eta_1 = \begin{bmatrix} \beta_k - \widehat{\beta}_k \\ \widehat{\beta}_z - \beta_z \end{bmatrix},$$

where

$$\sigma = \begin{bmatrix} (\sigma_k)' \\ (\sigma_z)' \end{bmatrix}.$$

Relative entropy $\frac{\mathsf{q}^2}{2}$ emerges from applying the method of undetermined coefficients to solve differential equation (10)

$$\frac{d\rho}{dz}(z)[-\widehat{\beta}_z(z - \bar{z}) + \sigma_z \cdot \eta(z)] + \frac{|\sigma_z|^2}{2}\frac{d^2\rho}{dz^2}(z) - \frac{\mathsf{q}^2}{2} + \frac{|\eta(z)|^2}{2} = 0, \tag{20}$$

where under the parametric alternatives (18), ρ is quadratic in $z - \bar{z}$:

$$\rho(z) = \rho_1(z - \bar{z}) + \frac{1}{2}\rho_2(z - \bar{z})^2.$$

We compute ρ_1 and ρ_2 by matching coefficients on the terms $(z - \bar{z})$ and $(z - \bar{z})^2$, respectively. Matching constant terms then implies $\frac{\mathsf{q}^2}{2}$.

Suppose that the robust planner's instantaneous utility function is logarithmic. Guess that the value function is then of the additively separable form $\Phi(x) = \log k + \widehat{\Psi}(z)$, where (k, z) are potential realizations of the state vector (K_t, Z_t). If misspecifications of the structured models were not of concern, we would be led to solve the following Hamilton-

Jacobi-Bellman (HJB) equation:

$$
\begin{aligned}
0 = \max_{i} \min_{s} \Big\{ & \delta \log(\kappa - i) - \delta \widehat{\Psi}(z) + \widehat{\alpha}_k + \widehat{\beta}_k z + i - \phi(i) + \sigma_k \cdot s \\
& + [-\widehat{\beta}_z (z - \bar{z}) + \sigma_z \cdot s] \frac{d\widehat{\Psi}}{dz}(z) + \frac{1}{2} |\sigma_z|^2 \frac{d^2 \widehat{\Psi}}{dz^2}(z) \Big\}, \qquad (21)
\end{aligned}
$$

where i is a potential choice of the investment-capital ratio and s is a potential choice of the structured drift distortion. To restrict s, we impose:

$$
[\rho_1 + \rho_2 (z - \bar{z})] \left[-\widehat{\beta}_z (z - \bar{z}) + \sigma_z \cdot s \right] + \frac{|\sigma_z|^2}{2} \rho_2 - \frac{\mathsf{q}^2}{2} + \frac{s \cdot s}{2} \leqslant 0. \qquad (22)
$$

The boundary of the constraint set is an ellipsoid.

By pre-specifying $(\rho_1, \rho_2, \mathsf{q})$, we trace out a one-dimensional family of parametric models with the same relative entropy. For instance, we can solve equation (20) for η_0 and η_1. By matching a constant, a linear term, and a quadratic term in $z - \bar{z}$, we obtain three equations in four unknowns that imply a one dimensional curve for η_0 and η_1 that imply nonlinear S_t's. In this way, nonlinear structured models emerge endogenously. These nonlinear models also have relative entropy $\frac{\mathsf{q}^2}{2}$. We can represent the resulting nonlinear model as a time-varying coefficient model by solving

$$
r^*(z) = \sigma \left[\eta_0 + \eta_1 (z - \bar{z}) \right]
$$

for η_0 and η_1, z by z, along the one-dimensional curve, as shown in Illustration 4.1.

We will feature the following special case in some of our calculations.

Illustration 4.1. *Suppose that*

$$
\eta(z) = \eta_1 (z - \bar{z}),
$$

which focuses structured uncertainty on how drifts for (K, Z) *respond to the state variable* Z*. In this case,* $\rho_1 = 0$ *and*

$$
-\frac{\mathsf{q}^2}{2} + \frac{|\sigma_z|^2}{2} \rho_2 = 0
$$

or equivalently,

$$
\rho_2 = \frac{\mathsf{q}^2}{|\sigma_z|^2}.
$$

Notice that restriction (22) *implies that*

$$
s = 0
$$

when $z = \bar{z}$*. Also given* $|\sigma_z|^2$*, the value of* ρ_2 *is determined by* q*. More generally,* q *and* ρ *cannot be specified independently.*

To connect to a time-varying parameter specification, first construct the convex set of η_1*'s that satisfy*

$$\frac{1}{2}\eta_1 \cdot \eta_1 + \left(\frac{\mathsf{q}^2}{|\sigma_z|^2}\right)\left(-\widehat{\beta}_z + \sigma_z \cdot \eta_1\right) \leqslant 0. \tag{23}$$

Next form the boundary of a convex set Π *of alternative parameter configurations constrained by* (23)

$$\sigma\eta_1 = \begin{bmatrix} \beta_k - \widehat{\beta}_k \\ \widehat{\beta}_z - \beta_z \end{bmatrix}$$

for (β_k, β_z) *associated with alternative choices of* η_1*.*

For a given $\widehat{\Psi}$ and state realization z, the component of the objective for the HJB equation (21) that depends on s is the inner product

$$\begin{bmatrix} 1 & \frac{d\widehat{\Psi}}{dz}(z) \end{bmatrix} \sigma s.$$

It is pedagogically convenient to set $r = \sigma s$. With this transformation, the two distinct entries of r alter each of the two state dynamics equations and the objective component of interest. The first entry, r_1 shifts the log capital evolution equation and the second entry, r_2, shifts the evolution for the exogenous state dynamics as is evident from HJB equation (21). The objective used in this HJB equation remains linear in r with a translation, and this this linearity pushes the minimizing r to an ellipsoid that is the boundary of the convex constraint set for each z. Under parameters for the baseline model that we present in section 6, Figure 1 shows ellipsoids associated with two alternative values of z.

For every feasible choice of r_2, two choices of r_1 satisfy the implied quadratic equation. Provided that $\frac{d\widehat{\Psi}}{dz}(z) > 0$, which is true in our calculations, we take the lower of the two solutions for r_1 because the objective has positive weights on the two entries of r. The minimizing solution occurs at a point on the lower left of the ellipsoid where $\frac{dr_1}{dr_2} = -\frac{d\widehat{\Psi}}{dz}(z)$ and depends on z, as Figure 1 indicates.

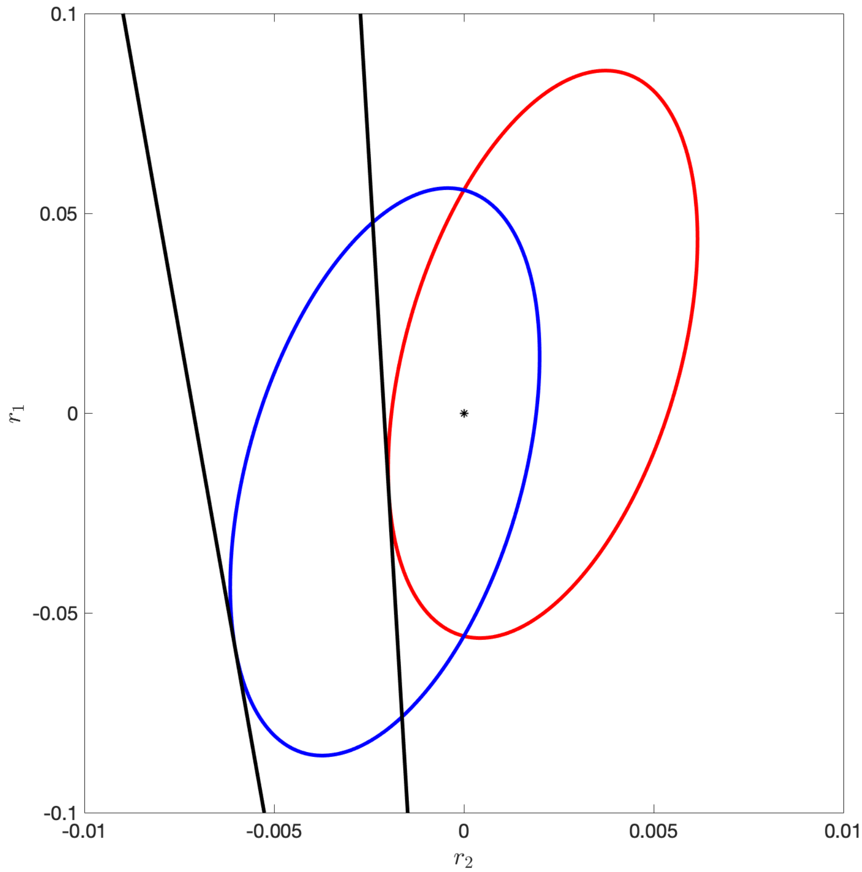


Figure 1: An illustration for section 6 parameter configuration, the figure 3 configuration for $\mathsf{q}_{s,0} = .1$ and $\mathsf{q}_{u,s} = .2$. The figure displays parameter contours for (r_1, r_2), holding relative entropy fixed. The upper right contour depicted in red is for z equal to the .1 quantile of the stationary distribution under the baseline model and the lower left contour is for z at the .9 quantile. The dot depicts the $(r_1, r_2) = (0, 0)$ point corresponding to the baseline model. Tangency points denote worst-case structured models.

4.3.1 Step 3

We now alter the HJB equation in a way that acknowledges the decision maker's fear that all of his structured models are misspecified. He does this by adding unstructured models via a penalized entropy term. This results in the modified version of HJB equation (21):

$$0 = \max_i \min_{u,s} \Big\{ \delta \log(\kappa - i) - \delta \widehat{\Psi}(z) + \widehat{\alpha}_k + \widehat{\beta}_k z + i - \phi(i) + \sigma_k \cdot u \Big]$$

$$+ \left[-\widehat{\beta}_z(z-\bar{z}) + \sigma_z \cdot u\right]\frac{d\widehat{\Psi}}{dz}(z) + \frac{1}{2}|\sigma_z|^2\frac{d^2\widehat{\Psi}}{dz^2}(z) + \frac{\theta}{2}|u-s|^2\Bigg\} \tag{24}$$

where s is constrained by (22). Consider minimization with respect to u first. The first-order conditions imply that

$$u = s + \sigma' \begin{bmatrix} 1 \\ \frac{d\widehat{\Psi}}{dz}(z) \end{bmatrix}.$$

Substituting this choice of u into HJB equation (24) leads us to state the following

Problem 4.2. *Robust planning problem*

$$0 = \max_i \min_s \Bigg\{\delta \log(\kappa - i) - \delta\widehat{\Psi}(z) + \widehat{\alpha}_k + \widehat{\beta}_k z + i - \phi(i) + \sigma_k \cdot s$$
$$+ \left[-\widehat{\kappa}(z-\bar{z}) + \sigma_z \cdot s\right]\frac{d\widehat{\Psi}}{dz}(z) + \frac{1}{2}|\sigma_z|^2\frac{d^2\widehat{\Psi}}{dz^2}(z) - \frac{\theta}{2}\begin{bmatrix} 1 & \frac{d\widehat{\Psi}}{dz}(z) \end{bmatrix}\sigma\sigma'\begin{bmatrix} 1 \\ \frac{d\widehat{\Psi}}{dz}(z) \end{bmatrix}\Bigg\}$$

subject to

$$[\rho_1 + \rho_2(z-\bar{z})]\left[-\widehat{\beta}_z(z-\bar{z}) + \sigma_z \cdot s\right] + \frac{|\sigma_z|^2}{2}\rho_2 - \frac{\mathsf{q}^2}{2} + \frac{s \cdot s}{2} \leqslant 0.$$

Notice that in the HJB equation in Problem 4.2, the objective is additively separable in i and s. This implies that the order of extremization is inconsequential, confirming a Bellman-Isaacs condition. Moreover, for this particular economic environment, the maximizing solution for i, which we denote i^*, is state independent, since the first-order conditions are:

$$\phi'(i) - 1 = \frac{\delta}{\kappa - i}.$$

Thus, the consumption-capital ratio is constant and the logarithms of consumption and capital share the same evolution under the benchmark model, namely,

$$d\log C_t = .01\left[\left(\widehat{\alpha}_c + \widehat{\beta}_c Z_t\right)dt + \sigma_c \cdot dW_t\right]$$

where the .01 scaling is used so that the implied parameters are represented as growth rates,

$$.01\widehat{\alpha}_c = \widehat{\alpha}_k + i^* - \phi(i^*) - \frac{|\sigma_k|^2}{2},$$

$.01\widehat{\beta}_c = \widehat{\beta}_k$, and $.01\sigma_c = \sigma_k$. This model illustrates again a result of Hansen et al. (1999) and

Tallarini (2000) for economies with a single capital stock, namely, that effects of concerns about robustness operate mostly on asset prices, not on allocations.[8]

5 Alternative entropy measures

Preference orderings described in section 4 use the penalty parameter θ in conjunction with relative entropy to limit the unstructured models that the decision maker uses to cope with his concerns that all of the structured models are misspecified. Good (1952) recommended that applications of a max-min expected utility approach should verify that a worst-case model is plausible.[9] We implement Good's suggestion in several ways: we characterize worst-case structured and unstructured models; we also explore how the planner's setting of θ in Problem 4.2 affects the implied relative entropy of the worst-case unstructured model. In calibrating θ in actual decision problems, we find it enlightening to measure the magnitude of the implied worst-case adjustment for misspecifications of the structured models. Also, although we use relative entropy in formulating the decision problems, we find it helpful also to consult another measure of statistical discrepancy called Chernoff entropy.

Let logarithms of two martingales M^S and M^U evolve according to appropriate versions of (7), namely,

$$\begin{aligned} d\log M_t^S &= -\frac{1}{2}|S_t|^2 dt + S_t \cdot dW_t \\ d\log M_t^U &= -\frac{1}{2}|U_t|^2 dt + U_t \cdot dW_t. \end{aligned}$$

Think of a pairwise model selection problem that statistically compares a structured model generated by a martingale M^S with an unstructured model generated by a martingale M^U. For a given value of θ, we compute worst-case structured and unstructured models in terms of the drift distortions

$$\begin{aligned} S_t &= \eta_s(Z_t) \\ U_t &= \eta_u(Z_t) \end{aligned}$$

[8] This result does not occur in environments with multiple capital stocks having different exposures to uncertainty. For a multiple capital stock example with a different specification of model ambiguity, see Hansen et al. (2018)

[9] See Berger (1994) and Chamberlain (2000) for related discussions.

implied for example by the minimization that appears in decision Problem (24).

5.1 Relative entropy

A gauge of divergence between two probability distributions is the following expected log likelihood ratio called relative entropy:

$$\Lambda(M^U, M^S) = \lim_{t \to \infty} \frac{1}{t} E\left[M_t^U \left(\log M_t^U - \log M_t^S\right) | \mathcal{F}_0\right].$$

Since worst-case structured and unstructured probability models are both Markovian, we can compute $\Lambda(M^U, M^S)$ using the same procedures that we applied in section 3.1 to compute entropy relative to the baseline model. In particular, we solve

$$\frac{d\rho}{dz}(z)\left(\widehat{\alpha}_z - \widehat{\beta}_z + \sigma\eta_u\right) + \frac{1}{2}|\sigma_z|^2 \frac{d^2\rho}{dz^2} + \frac{|\eta_u - \eta_s|^2}{2} \leqslant \frac{\mathsf{q}^2}{2}$$

for $\frac{\mathsf{q}^2}{2}$ and for ρ, up to a constant of translation.

Appendix A describes our computational approach. Entropy concept $\Lambda(M^U, M^S)$ is typically independent of date zero conditioning information when the Markov process is asymptotically stationary. In our application, we find it enlightening to report the following transformed object that measures the magnitude of the drift distortion:

$$\mathsf{q}_{u,s} = \sqrt{2\Lambda(M^U, M^S)}.$$

5.2 Chernoff entropy

A dynamic counterpart to an idea of Chernoff (1952) provides an alternative concept of discrepancies between probability measures. Chernoff entropy emerges from studying how, by disguising distortions of a baseline probability model, Brownian motions make it challenging to distinguish models statistically. Although Chernoff entropy's connection to a statistical decision problem makes it interesting, it is less tractable than relative entropy. Anderson et al. (2003) used Chernoff entropy measured as a local rate to make direct connections between magnitudes of market prices of uncertainty, on the one hand, and statistical discrimination between two models, on the other hand. That local rate is state-dependent and for diffusion models is proportional to the local drift in relative entropy. Quantitative differences arise when we measure statistical discrepancy globally as did Newman and

Stuck (1979). We now proceed to characterize a long-run version of Chernoff entropy and show how to compute it.

Think of a pairwise model selection problem that statistically compares a structured model generated by a martingale M^S with an unstructured model generated by a martingale M^U. Consider a statistical model selection rule based on a data history of length t that takes the form of a log likelihood comparison $\log M_t^U - \log M_t^S \geqslant \mathsf{h}$. This selection rule sometimes incorrectly chooses the unstructured model when the structured model governs the data. We can bound the probability of this incorrect selection outcome by using an argument from large deviations theory that starts from

$$\begin{aligned}\mathbf{1}_{\left\{\log M_t^U - \log M_t^S \geqslant \mathsf{h}\right\}} &= \mathbf{1}_{\left\{-\gamma\left(\mathsf{h} + \log M_t^U - \log M_t^S\right) \geqslant 0\right\}} \\ &= \mathbf{1}_{\left\{\exp(-\gamma \mathsf{h})(M_t^U)^\gamma (M_t^S)^{-\gamma} \geqslant 1\right\}} \\ &\leqslant \exp(-\gamma\mathsf{h})(M_t^U)^\gamma (M_t^S)^{-\gamma},\end{aligned}$$

where the inequality holds for $0 \leqslant \gamma \leqslant 1$. Under the structured model, the mathematical expectation of the term on the left side multiplied by M_t^S equals the probability of mistakenly selecting the alternative model when data are a sample of size t generated under the structured model. We bound this mistake probability for large t by following Donsker and Varadhan (1976) and Newman and Stuck (1979) and studying

$$\lim_{t\to\infty} \frac{1}{t} \log E\left[\exp(-\gamma\mathsf{h})\left(M_t^U\right)^\gamma \left(M_t^S\right)^{1-\gamma} | \mathcal{F}_0\right] = \lim_{t\to\infty} \frac{1}{t} \log E\left[\left(M_t^U\right)^\gamma \left(M_t^S\right)^{1-\gamma} | \mathcal{F}_0\right]$$

for alternative choices of γ. We apply these calculations for specifications of U and S, checking that the limits are well defined. The threshold h does not affect the limit. Furthermore, the limit is often independent of the initial conditioning information. To get the best bound, we compute

$$\inf_{0\leqslant\gamma\leqslant 1} \lim_{t\to\infty} \frac{1}{t} \log E\left[\left(M_t^U\right)^\gamma \left(M_t^S\right)^{1-\gamma} | \mathcal{F}_0\right],$$

which is typically negative because mistake probabilities decay with sample size. Chernoff entropy is then

$$\Gamma(M^U, M^S) = -\inf_{0\leqslant\gamma\leqslant 1} \liminf_{t\to\infty} \frac{1}{t} \log E\left[\left(M_t^U\right)^\gamma \left(M_t^S\right)^{1-\gamma} | \mathcal{F}_0\right]. \tag{25}$$

Setting $\Gamma(M^U, M^S) = 0$ would include only alternative models M^U that cannot be distinguished from M^S on the basis of histories of infinite length.[10] Because we want to include more possible alternative models than that, we entertain positive values of $\Gamma(M^U, M^S)$.

To interpret $\Gamma(M^U, M^S)$, note that if the decay rate of mistake probabilities were constant, say d, then mistake probabilities for two sample sizes $T_i, i = 1, 2$, would be

$$\text{mistake probability}_i = \frac{1}{2} \exp\left(-T_i \mathsf{d}_{u,s}\right)$$

for $\mathsf{d}_{u,s} = \Gamma(M^U, M^S)$. We define a half-life as an increase in sample size $T_2 - T_1 > 0$ that multiplies a mistake probability by a factor of one half:

$$\frac{1}{2} = \frac{\text{mistake probability}_2}{\text{mistake probability}_1} = \frac{\exp\left(-T_2 \overline{\chi}\right)}{\exp\left(-T_1 \mathsf{d}\right)},$$

so the half-life is approximately

$$T_2 - T_1 = \frac{\log 2}{\mathsf{d}}. \tag{26}$$

The bound on the decay rate should be interpreted cautiously because the actual decay rate is not constant. Furthermore, the pairwise comparison understates the challenge truly confronting the decision maker, which is statistically to discriminate among *multiple* models.

A symmetrical calculation reverses the roles of the two models and instead conditions on the perturbed model implied by martingale M^U. The limiting rate remains the same. Thus, when we select a model by comparing a log likelihood ratio to a constant threshold, the two types of mistakes share the same asymptotic decay rate.

To implement Chernoff entropy, we follow an approach suggested by Newman and Stuck (1979). Because our worst-case models are Markovian, we use Perron-Frobenius theory to characterize

$$\lim_{t \to \infty} \frac{1}{t} \log E\left[\left(M_t^U\right)^\gamma \left(M_t^S\right)^{1-\gamma} | \mathcal{F}_0\right]$$

for a given $\gamma \in (0, 1)$ as a dominant eigenvalue of a semigroup of linear operators. This limit does not depend on the initial state x and is characterized as a dominant eigenvalue associated with an eigenfunction that is strictly positive.

[10]That is what is done in extensions of the rational expectations equilibrium concept to self-confirming equilibria that allow probability models that are wrong only off equilibrium paths, i.e., for events that in equilibrium do not occur infinitely often. See Fudenberg and Levine (1993, 2009) and Sargent (1999). Our decision theory differs from that used in most of the literature on self-confirming equilibria because our decision maker acknowledges model uncertainty and wants to adjust decisions accordingly. But see Battigalli et al. (2015).

Appendix A describes how we evaluate Chernoff entropy numerically for the nonlinear Markov specifications that we use in subsequent sections.

6 Quantitative example

Our example builds on the physical technology and continuation value process described in section 4 and features a representative investor who wants to explore utility consequences of alternative models portrayed by $\{M_t^U\}$ and $\{M_t^S\}$ processes, some of which contribute a troublesome and difficult to detect predictable components of consumption growth.[11]

6.1 Baseline model

We think of capital in broad terms and base our quantitative application on an empirical calibration of the consumption dynamics. Our example blends elements of Bansal and Yaron (2004) and Hansen et al. (2008). Because we want to focus exclusively on fluctuations in uncertainty prices that are induced by a representative investor's specification concerns, we assume no stochastic volatility, in contrast to Bansal and Yaron (2004). We use a vector autoregression (VAR) to construct a quantitative version of a baseline model like (17) that approximates responses of consumption to permanent shocks. Our VAR follows Hansen et al. (2008) in using various macroeconomic time series to infer information about long-term consumption growth. We deduce a calibration of our baseline model (17) from a trivariate VAR for the first difference of log consumption, the difference between logs of business income and consumption, and the difference between logs of personal dividend income and consumption. This specification makes consumption, business income, and personal dividend income be cointegrated.[12] Since we presume that all three time series grow, we know the coefficients in the cointegrating relation. In Appendix B we tell how

[11] While we appreciate the value of a more comprehensive empirical investigation with multiple macroeconomic time series, here our aim is to illustrate a mechanism within the context of relatively simple time series models of predictable consumption growth.

[12] Business income is measured as proprietor's income plus corporate profits per capita. Dividends are personal dividend income per capita. The time series are quarterly data from 1948 Q1 to 2018 Q3. Our consumption measure is nondurables plus services consumption per capita. The nominal consumption data come from BEA's NIPA Table 1.1.5 and their deflators from BEA's NIPA Table 1.1.4. The business income data with IVA and CCadj are from BEA's NIPA Table 1.12. Personal dividend income data were obtained from from FRED's B703RC1Q027SBEA. Population data comes from FRED's CNP16OV. By including proprietors' income in addition to corporate profits, we use a broader measure of business income than Hansen et al. (2008) who used only corporate profits. Hansen et al. (2008) did not include personal dividends in their VAR analysis.

we obtained discrete time VAR estimates and used these estimates to deduce the following parameters for the baseline model (17):

$$
\begin{aligned}
\widehat{\alpha}_c &= .484 \qquad \widehat{\beta}_c = 1 \\
\widehat{\alpha}_z &= 0 \qquad \widehat{\beta}_z = .014 \\
(\sigma_c)' &= \begin{bmatrix} .477 & 0 \end{bmatrix} \\
(\sigma_z)' &= \begin{bmatrix} .011 & .025 \end{bmatrix}
\end{aligned}
\tag{27}
$$

We suppose that $\delta = .002$. Under this model, the standard deviation of the Z process in the implied stationary distribution is .163.

6.2 Structured models and a robust plan

We solve HJB equation (21) for three different configurations of structured models. We describe our numerical implementation in Appendix C.

6.2.1 Uncertain growth rate responses

We compute a solution by first focusing on an Illustration 4.1 specification in which $\rho_1 = 0$ and ρ_2 satisfies:

$$\rho_2 = \frac{\mathsf{q}^2}{|\sigma_z|^2}.$$

When η is restricted to be $\eta_1(z - \bar{z})$, a given value of q imposes a restriction on η_1 and implicitly on (β_c, β_k). Figure 2 plots iso-entropy contours for (β_c, β_z) for $\mathsf{q} = .1$ and $\mathsf{q} = .05$.

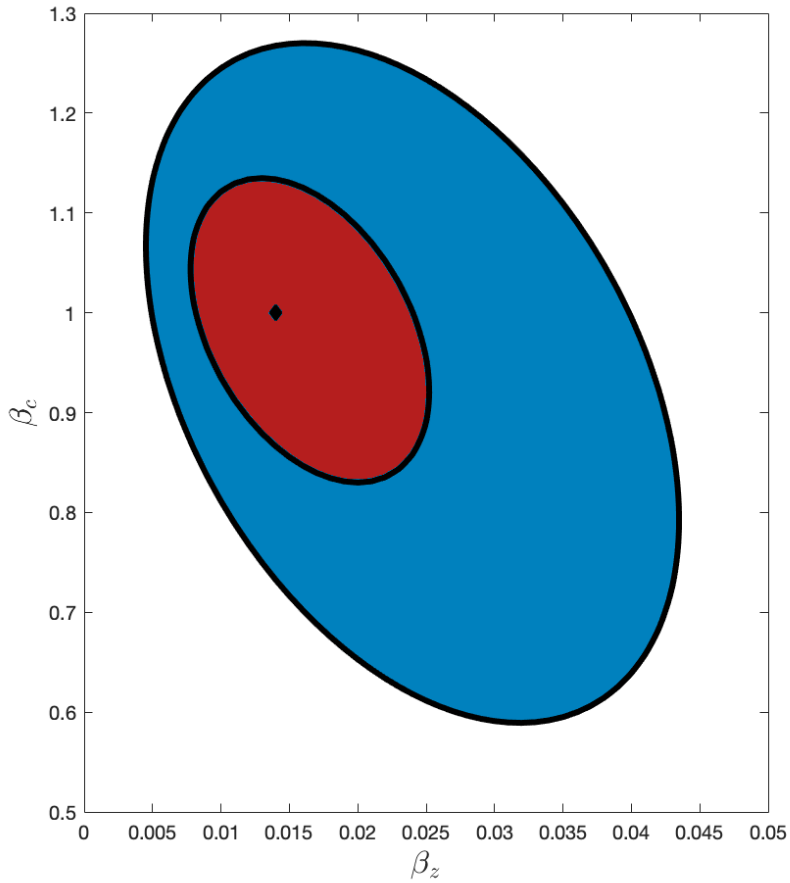


Figure 2: Parameter contours for (β_c, β_k) holding relative entropy fixed. The outer curve depicts $\mathsf{q}_{s,0} = .1$ and the inner curve $\mathsf{q}_{s,0} = .05$. The small diamond depicts the baseline model.

While Figure 2 displays contours of time-invariant parameters with the same relative entropy, the robust planner actually chooses a two-dimensional vector of drift distortions r for a structured model in a more flexible way. As happens when there is uncertainty about (β_c, β_z), sets of possible r's differ depending on the state z. As we remarked earlier, when $z = 0$ the only feasible r is $r = 0$. Figure 1 also reported iso-entropy contours when z is at the .1 and .9 quantile of the stationary distribution under the baseline model. The larger value of z results in a downward shift of the contour relative to the smaller value of z. The points of tangency in Figure 1 are the worst-case structured models. A tangency point occurs at a lower drift distortion for the .9 quantile than for the .1 quantile.

Consider next the adjustment for model misspecification. Since

$$\sigma(u^* - s^*) = -\frac{1}{\theta}\sigma\sigma' \begin{bmatrix} 1 \\ \frac{d\hat{\Psi}}{dz} \end{bmatrix}$$

and entries of $\sigma\sigma'$ are positive, the adjustment for model misspecification is smaller in magnitude for larger values of the state z. Taken together, the vector of drift distortions is:

$$\sigma u^* = \sigma(u^* - s^*) + r^*.$$

The first term on the right is smaller in magnitude for a larger z and conversely, the second term is larger in magnitude for smaller z.

Under the restrictions on structured models now under study, namely, that $\rho_1 = 0$ and $\rho_2 = \frac{\mathsf{q}^2}{|\sigma_z|^2}$ and that η is restricted to be $\eta_1(z - \bar{z})$, the first derivative of the value function is not differentiable at $z = \bar{z}$. We can compute the value function and the worst-case models by solving two coupled HJB equations, one for $z < \bar{z}$ and another for $z > \bar{z}$. We obtain two second-order differential equations in value functions and their derivatives; these value functions coincide at $z = 0$, as do their first derivatives.

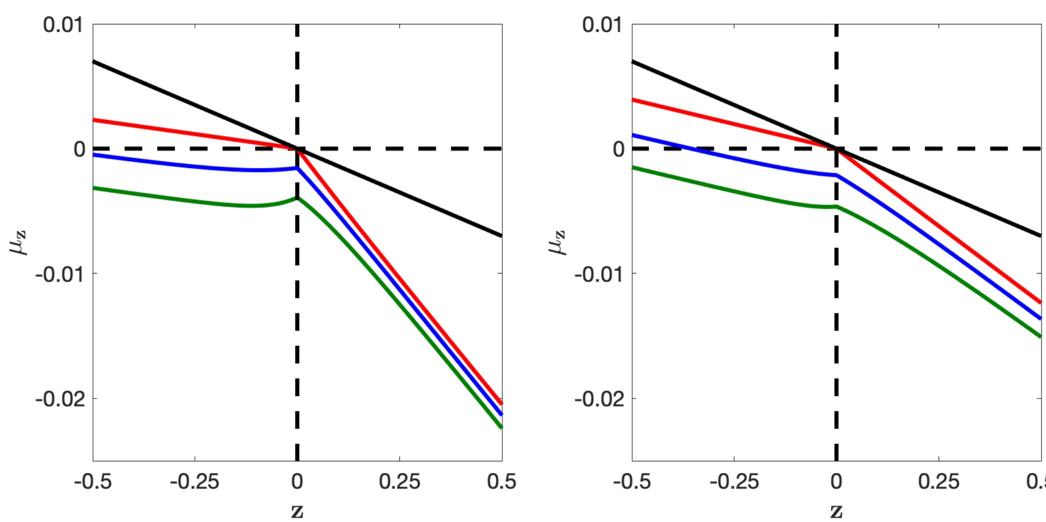


Figure 3: Worst-case structured model growth rate drifts. Left panel: larger structured entropy ($\mathsf{q}_{s,0} = .1$). Right panel: smaller structured entropy ($\mathsf{q}_{s,0} = .05$). The penalty parameter θ set to hit targeted values of $\mathsf{q}_{u,s}$. **Red**: worst-case structured model; **blue**: $\mathsf{q}_{u,s} = .1$; and **green**: $\mathsf{q}_{u,s} = .2$.

Figure 3 shows adjustments of the drifts due to aversion to not knowing which structured model is best and to concerns about misspecifications of the structured models. Setting $\theta = \infty$ silences concerns about misspecification of the structured models, all of which are expressed through minimization over s. When we set $\theta = \infty$, the implied worst-case structured model has state dynamics that take the form of a threshold autoregression with a kink at zero. The distorted drifts again show less persistence than does the baseline model for negative values of z and more persistence for larger values of z. We activate a concern for misspecification of the structured models by setting θ to attain targeted values of $\mathsf{q}_{u,s}$ computed using the structured and unstructured worst-case models. This adjustment shifts the implied worst-case drift as a function of the state downwards, more for negative values of z than for positive ones. The impact of the drift for y is much more modest.

$\mathsf{q}_{s,0}$	$\mathsf{q}_{u,s}$	$\mathsf{d}_{u,s}$	half life u, s	$\mathsf{q}_{u,0}$	$\mathsf{d}_{u,0}$	half life $u, 0$
.10	.10	.0010	668	.33	.0035	198
.10	.20	.0049	142	.62	.0116	60
.05	.10	.0011	631	.19	.0024	289
.05	.20	.0048	144	.36	.0082	84

Table 1: Entropies and half lives. $\frac{1}{2}\mathsf{q}^2$ measures relative entropy and d measures Chernoff entropy. The subscripts denote the probability models used in performing the computations.

Table 1 reports Chernoff and relative entropies implied by structured and unstructured worst-case models. The first two columns tell relative entropy magnitudes that we imposed by adjusting the value of θ. The remaining columns report other measures of entropy as implied by these settings. Recall that the q's measure magnitudes of the drift distortions under associated distorted measures. Thus, $\mathsf{q}_{u,0}$ measures how large the drift distortion is relative to the baseline model. As expected, increasing the targeted values of $\mathsf{q}_{s,0}$ and $\mathsf{q}_{u,s}$ increases the implied values $\mathsf{q}_{u,0}$. There is one curious finding. From table 1, we see that

$$\mathsf{q}_{u,s} + \mathsf{q}_{s,0} < \mathsf{q}_{u,0},$$

which does not satisfy a Triangle Inequality because, while $\mathsf{q}_{u,s}$ and $\mathsf{q}_{u,0}$ are computed under the stationary probability measure implied by the worst-case unstructured model induced by U, $\mathsf{q}_{s,0}$ is computed under the measure implied by worst-case structured model.

Table 1 also reports Chernoff entropies and their implied half lives. These numbers indicate that statistical discrimination is challenging for all four configurations, since even the smallest half-life exceeds 65 quarters. Discrimination is especially challenging when we limit the extent of model misspecification by setting $q_{u,s} = .1$. How are the entropy measures are related? We know no formula that transforms relative entropy into long-run Chernoff entropy, but a formula from by Anderson et al. (2003) is valid locally and leads us to expect that

$$\frac{\mathsf{q}^2}{2} \approx 4\mathsf{d},$$

an approximation that becomes exact when relative drift distortions are constant. This is evidently a good approximation for computed $\mathsf{q}_{u,s}$ and $\mathsf{d}_{u,s}$, but not for $\mathsf{q}_{u,0}$ and $\mathsf{d}_{u,0}$. As we have seen, the composite drift distortions show substantial state dependence via the worst-case structured model.

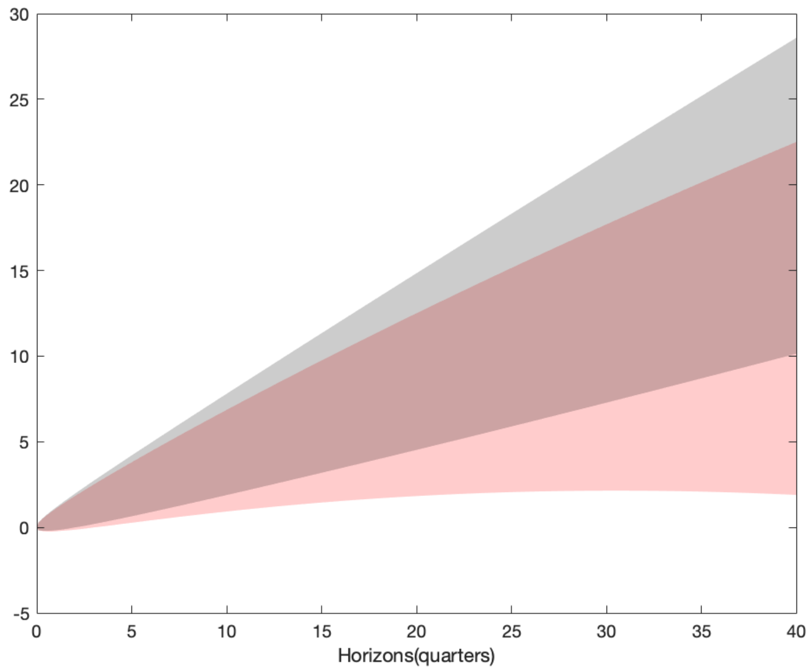


Figure 4: Distribution of $Y_t - Y_0$ under the baseline model and worst-case model for $\mathsf{q}_{s,0} = .1$ and $\mathsf{q}_{u,s} = .2$. The gray shaded area depicts the interval between the .1 and .9 deciles for every choice of the horizon under the baseline model. The red shaded area gives the region within the .1 and .9 deciles under the worst-case model.

Figure 4 extrapolates impacts of the drift distortion on distributions of future consumption growth over alternative horizons. It shows how the consumption growth distribution adjusted for not knowing the best structured model and for distrusting all of the structured models tilts down relative to the baseline distribution.

6.2.2 Altering the scope of uncertainty

Until now, we have imposed

$$\rho_2 = \frac{\mathsf{q}}{|\sigma_z|^2},$$

with the implication that the alternative structured models have no drift distortions for Z at $Z_t = \bar{z}$. We now alter this restriction by cutting the value of ρ_2 in half. Consequences of this change are depicted in the right panel of Figure 5. For sake of comparison, this figure includes the previous specification in the left panel. The worst-case structured drifts no longer coincide with the baseline drift at $z = \bar{z}$ and vary smoothly in the vicinity of $z = \bar{z}$.

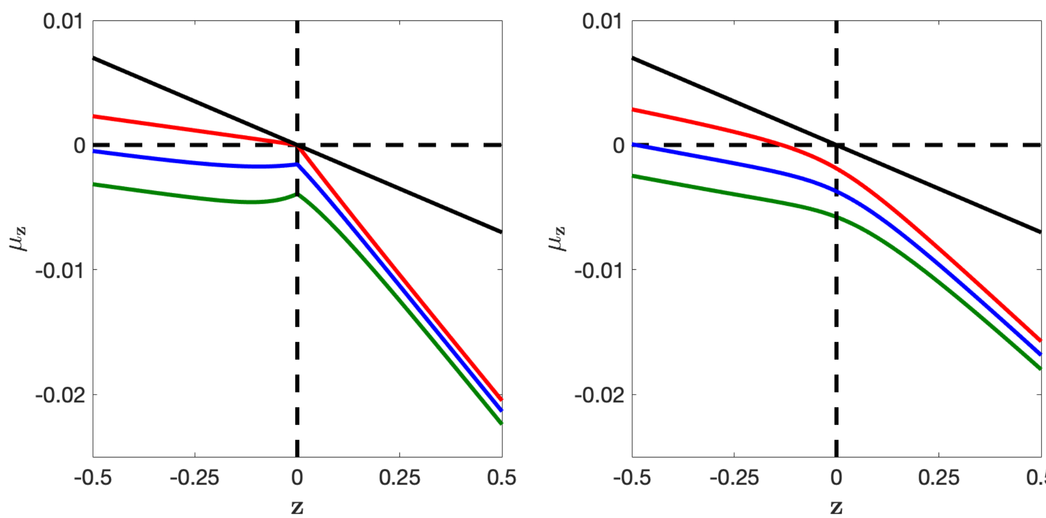


Figure 5: Distorted growth rate drift for Z. Relative entropy $\mathsf{q}_{s,0} = .1$. Left panel: $\rho_2 = \frac{(.01)}{|\sigma_z|^2}$. Right panel: $\rho_2 = \frac{(.01)}{2|\sigma_z|^2}$. **red**: worst-case structured model; **blue**: $\mathsf{q}_{u,s} = .1$; and **green**: $\mathsf{q}_{u,s} = .2$.

Adding the restriction that $\rho_2 = 0$ makes the robust planner's value function become linear and the minimizing s and u become constant and therefore independent of z. Specifically,

$$\frac{d\widehat{\Phi}}{dz} = .01 \frac{\widehat{\beta}}{\delta + \widehat{\beta}_z},$$

and

$$s^* \propto -\sigma' \begin{bmatrix} .01 \\ \frac{.01}{\delta + \widehat{\beta}_z} \end{bmatrix}$$

$$u^* - s^* = -\frac{1}{\theta}\sigma' \begin{bmatrix} .01 \\ \frac{.01}{\delta + \widehat{\beta}_z} \end{bmatrix}$$

The constant of proportionality for s^* is determined by the constraint $|s^*| = \mathsf{q}$. So setting ρ_1 and ρ_2 to zero results in parallel downward shifts of worst-case drifts for both Y and Z. This amounts to changing the coefficients α_y and α_z in ways that are time invariant and leave $\beta_y = \widehat{\beta}_y$ and $\beta_z = \widehat{\beta}_z$.

7 Uncertainty prices

In this section, we construct equilibrium prices that a representative investor receives for bearing ill-understood risks. These equal shadow prices for the robust planning problem of section 4. We decompose equilibrium risk prices into distinct compensations for bearing risk and for bearing model uncertainty. Appendix D describes in detail how we use competitive markets to decentralize implementation of the allocation chosen by a robust planner.[13]

7.1 Local uncertainty prices

The equilibrium stochastic discount factor process Sdf for our robust representative investor economy is

$$d \log Sdf_t = -\delta dt - .01\left(\widehat{\alpha}_c + \widehat{\beta}_c Z_t\right) dt - .01\sigma_y \cdot dW_t + U_t^* \cdot dW_t - \frac{1}{2}|U_t^*|^2 dt. \tag{28}$$

Components of the vector $\omega^*(Z_t) = (.01)\sigma_c - \eta^*(Z_t)$ equal minus the local exposures to the Brownian shocks.[14] While these are usually interpreted as local "risk prices," we shall

[13]We evaluate risk and uncertainty prices relative to the baseline model (1), which we regard as approximating the data well. The planner's and the representative investor's doubts about that model are reflected in the computed compensations.

[14]Please see equation (39) for derivation of this formula for $\omega^*(z)$.

reinterpret them because the decomposition

$$\underbrace{\textbf{minus stochastic discount factor exposure}}_{} = \underbrace{.01\sigma_c}_{\textbf{risk price}} \quad \underbrace{-U_t^*}_{\textbf{uncertainty price}},$$

motivates us to think of $.01\sigma_y$ as risk prices induced by the curvature of log utility and $-U_t^*$ as "uncertainty prices" induced by a representative investor's doubts about the baseline model. Here U_t^* is state dependent. Local prices are large in both good and bad macroeconomic growth states. Prices of uncertainty at longer horizons display more complicated responses to shocks to the macro growth state.

7.2 Uncertainty prices over alternative investment horizons

In the previous subsection, we interpreted $-U_t^*$ as a local price of uncertainty. In this subsection, we provide a corresponding family of conditional expectations:

$$-E\left(M_t^{U^*} U_t^* \mid X_0 = x\right) = \underbrace{-E\left(M_t^{U^*} S_t^* | X_0 = x\right)}_{\textbf{ambiguity price}} - \underbrace{E\left[M_t^{U^*}\left(U_t^* - S_t^*\right) \mid X_0 = x\right)}_{\textbf{misspecification price}}. \tag{29}$$

We interpret the first term on the right-hand side as coming from not knowing the best the structured model and the second term as coming from concerns that all of the structured models might be misspecified. We motivate these measures by constructing "shock price elasticities" for being exposed to future shocks.

The shock elasticities that we construct here are related to but distinct from objects computed by Borovička et al. (2014) because they differ in the timing of the shocks being studied. We study shocks arriving at a different time than do Borovička et al. because we want to answer a different set of questions. In particular, Borovička et al. report elasticities that tell how changing exposures to a shock next period affects the expected return today of an asset that pays off τ periods into the future. In contrast, here we shift the date of the asset's exposure to a shock τ time periods into the future, the same time that the asset pays off. We then study how the expected return as of today varies as we alter the investment horizon $\tau > 0$. We express responses of the expected rate of return as elasticities by normalizing the change in exposure to the shock to be a unit standard deviation and studying effects on logs of expected returns. Shock-price elasticities constructed in this way enlighten us about how state dependence in exposures to future shocks affects returns expected expected as of today of payoffs that materialize across different investment horizons. We shall show that in addition to being intrinsically interesting, elasticities defined in this way link uncertainty prices to relative entropy.

We let consumption be the hypothetical payoff of interest. The logarithm of the expected return from a consumption payoff at date t consists of two terms:

$$\log E\left(\frac{C_t}{C_0}\middle| X_0 = x\right) - \log E\left[Sdf_t\left(\frac{C_t}{C_0}\right)\middle| X_0 = x\right], \tag{30}$$

where $\log C_t = Y_t$. The first term is an expected payoff and the second is the cost of purchasing that payoff. Because our example model imposes a unitary elasticity of substitution via $Sdf_t\left(\frac{C_t}{C_0}\right) = M_t^{U*}$, the second term features a martingale contributed by the representative investor's concern that he does not know which member of his set of structured models is correct and also his concern that all of the structured models are misspecified.

An elasticity tells changes in an expected return that result from a local change in the exposure of consumption to the underlying Brownian motion. Malliavin derivatives are important inputs into calculating a shock-price elasticity. These derivatives measure how a shock at a given date affects consumption and stochastic discount factor processes. The Sdf_t and C_t processes both depend on the same Brownian motion between dates zero and t. We are particularly interested in the consequences now of being exposed to shock at date t. Computing the derivative of the logarithm of the expected return given in (30) results in

$$\frac{E\left[\mathcal{D}_t C_t | \mathcal{F}_0\right]}{E\left[C_t | \mathcal{F}_0\right]} - E\left[\mathcal{D}_t M_t^{U*} | \mathcal{F}_0\right],$$

where $\mathcal{D}_t C_t$ and $\mathcal{D}_t M_t^{U*}$ denote two-dimensional vectors of Malliavin derivatives with respect to the two-dimensional Brownian increment at date t for consumption and the worst-case martingale, respectively.

A formula familiar from other forms of differentiation implies

$$\mathcal{D}_t C_t = C_t\left(\mathcal{D}_t \log C_t\right).$$

The Malliavin derivative of $\log C_t = Y_t$ is the vector $.01\sigma_y$, which is the exposure vector of $\log C_t$ to the Brownian increment dW_t:

$$\mathcal{D}_t C_t = .01 C_t \sigma_c,$$

so

$$\frac{E\left(\mathcal{D}_t C_t | \mathcal{F}_0\right)}{E\left(C_t | \mathcal{F}_0\right)} = .01\sigma_c.$$

Similarly,

$$\mathcal{D}_t M_t^{U*} = U_t^*.$$

Therefore, the term structure of prices that interests us is

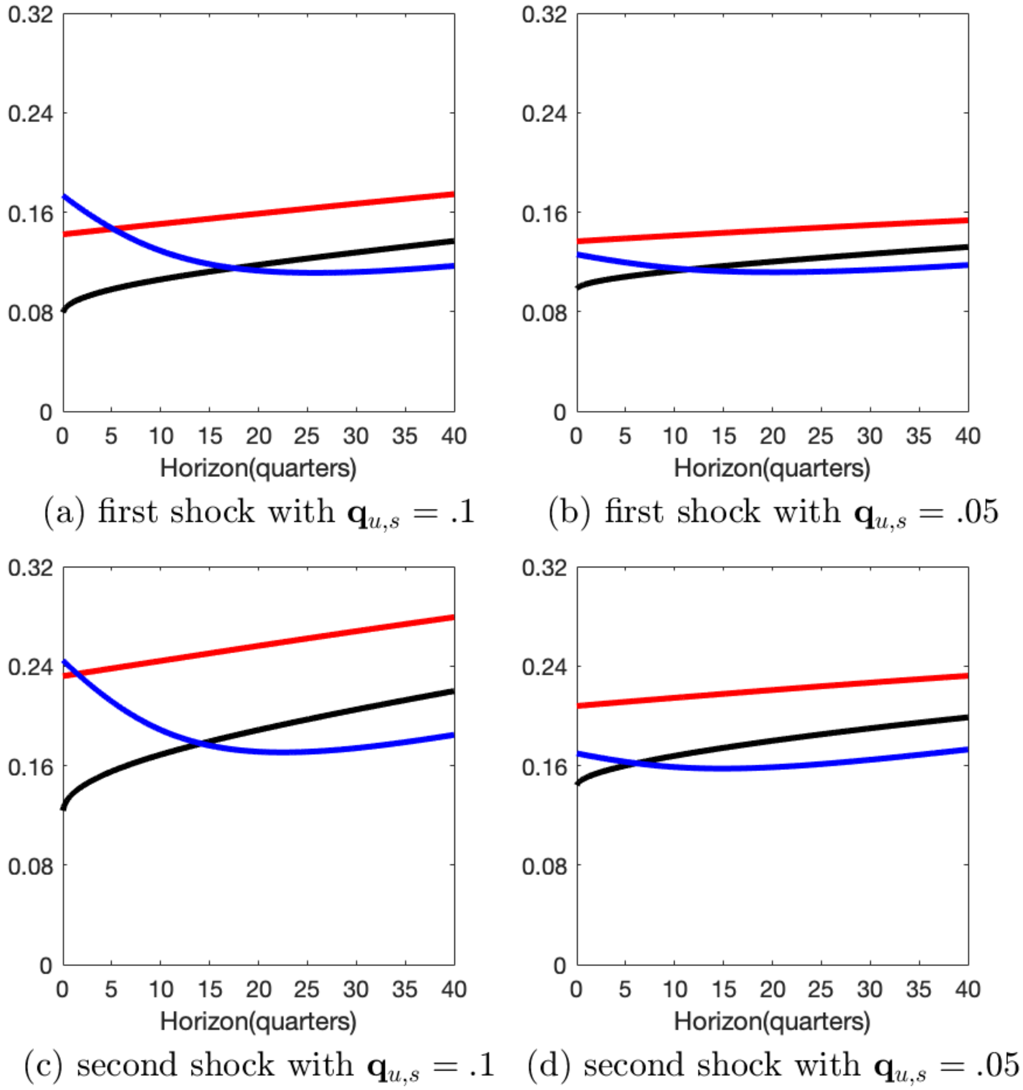
$$.01\sigma_c - E\left(M_t^{U*} U_t^* | \mathcal{F}_0\right). \tag{31}$$

The first term is the risk price familiar from consumption-based asset pricing. It is a (small) state independent-term that is also independent of the horizon. In contrast, the equilibrium drift distortion in the second term contains a state-dependent component, namely, the conditional expectation of the worst-case drift distortion under the distorted probability measure.

Proposition 7.1. *Including contributions from both worst-case structured and unstructured models, horizon-dependent uncertainty prices are:*

$$\upsilon^t(x) \equiv -E\left(M_t^{U*} U_t^* | X_0 = x\right),$$

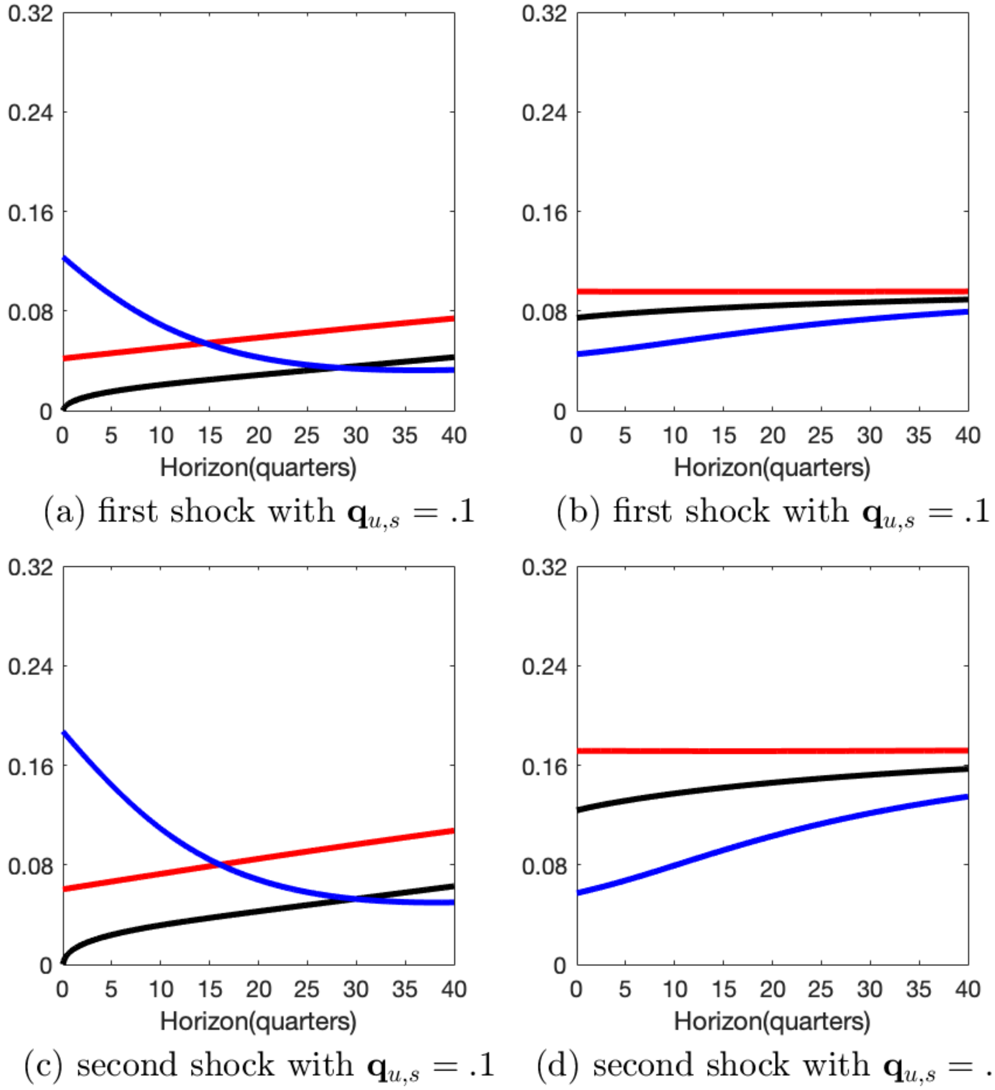
which depend on the horizon t and the initial state x. The limiting uncertainty price vector is the unconditional expectation of the composite drift distortion under the distorted probability distribution.



(a) first shock with $\mathbf{q}_{u,s} = .1$ (b) first shock with $\mathbf{q}_{u,s} = .05$

(c) second shock with $\mathbf{q}_{u,s} = .1$ (d) second shock with $\mathbf{q}_{u,s} = .05$

Figure 6: Shock price elasticities $\upsilon^t(x)$ for alternative horizons. The change in exposure occurs at the same future date as the consumption payoff. The figure reports the median and deciles for the section 6 specification with (β_c, β_z) structured uncertainty. **Black**: median of the Z stationary distribution **red**: .1 decile; and **blue**: .9 decile.



(a) first shock with $\mathbf{q}_{u,s} = .1$ (b) first shock with $\mathbf{q}_{u,s} = .1$

(c) second shock with $\mathbf{q}_{u,s} = .1$ (d) second shock with $\mathbf{q}_{u,s} = .1$

Figure 7: Structured and unstructured contributions to shock price elasticities for alternative horizons. The panels in the left-hand side column plot the ambiguity component in equation (29). The panels in the right-hand side column plot the misspecification component in equation (29). The change in exposure occurs at the same future date as the consumption payoff. The figure reports the median and deciles for the section 6 specification with (β_c, β_z) structured uncertainty. **Black**: median of the Z stationary distribution **red**: .1 decile; and **blue**: .9 decile.

Figure 6 shows shock price elasticities for our section 6 economy, Figure 7 plots separate components of these elasticities given by the right-hand side of equation (7). We feature the case in which $\mathsf{q}_{u,s} = .1$. Notice that although the price elasticity is initially smaller for the median specification of z than for the .9 quantile, this inequality is eventually reversed as the horizon increases. Figure 7 reveals a similar pattern for the instantaneous uncertainty prices: especially for the second shock, instantaneous uncertainty prices are high for the .1 and .9 quantiles of the z distribution relative to the median growth state. Over longer investment horizons, elasticities diminish for the .9 quantiles to magnitudes that are eventually lower than the median elasticities for the same investment horizons. (The blue and black curves cross.) Notice that the misspecification components plotted in Figure 7 are ordered according to quantile, with the lowest quantile have the highest contribution. In contrast, the contribution from ambiguity about the structured models is substantially higher for the .9 quantile than for the other two, with median contributions starting at zero. The misspecification contributions are thus important for understanding both the magnitudes and initial orderings as well as the subsequent reversals of the uncertainty price elasticities. The structured uncertainty components of the elasticities and hence the elasticities themselves diminish with horizon because the probability measure implied by the martingale M_t^{U*} has reduced persistence for positive growth states. Under the M_t^U probability, the growth rate state variable is expected to spend less time in the positive region. This is reflected in smaller ambiguity components of price elasticities at the .9 quantile than at the median over longer investment horizons. For longer investment horizons, but not necessarily for very short ones, an endogenous nonlinearity makes uncertainty prices larger for negative values than for positive values of z. Horizon dependence is an important avenue through which concerns about misspecification and ambiguity aversion influence valuations of assets.

There is an intriguing connection between long-horizon prices and relative entropy. While the uncertainty price trajectories do not converge over the time span reported in Figure 6, well defined limiting uncertainty prices do emerge over longer time horizons.[15] These limits equal $E^{M^{U^*}}(-U_t^*)$, i.e., the unconditional expectation of the corresponding drift distortion vector computed under the worst-case stationary probability measure. In table 2, we compare these limit prices to the relative entropy divergence $\mathsf{q}_{u,0}$, which measures the overall magnitude of these distortions by $\sqrt{2E^{M^{U*}}[|U_t^*|^2]}$, i.e., the square root of twice

[15] Hansen and Scheinkman (2012) study a limiting growth rate risk price that is based on a different conceptual experiment but leads to a similar characterization. Whereas formula (31) has an adjustment for current consumption's exposure to shocks, the limiting Hansen and Scheinkman measure replaces this term by the proportionate exposure of the martingale component of consumption. Both adjustments are small in our quantitative example.

the expected square of the absolute value of the vector of drift distortions, also under worst-case stationary probability measures. Indeed, these mean contributions account for most of the relative entropy measures. This is evident by comparing the square of the number in the third column of table 2 to the sum of the squares in the fourth and fifth columns. Thus, the square root of twice relative entropy provides a measure of the magnitude of long-run uncertainty prices.

$\mathsf{q}_{s,0}$	$\mathsf{q}_{u,s}$	$\mathsf{q}_{u,0}$	shock one price	shock 2 price
.10	.20	.62	.34	.52
.05	.20	.36	.20	.30

Table 2: Entropies and limit prices. $\frac{1}{2}\mathsf{q}^2$ denotes relative entropy. The limiting long-horizon prices are the expectations of $-U^*$ under the probability model implied by U^*.

We have designed our quantitative example to activate a particular mechanism that causes statistically plausible amounts of uncertainty to generate fluctuations in uncertainty prices. We inferred parameters of the baseline model for these examples solely from time series of macroeconomic quantities, thus completely ignoring asset prices during calibration. As a consequence, we do not expect to track high-frequency movements in financial markets well. By limiting our empirical inputs, we respect concerns that Hansen (2007) and Chen et al. (2015) expressed about using asset market data to calibrate macro-finance models that assign a special role to investors' beliefs about future asset prices.[16]

8 Concluding remarks

This paper formulates and applies a tractable model of the effects of macroeconomic uncertainties on equilibrium prices. We quantify investors' concerns about model misspecification in terms of the consequences of alternative statistically plausible models for discounted expected utilities. We characterize the effects of concerns about misspecification of a baseline stochastic process for individual consumption as shadow prices for a planner's problem that supports competitive equilibrium prices.

[16]Hansen (2007) and Chen et al. (2015) describe situations in which it is the behavior of expected rates of return on assets that, through the cross-equation restrictions, lead an econometrician to make inferences about the behavior of macroeconomic quantities like consumption that are much more confident than can be made from the quantity data alone. How could investors put those cross-equation restrictions from returns into quantity processes before *they* had observed returns?

To illustrate our approach, we have focused on the growth rate uncertainty featured in the "long-run risk" literature initiated by Bansal and Yaron (2004). Other applications seem natural. For example, the tools developed here could shed light on a recent public debate between two groups of macroeconomists, one prophesying secular stagnation because of technology growth slowdowns, the other discounting those pessimistic forecasts. The tools that we describe can be used, first, to quantify how challenging it is to infer persistent changes in growth rates, and, second, to guide macroeconomic policy in light of evidence.

Specifically, we have produced a model of a log stochastic discount factor whose uncertainty prices reflect a robust planner's worst-case drift distortions U^* and have shown that these drift distortions can be interpreted as prices of model uncertainty. The dependence of uncertainty prices U^* on the growth state z is shaped partly by alternative parametric models that the decision maker entertains. Thus, our theory of state dependence in uncertainty prices focuses on how our robust investor responds to the presence of the alternative parametric models among a huge set of unspecified alternative models that also concern him.

It is worthwhile comparing this paper's way of inducing time-varying prices of risk with three other macro/finance models that also get them. Campbell and Cochrane (1999) proceed in the standard rational expectations single-known-probability-model tradition and so exclude fears of model misspecification from the mind of their representative investor. Campbell and Cochrane construct a utility function in which the history of consumption expresses an externality. This history dependence makes the investor's local risk aversion depend in a countercyclical way on the economy's growth state. Ang and Piazzesi (2003) use an exponential-quadratic stochastic discount factor in a no-arbitrage statistical model and explore links between the term structure of interest rates and other macroeconomic variables. Their approach allows movements in risk prices to be consistent with historical evidence without specifying an explicit general equilibrium model. A third approach introduces stochastic volatility into the macroeconomy by positing that the volatilities of shocks driving consumption growth are themselves stochastic processes. A stochastic volatility model induces time variation in risk prices via exogenous movements in the conditional volatilities of shocks that impinge on macroeconomic variables. A related approach is implemented by Ulrich (2013) and Ilut and Schneider (2014), who use exogenous stochastic fluctuations in ambiguity concerns to induce additional macroeconomic fluctuations.

In Hansen and Sargent (2010), countercyclical uncertainty prices are driven by a representative investor's robust model averaging. The investor carries along two difficult-to-distinguish models of consumption growth, one with substantial growth rate persistence and the other with little such persistence. The investor uses observations on consumption growth to update a Bayesian posterior over these models and expresses his specification distrust by pessimistically exponentially twisting a posterior over alternative models. That leads the investor to act as if good news is temporary and bad news is persistent, an outcome that is qualitatively similar to what we have found here. Learning occurs in Hansen and Sargent's analysis because the parameterized structured models are time invariant and hence learnable.

In this paper, we propose a different way to make uncertainty prices vary in a way that turns out to be qualitatively similar. We exclude learning by including alternative models with parameters whose future variations cannot be inferred from historical data. These time-varying parameter models differ from the decision maker's baseline model, a fixed parameter model whose parameters can be well estimated from historical data. The alternative models include ones that allow parameters persistently to deviate from those of the baseline model in statistically subtle and time-varying ways. In addition to this parametric class of alternative models, the decision maker also worries about many other specifications. The robust planner's worst-case model responds to these forms of model uncertainty partly by having more persistence in bad states and less persistence in good states. Adverse shifts in a worst-case shock distribution that increase the absolute magnitudes of uncertainty prices were also present in some of our earlier work (for example, see Hansen et al. (1999) and Anderson et al. (2003)). But in this paper, we induce state dependence in uncertainty prices in a new way, namely, by specifying a set of alternative models in a way that captures concerns about the baseline model's specification of persistence in consumption growth.

Our continuous-time formulation (15) exploits mathematically convenient properties of a Brownian information structure. There is a discrete-time version of our formulation that starts from a baseline model cast in terms of a nonlinear stochastic difference equation. In that formulation, there are counterparts to structured and unstructured models that play their roles here. Furthermore, preference orderings defined in terms of continuation values are dynamically consistent.

While our example used entropy measures to restrict the decision maker's set of structured models, two other approaches could be explored instead. One would use a more direct implementation of a robust Bayesian approach; the other would refrain from imposing absolute continuity when constructing a family of structured models.

Turning first to a Bayesian approach, we could start with a set of structured models with time-invariant parameters and a convex set of priors over those parameters. A model-by-model Bayesian approach might be tractable if the implied set of posteriors can be characterized date-by-date and be computed easily, say through the use of conjugate priors. However, after a rectangular augmentation of a set of probabilities of the kind recommended by Epstein and Schneider (2003), the implied worst-case structured model would typically not emerge from applying Bayes' rule to a single prior. That prevents applying Good's advice about assessing the plausibility of max-min choice theory. On the other hand, a rectangular structure may place models on the table that are substantively interesting in their own right, including possibly the worst-case structured model. By incorporating a concern for misspecification, this approach could provide an alternative to the robust learning in Hansen and Sargent (2007).

We turn now to an approach that would abandon the absolute continuity that we have built in when we assumed that the structured model probabilities can be represented as martingales with respect to a baseline model. The approach invented by Peng (2004) uses a theory of stochastic differential equations under a broad notion of model ambiguity that is rich enough to allow for uncertainty about the conditional volatility of Brownian increments. Alternative probability specifications there fail to be absolutely continuous and standard likelihood ratio analysis ceases to apply. If we could construct bounds on uncertainty under a non-degenerate rectangular embedding, we could extend the construction of worst-case structured models and still restrain relative entropy as a way to limit the decision maker's set of unstructured models.[17]

[17]See Epstein and Ji (2014) for an application of the Peng analysis to asset pricing that does not use relative entropy.

Appendices

A Computing Chernoff and relative entropy

We show how to compute Chernoff and relative entropies for Markov specifications where the associated S's and U's take the forms

$$U_t = \eta_u(Z_t)$$
$$S_t = \eta_s(Z_t).$$

A.1 Chernoff entropy

The Markov structures of both models allows us to compute Chernoff entropy by using an eigenvalue approach of Donsker and Varadhan (1976) and Newman and Stuck (1979). We start by computing the drift of $\left(M_t^U\right)^\gamma \left(M_t^S\right)^{1-\gamma} g(Z_t)$ for $0 \leqslant \gamma \leqslant 1$ at $t = 0$:

$$[\mathbb{G}(\gamma)g](z) \doteq -\frac{\gamma(1-\gamma)}{2}|\eta_u(z) - \eta_s(z)|^2 g(z) + g(z)'\sigma \cdot [\gamma\eta_u(z) + (1-\gamma)\eta_s(z)] + g'(z)\left(\widehat{\alpha}_z - \widehat{\kappa} z\right) + \frac{g''(z)}{2}|\sigma_z|^2,$$

where $[\mathbb{G}(\gamma)g](x)$ is the drift given that $Z_0 = z$. Next we solve the eigenvalue problem

$$[\mathbb{G}(\gamma)]e(z,\gamma) = -\lambda(\gamma)e(z,\gamma).$$

We seek the eigenvalue for which $\exp[-\lambda(\gamma)]$ is largest in magnitude; the associated eigenfunction is positive.

We compute Chernoff entropy by solving

$$\Gamma(M^H, M^S) = \max_{\gamma\in[0,1]} \lambda(\gamma),$$

where we compute $\lambda(\gamma)$ numerically using a finite-difference approach. For a pre-specified γ, we evaluate $[\mathbb{G}(\gamma)]g$ at each of n grid points and replace derivatives by two-sided symmetric differences except at the edges, where we use corresponding one-sided differences. This procedure yields a linear transformation of g evaluated at the n grid points. The outcome of this calculation is an n by n matrix applied to a vector containing the entries of g

evaluated at the n grid points. The eigenvalue of the resulting matrix that has the largest exponential equals $-\eta(\gamma)$. We we use a grid for z over the interval $[-2.5, 2.5]$ with grid increments equal to .01, choices that imply that $n = 501$.

A.2 Relative entropy

We solve

$$\frac{\mathsf{q}^2}{2} - \frac{d\rho}{dz}(z)[\widehat{\alpha}_z - \widehat{\beta}_z z + \sigma_z \cdot \eta_u(z)] - \frac{|\sigma_z|^2}{2}\frac{d^2\rho}{dz^2}(z) = \frac{|\eta_u(z) - \eta_s(z)|^2}{2} \tag{32}$$

for q numerically using a finite difference approach like that described in section A.1. Notice that the left-hand side of (32) is linear in $\left(\rho, \frac{\mathsf{q}^2}{2}\right)$. We evaluate equation (32) at the n grid points for z and use a finite difference approximation for the derivatives. We write the resulting left-hand side equations as a matrix times a vector containing $\frac{\mathsf{q}^2}{2}$ and ρ evaluated at $n-1$ grid points omitting $z = 0$ because we set $\rho(0) = 0$ for convenience. We write the right-hand side as a vector evaluated at the n grid points and solve the resulting equation system via matrix inversion.

B Statistical calibration

We fit a trivariate VAR with the following variables:

$$Y_{t+1} = \begin{bmatrix} \log C_{t+1} - \log C_t \\ \log G_{t+1} - \log C_{t+1} \\ \log D_{t+1} - \log C_{t+1} \end{bmatrix}$$

where C_t is consumption, G_t is the sum of corporate profits and proprietors' income, and D_t is personal income.

Provided that the VAR has stable coefficients, this is a co-integrated system. All three time series have stationary increments, but there is one common martingale process. The innovation to this martingale process is identified as the only shock having long-term consequences.

B.1 Calibrating the approximating model

We set $\widehat{\alpha}_z = 0$ and $\widehat{\beta}_c = 1$. For the remaining parameters we:

i) Fit a VAR with a constant and four lags of C_t and five lags of G_t and D_t. Write this as:

$$X_{t+1} = H + AX_t + BW_{t+1}$$

where BB' is the one-period covariance matrix and write $Y_{t+1} = JX_{t+1}$ for a pre-specified matrix J. Identify B by assuming that the square matrix JB is lower triangular.

ii) Compute the implied mean, μ, and the covariance matrix Σ of the stationary distribution for X by solving:

$$\mu = H + A\mu$$
$$\Sigma = A\Sigma A' + BB'$$

where the the second equation can be solved by a doubling algorithm.

iii) Compute the implied mean for $\log C_{t+1} - \log C_t = \mathsf{u}'\mu$ and set it to $.01\widehat{\alpha}_c$; here u' selects the consumption growth rate from the vector X_{t+1}.

iv) Compute the state dependent component of the expected long-term growth rate by evaluating:

$$Z_t^p = \lim_{j\to\infty} E\left(\log C_{t+j} - \log C_t - j\widehat{\alpha}_y | \mathcal{F}_t\right) = \mathsf{u}'(I-A)^{-1}\left[X_t - (I-A)^{-1}H\right]$$

implied by the VAR estimates. This is the counterpart to it from the long-run risk specification:

$$Z_t^p = \lim_{j\to\infty} E\left(\log C_{t+j} - \log C_t - j\widehat{\alpha}_y | Z_t\right) = \frac{.01}{\widehat{\beta}_z} Z_t$$

in the continuous-time model.

v) compute the implied autoregressive coefficient for $\log C_t^p$ in the discrete-time specification using the VAR parameter estimates and equate it to $1 - \widehat{\beta}_z$:

$$1 - \widehat{\beta}_z = \frac{\mathsf{u}'A(I-A)^{-1}A\Sigma(I-A')^{-1}A'\mathsf{u}}{\mathsf{u}'A(I-A)^{-1}\Sigma(I-A')^{-1}A'\mathsf{u}}.$$

vi) Compute the VAR implied covariance matrix for the one-step-ahead forecast error for $\log C^p$ and form the covariance matrix for the growth rate process for consumption and for Z_{t+1}^p.

$$\begin{bmatrix} \mathsf{u}'BB'\mathsf{u} & \mathsf{u}'BB'(I-A')^{-1}A'\mathsf{u} \\ \mathsf{u}'A(I-A)^{-1}BB'\mathsf{u} & \mathsf{u}'A(I-A)^{-1}BB'(I-A')^{-1}A'\mathsf{u} \end{bmatrix} = .0001 \begin{bmatrix} (\sigma_c)' \\ \frac{1}{\hat{\beta}_z}(\sigma_z)' \end{bmatrix} \begin{bmatrix} (\sigma_c) & \frac{1}{\hat{\beta}_z}(\sigma_z) \end{bmatrix};$$

we achieve identification of σ_z and σ_c by imposing a zero restriction on the second entry of σ_c and positive signs on the first coefficient of σ_c and on the second coefficient of σ_z.

B.2 Estimation and inference

Consider the stable VAR:

$$X_{t+1} = H + AX_t + BW_{t+1}$$

where W_{t+1} is a multivariate standard normal and data are available for $X_0, X_1, ..., X_N$. We use importance sampling to construct medians and deciles for the parameters of interest by using formulas in the previous subsection.

i) Construct a "posterior" for the coefficients of the VAR using special case of a method described by Zha (1999). Following Zha, we exploit the lower triangularity of JB by first transforming the equation system to make the implied population residuals are not correlated. We impose "convenient" conjugate priors on the transformed system and initialize them at a "non-informative" prior. This method conditions on X_0. We use Monte Carlo simulation to produce a sequence $\{\theta_j := 1, 2, ..., N\}$ where N is the sample size of the simulated data. We usse this simulation to form an "empirical distribution" with $\frac{1}{N}$ probability assigned to each θ_j rejecting all draws that do not imply a stationary VAR.

ii) Let $f(\cdot \mid \mu, \Sigma)$ be the multivariate normal density and assign weight

$$\frac{f(X_0 \mid \mu_j, \Sigma_j)}{\sum_{j=1}^{N} f(X_0 \mid \mu_j, \Sigma_j)}$$

to outcome θ_j where μ_j and Σ_j are the mean vector and covariance matrix for the stationary distribution implied by θ_j. This adjusts the empirical distribution based on the contribution to the likelihood function from the initial state X_0. Construct medians and deciles from this discrete distribution.

In our computations, we set $N = 10,000,000$. The resulting medians and .1 and .9 deciles are:

Parameter	10^{th} percentile	50^{th} percentile	90^{th} percentile
α_c	.321	.484	.646
β_z	.005	.014	.037
σ_c^1	.452	.477	.501
σ_z^1	.003	.011	.029
σ_z^2	.013	.025	.039

We used medians in computations underlying figures and tables in the text.

C Solving the ODE's

For large $|z|$, the value function is approximately linear in the state variable. This gives a good Neumann boundary condition to use in an approximation that restricts z to be in a compact interval that includes $z = \bar{z}$. Recall the constraint:

$$\frac{1}{2} r' \Lambda r + [\rho_1 + \rho_2 (z - \bar{z})] \left[-\widehat{\beta}_z (z - \bar{z}) + r_2 \right] + \frac{|\sigma_z|^2}{2} \rho_2 - \frac{\mathsf{q}^2}{2} \leqslant 0.$$

Consider an affine solution $r = r_0 + r_1 (z - \bar{z})$. The vector r_1 satisfies

$$\frac{1}{2} (r_1)' \Lambda r_1 - \rho_2 \widehat{\beta}_z + \rho_2 r_{1,2} = 0 \tag{33}$$

where $r_1 = (r_{1,1}, r_{1,2})'$. When we view (33) as a quadratic equation in $r_{1,1}$ given $r_{1,2}$, there are two solutions. We pick the solution that makes $r_{1,1}(z - \bar{z})$ the smallest; the solution depends on whether we use a left boundary point $z^- << \bar{z}$ or a right boundary point $z^+ >> \bar{z}$.

It remains to pick boundary conditions ψ^- and ψ^+ for the derivative of the value function. From the HJB equation:

$$\begin{aligned} (-\delta - \widehat{\beta}_z + r_{1,2}) \psi + .01 (\widehat{\beta}_k + r_{1,1}) = 0 \\ \Lambda r_1 + \begin{bmatrix} 0 \\ \rho_2 \end{bmatrix} \propto \begin{bmatrix} .01 \\ \psi \end{bmatrix}. \end{aligned} \tag{34}$$

The first equation in equation (34) is the derivative of the value function for constant coefficients, putting minimization aside. The second block in (34) consists of two equations derived as the large z approximation to the first-order conditions implied by (24). By taking ratios of these two latter equations we can cancel out the constant of proportionality (the multiplier on the constraint) leaving us with one equation that emerges from the second block. Combining equation (33) and the two equations that emerge from (34), we are left with three equations that determine $(r_{1.1}^{-}, r_{1,2}^{-}, \psi^{-})$ and $(r_{1,1}^{+}, r_{1,2}^{+}, \psi^{+})$, where ψ^{-} and ψ^{+} are the two approximate boundary conditions for the derivative of the value function. We used `bvp4c` in Matlab to solve the ode's over the two intervals $[-2.5, 0]$ and $[0, 2.5]$ where $\overline{z} = 0$.

D Decentralization

D.1 Robust investor portfolio problem

A representative investor solves a continuous-time Merton portfolio problem in which individual wealth A evolves as

$$dA_t = -C_t dt + A_t \iota(Z_t) dt + A_t F_t \cdot dW_t + A_t \omega(Z_t) \cdot D_t dt, \tag{35}$$

where $F_t = f$ is a vector of chosen risk exposures, $\iota(z)$ is an instantaneous risk-free rate, and $\omega(z)$ is a vector of risk prices evaluated at state $Z_t = z$. Initial wealth is A_0. The investor discounts the logarithm of consumption and distrusts his probability model.

Key inputs to a representative investor's robust portfolio problem are the baseline model (1), the wealth evolution equation (35), the vector of risk prices $\omega(z)$, and the quadratic function ρ and relative entropy $\frac{\mathsf{q}^2}{2}$ that define alternative structured models.

Under a guess that the value function takes the form $\widetilde{\Psi}(z) + \log a + \log \delta$, the HJB equation for the robust portfolio allocation problem is

$$\begin{aligned} 0 = \max_{c,f} \min_{u,s} & -\delta \widetilde{\Psi}(z) - \delta \log a - \delta \log \delta + \delta \log c - \frac{c}{k} + \iota(z) \\ & + \omega(z) \cdot f + f \cdot u - \frac{|f|^2}{2} + \frac{d\widetilde{\Psi}}{dz}(z) \left[-\widehat{\beta}_z (z - \bar{z}) + \sigma_z \cdot u \right] \\ & + \frac{1}{2} |\sigma_z|^2 \frac{d^2 \widetilde{\Psi}}{dz^2}(z) + \frac{\theta}{2} |u - s|^2 \end{aligned} \tag{36}$$

where extremization is subject to

$$\frac{|s|^2}{2} + \frac{d\rho}{dz}(z)[-\widehat{\beta}_z(z-\bar{z}) + \sigma_z \cdot s] + \frac{|\sigma_z|^2}{2}\frac{d^2\rho}{dz^2}(z) - \frac{\mathsf{q}^2}{2} = 0. \tag{37}$$

First-order conditions for consumption are

$$\frac{\delta}{c^*} = \frac{1}{a},$$

which imply that $c^* = \delta a$, an implication that follows from the unitary elasticity of intertemporal substitution associated with the logarithmic instantaneous utility function. First-order conditions for a and u are

$$\omega(z) + u^* - f^* = 0 \tag{38a}$$

$$f^* + \theta(u^* - s^*) + \frac{d\widetilde{\Psi}}{dz}(z)\sigma_z = 0. \tag{38b}$$

These two equations determine a^* and $u^* - s^*$ as functions of $\omega(z)$ and the value function $\widetilde{\Psi}$. We determine s^* as a function of u^* by solving

$$\min_s \frac{\theta}{2}|u-s|^2$$

subject to (37). Taken together, these determine (f^*, u^*, s^*). We can appeal to arguments like those of Hansen and Sargent (2008, ch. 7) to justify stacking first-order conditions as a way to collect equilibrium conditions for the two-person zero-sum game that the robust portfolio problem solves.[18]

D.2 Competitive equilibrium prices

We show that the drift distortion η^* that emerges from a robust planner's problem determines prices that a competitive equilibrium awards for bearing model uncertainty. In particular, we compute a vector $\omega(x)$ of competitive equilibrium risk prices by finding a robust planner's marginal valuations of exposures to the W shocks. We decompose that price vector into separate compensations for bearing *risk* and for accepting *model uncertainty.*

[18] An alternative timing protocol that allows the maximizing player to take account of the impact of its decisions on the minimizing agent implies the same equilibrium decision rules described in the text. See Hansen and Sargent (2008, ch. 5).

We verify that the plan for $\log C$ that emerges from the robust planner's problem coincides with the plan for consumption that solves the portfolio problem of a robust investor who takes those prices as given.

Noting from the robust planning problem that the shock exposure vectors for $\log A$ and $\log C$ must coincide implies

$$f^* = (.01)\sigma_c x.$$

From (38b) and the solution for s^*

$$u^* = \eta^*(z),$$

where η^* can be shown to be the worst-case drift from the robust planning problem if we can show that $\widetilde{\Psi} = \widehat{\Psi}$, where $\widehat{\Psi}$ is the value function for the robust planning problem. Thus, from (38a), $\omega = \omega^*$, where

$$\omega^*(z) = (.01)\sigma_c - \eta^*(z). \tag{39}$$

Similarly, in the problem faced by a representative investor within a competitive equilibrium, the drifts for $\log A$ and $\log C$ coincide:

$$-\delta + \iota(z) + [(.01)\sigma_c - \eta^*(z)] \cdot a^* - \frac{.0001}{2}\sigma_c \cdot \sigma_c = (.01)(\widehat{\alpha}_c + \widehat{\beta}_c z),$$

so that $\iota = \iota^*$, where

$$\iota^*(z) = \delta + .01(\widehat{\alpha}_c + \widehat{\beta}_c z) + .01\sigma_y \cdot \eta^*(z) - \frac{.0001}{2}\sigma_c \cdot \sigma_c. \tag{40}$$

By setting $\widetilde{\Psi} = \widehat{\Psi}$, we use these formulas for equilibrium prices to construct a solution to the HJB equation of a representative investor in a competitive equilibrium.

References

Anderson, Evan W., Lars P. Hansen, and Thomas J. Sargent. 1998. Risk and Robustness in Equilibrium. Available on webpages.

———. 2003. A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection. *Journal of the European Economic Association* 1 (1):68–123.

Ang, Andrew and Monika Piazzesi. 2003. A No-Arbitrage Vector Autoregression of the Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics* 50:745–787.

Bansal, Ravi and Amir Yaron. 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance* 59 (4):1481–1509.

Battigalli, Pierpaolo, Simone Cerreia-Vioglio, Fabio Maccheroni, and Massimo Marinacci. 2015. Self-Confirming Equilibrium and Model Uncertainty. *American Economic Review* 105 (2):646–677.

Berger, James O. 1994. An Overview of Robust Bayesian Analysis (with discussion). *Test* 3 (1):5–124.

Borovička, Jaroslav, Lars Peter Hansen, and José A. Scheinkman. 2014. Shock Elasticities and Impulse Response Functions. *Mathematics and Financial Economics* 8 (4):333–354.

Campbell, John Y. and John Cochrane. 1999. Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy* 107 (2):205–251.

Chamberlain, Gary. 2000. Econometric Applications of Maxmin Expected Utility. *Journal of Applied Econometrics* 15 (6):625–644.

Chen, Hui, Winston Wei Dou, and Leonid Kogan. 2015. Measuring the 'Dark Matter' in Asset Pricing Models. MIT Sloan School of Management.

Chen, Zengjing and Larry Epstein. 2002. Ambiguity, Risk, and Asset Returns in Continuous Time. *Econometrica* 70:1403–1443.

Chernoff, Herman. 1952. A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations. *Annals of Mathematical Statistics* 23 (4):pp. 493–507.

Donsker, Monroe E. and S. R. Srinivasa Varadhan. 1976. On the Principal Eigenvalue of Second-Order Elliptic Differential Equations. *Communications in Pure and Applied Mathematics* 29:595–621.

Eberly, Janice C and Neng Wang. 2011. Reallocating and Pricing Illiquid Capital: Two Productive Trees.

Epstein, Larry G. and Shaolin Ji. 2014. Ambiguous Volatility, Possibility and Utility in Continuous Time. *Journal of Mathematical Economics* 50:269 – 282.

Epstein, Larry G. and Martin Schneider. 2003. Recursive Multiple-Priors. *Journal of Economic Theory* 113 (1):1–31.

Fudenberg, Drew and David K. Levine. 1993. Self-Confirming Equilibrium. *Econometrica* 61:523–46.

———. 2009. Self-confirming Equilibrium and the Lucas Critique. *Journal of Economic Theory* 144 (6):2354–2371.

Good, Irving J. 1952. Rational Decisions. *Journal of the Royal Statistical Society. Series B (Methodological)* 14 (1):pp. 107–114.

Hansen, Lars Peter. 2007. Beliefs, Doubts and Learning: Valuing Macroeconomic Risk. *American Economic Review* 97 (2):1–30.

Hansen, Lars Peter and Massimo Marinacci. 2016. Ambiguity Aversion and Model Misspecification: An Economic Perspective. *Statistical Science* 31 (4):511–515.

Hansen, Lars Peter and Thomas J. Sargent. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91 (2):60–66.

———. 2007. Recursive Robust Estimation and Control Without Commitment. *Journal of Economic Theory* 136 (1):1 – 27.

———. 2008. *Robustness*. Princeton, New Jersey: Princeton University Press.

———. 2010. Fragile Beliefs and the Price of Uncertainty. *Quantitative Economics* 1 (1):129–162.

———. 2018. Structured Uncertainty and Model Misspecification. Unpublished working paper.

Hansen, Lars Peter and José A. Scheinkman. 2012. Pricing Growth-rate Risk. *Finance and Stochastics* 16 (1):1–15.

Hansen, Lars Peter, Thomas J. Sargent, and Thomas D. Tallarini. 1999. Robust Permanent Income and Pricing. *The Review of Economic Studies* 66 (4):873–907.

Hansen, Lars Peter, Thomas J. Sargent, Gauhar A. Turmuhambetova, and Noah Williams. 2006. Robust Control and Model Misspecification. *Journal of Economic Theory* 128 (1):45–90.

Hansen, Lars Peter, John C. Heaton, and Nan Li. 2008. Consumption Strikes Back?: Measuring Long Run Risk. *Journal of Political Economy* .

Hansen, Lars Peter, Bálint Szőke, Lloyd S. Han, and Thomas J. Sargent. 2018. Twisted Probabilities, Uncertainty, and Prices. Unpublished manuscript.

Ilut, Cosmin L. and Martin Schneider. 2014. Ambiguous Business Cycles. *American Economic Review* 104 (8):2368–99.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini. 2006. Dynamic Variational Preferences. *Journal of Economic Theory* 128:4–44.

Newman, Charles M. and Barton W. Stuck. 1979. Chernoff Bounds for Discriminating between Two Markov Processes. *Stochastics* 2 (1-4):139–153.

Peng, Shige. 2004. *Nonlinear Expectations, Nonlinear Evaluations and Risk Measures.* Stochastic Methods in Finance: Lectures given at the C.I.M.E.-E.M.S. Summer School held in Bressanone/Brixen, Italy, July 6-12, 2003. Berlin, Heidelberg: Springer Berlin Heidelberg.

Sargent, Thomas J. 1999. *The Conquest of American Inflation.* Princeton, New Jersey: Princeton University Press.

Tallarini, Thomas D. 2000. Risk-Sensitive Real Business Cycles. *Journal of Monetary Economics* 45 (3):507–532.

Ulrich, Maxim. 2013. Inflation Ambiguity and the Term Structure of U.S. Government Bonds. *Journal of Monetary Economics* 60 (2):295 – 309.

Watson, James and Chris Holmes. 2016. Approximate Models and Robust Decisons. *Statistical Science* 31 (4):465–489.

Zha, Tao. 1999. Block recursion and structural vector autoregressions. *Journal of Econometrics* 90 (2):291–316.